Assessing the Risk to Deployed Personnel on Military Operations: a Discussion of Qualitative and Quantitative Approaches

Ian Graves and Neville J. Curtis
Defence Science & Technology Organisation (DSTO), AUSTRALIA
Ian.Graves@dsto.defence.gov.au Neville.Curtis@dsto.defence.gov.au

Abstract

The nature of modern conflicts means that previous definitions of deployments as "warlike" or "non war-like" are not helpful in identifying the potential risk to servicemen on operations and thus determination of appropriate pay and conditions needs to be re-examined.

We have developed an updated risk matrix based on physical combat factors, the physical environment, the operational setting and potential psychological harm. Defence subject matter experts identified and weighted specific harm factors within each of these groupings. For any new deployment, an overall risk figure can be determined and placed into one of five bands.

This was tested in two ways. First, a workshop of experienced servicemen was asked to place a set of previous operations into these bands using a top-down qualitative approach. This was then compared with a workshop of different servicemen who went through the quantitative scoring system described above (bottom-up). The results were mostly consistent though calibration was required to equate the two. In some cases, however, significant differences were met and these were for those cases at the "low end" of deployments where differences in opinion led to disagreement on the actual level of risk.

1. Introduction

Australian Defence Force (ADF) personnel are usually in receipt of benefits such as additional pay, tax concessions or leave entitlements when deployed on operations (usually overseas). Currently, the level of these benefits depends on whether the operation is classified as "warlike" or "non-warlike". An additional classification - "peacetime" is not subject to these conditions of service, though additional benefits may be granted. Defence decision makers however are challenged on how to assess these benefits, particularly as the traditional description and classification of operations as warlike and non-warlike has become difficult to apply. To this end Australia has been developing a rigorous tool that can be used to assess any upcoming operation and thus provide guidance on the level of benefits that should be considered. The tool is based on the notion of risk whether from physical combat, the physical environment, the operational setting or potential psychological harm. These terms thus provide a wider inspection of an operation rather than the rather black and white warlike versus non-warlike choice.

The essence of the proposed Defence Operational Risk Assessment (DORA) model is that a set of harm factors were identified and weighted to create a template. A new deployment would be judged through inserting the importance of each particular harm factor and using this to gain an overall figure for the operation. A scale can be developed that places each new operation into a category thus allowing a case to made to the Minister. This can be considered a bottom-up process arrived at by considering a set of specific aspects to the operation. Table 1 shows both the scale and the extent of the problem. Whereas previous conflicts such as World War 2 can be classified at the highest level, and off-shore border security perhaps at

the lower end, how might intermediate operations be treated? Similarly, what language should be used to replace "non-warlike" and where should the boundary be between those operations of the highest and next highest risk?

Table 1: The Defence Operational Risk Assessment (DORA) scale for deployments

Type of Operation	*Category*	*Illustrative Examples*
Warlike	5	World War 1 & 2, Vietnam
Hazardous (Non-Warlike)	4	*???*
	3	
	2	
	1	Border Security
Peacetime	N/A	Humanitarian Operations (i.e. Aceh earthquake/tsunami, Pakistan floods), Domestic Disaster Relief (i.e. Victorian bush fires, Queensland floods), Security Operations (i.e. Sydney Olympics)

While this semi-quantitative approach can be seen as applying a necessary aspect of transparency through a auditable process, the subject of off-shore benefits is very much an emotive topic with strong opinions from within Defence, other parts of the Public Service and from outside Government such as ex-Serviceman's associations. Thus, results from this bottom-up process need to be consistent with expectations and judgement of these other stakeholders. Our approach, in the first instance, has been to compare and contrast the "bottom-up" calculations with a largely subjective "top-down" appreciation by ADF personnel. In part this was done to attempt some validation of the bottom-up process but also to provide calibration (probably moderation) of the numbers (scores) used to classify the categories in Table 1. In addition, we have undertaken a number of case studies of specific operations to provide lessons in scoring of future actions. It is hoped that comparison of the top-down and bottom-up approaches will produce enough compelling evidence for formal take up of the model as part of the Nature of Service process.

In this paper, we will provide an overview of the process used to develop the harm factors, the engagement with Subject Matter Experts (SMEs) and how we compared the bottom-up and top-down approaches to refine the DORA model. In particular the ability to determine the quantitative numbers from the model and compare these with both pre- and post-deployment impressions has been extremely valuable both in formulating the scale and in providing confidence to what could otherwise be a sceptical audience. We now have details of about thirty deployments on the data base to allow some detailed examination of opportunities and threats for the model. It is also pertinent to note that at present Australia participates in around a dozen operations, many of which are of low public profile but which nevertheless still present risk of harm. Current and recent deployment details are shown in the publicly accessible website of the Department of Defence (http://www.defence.gov.au/op/index.htm). Operations span a range of types including intervention in Afghanistan, offshore evacuations of Australian citizens, peace observers, humanitarian relief and border protection. Some of these have been in place for many years.

2. The Current Process

The current process used in Australia is that the nature of service of a deployment is decided by the Minister of Defence, in consultation with the Prime Minister. The basis for the decision is the interpretation of the deployment against the definitions of "warlike" and "non-warlike"; the key differences are in the application of force, and the expectation of casualties. For example, for warlike operations the application of force is authorised and there is an expectation of casualties, whilst for non-warlike operations the application of force is limited to self-defence and there is a risk assigned to the tasks undertaken.

A shortcoming of the above approach is that serving personnel can effectively find themselves in a war zone that is considered to be non-warlike. A good example is personnel

deployed to the United Nations Truce Supervision Organisation (UNTSO), which is an organisation founded in 1948 for peacekeeping in the Middle East. On several occasions fighting has flared in the region, and personnel have been killed or injured; a total of fifty fatalities (including civilians) have occurred over the course of the mission.

Another shortcoming is the treatment of "peacetime" operations. Currently there is no definition of peacetime operations, other than to say any operation which is not warlike or non-warlike is considered peacetime. Effectively it is a political decision to classify an operation as a peacetime operation, even though the area that the personnel are deploying to may be far from peaceful and/or safe.

3. The Proposed Risk Based Approach

The pertinent features are that a set of "harm factors" were identified and weighted by a panel of experts. These harm factors have evolved since our initial examination (Graves and Curtis, 2005) and now (Graves and Curtis, 2011) include 10 items grouped into 4 (Table 2). Note that we used the term "matrix" on the basis of the AHP approach, and the term has become part of the language used by the Nature of Service Branch.

Table 2: Risk Matrices and Harm Factors

	Risk Matrices			
	Physical	**Health**	**Operational**	**Psychological**
Harm Factors	Opposing Forces	Communicable Diseases	Mission	Threat to Self
		Environmental Threats	Reliance on Allies	Exposure to Trauma
		Health Infrastructure	Operational Tempo	Operational Stressors

Four key steps were undertaken in development of this model:

- identification and engagement of SMEs from the ADF
- definition of the harm factors
- weightings of the harm factors
- application in practice and development of a scale

The notion of a semi-quantitative method has been used informally by the Nature of Service Branch for several years and for this reason it was comparatively straightforward to recruit genuine experts from the ADF, who both understood the area and the proposed model. In addition, the risk matrices comprise discrete elements that are readily identifiable within military units and thus allow for the to creation of discrete and distinct assessment groups (i.e. no SMEs were in more than one group). The initial harm factors from the original study were presented to the assessment groups and workshopped to the list shown above. Definitions were detailed in a narrative form. Associated with each harm factor are "points to consider" which are intended to assist the SME when assessing the factor for a particular operation. For example, for the Health Infrastructure harm factor the guidance is to consider the level of health services available from local civil and military facilities (i.e. not ADF or coalition), and the quality of blood supplies, food, water, shelter and sanitation.

Next a weighting system was introduced. The first phase was for the assessment groups to rank their specific harm factors in a workshop, using the analytical hierarchy process (Saaty, 1980). Consensus with the group led to acceptable consistency scores. Next the four risk matrices were ranked. For this each assessment group was only allowed to rank the other three areas thus eliminating potential bias. Again, discussion led to agreement on the relative weightings. To be used in practice, an operation would be judged against these harm factors, weightings applied and an overall figure obtained (e.g. Table 3). The scoring system involves SME groups assessing the exposure to the risk of harm (for the risks they are responsible for), on a 0 (no risk) to 10 scale, for the specific operation. For example, if the risk of "threat to self" was considered high then a rating of 7 or 8 might be appropriate. Each harm factor score

is multiplied by the appropriate weighting and the total sum determines the suggested operation type and category. For the example in Table 3 using indicative weightings, a score of 4.75 was obtained and this can be referred though its position on the ten point scale to one of the risk categories (Table 4), in this case "Hazardous – Category 3".

Table 3: Example of the weighting (i.e. DORA) system

Harm factor	*Indicative weighting*	*Example Scores*	*Weighted(DORA) score*
Opposing Forces	0.3	4	1.2
Communicable Diseases	0.05	5	0.25
Environmental Threats	0.05	6	0.3
Health Infrastructure	0.05	3	0.15
Mission	0.2	5	1.0
Reliance on Allies	0.1	6	0.6
Operational Tempo	0.1	3	0.3
Threat to Self	0.05	6	0.3
Exposure to Trauma	0.05	8	0.4
Operational Stressors	0.05	5	0.25
Overall	*1.0*		*4.75*

Whilst it can be argued that harm factors with a low weighting (i.e. less than 0.1) could be ignored it is our view that is it important to capture all the harm factors identified above. Further, the expectation is that the weightings will be reviewed every 2-3 years and it may be the case some of the weightings, especially for the psychology harm factors, are increased.

Table 4: Initial DORA Category Boundaries

Type of Operation	*Operational Category*	*Initial Boundaries*
Warlike	5	8.01 - 10
	4	6.01 – 8.0
Hazardous	3	4.01 – 6.0
	2	2.01 – 4.0
	1	0 – 2.0
Peacetime	N/A	N/A

Experience (see next section) showed that the model was practicable to apply and that an overall value was straightforward to obtain.

4. Comparison of the bottom-up and top-down scores

4a. The bottom-up analysis (quantitative using the DORA model)

A wide variety of different types of operations was considered to test the model. During the initial model development only past operations were considered, but a mixture of 15 past and 12 current operations were initially considered during the update of the model, with several other current operations subsequently added.

A workshop was held involving representatives from the four SME assessment groups to score the operations. For the current operations, a representative from the operational planners groups (operational and physical risks) provided an overview of each operation in turn and the SME from the other groups adding additional information or requested clarification. Once the discussion was complete scores were requested for each harm factor from the relevant assessment group. Usually the SME groups had already scored and operations prior to the workshop but were able to reconsider their scores in light of any new information. After the operation was scored the next one was discussed and scored. For the past operations the overview of the operation was provided by a representative from the Nature of Service Branch.

An advantage of the workshop approach used above is worth mentioning. Prior to a workshop the Psychology SME group assigned a very high score for “exposure to trauma” for a particular humanitarian operation on the basis that the deployed personnel would be “outside the wire” and extremely likely to encounter very distressing scenes. However, during the overview of the operation and subsequent discussions during the workshop it became clear that the deployed personnel would be “inside the wire” and, whilst dealing with locals who would be dealing with tragic circumstances, the “exposure to trauma” would be much lower.

Once all operations were scored the SME groups were given the opportunity to modify any scores to ensure consistency within each harm factor. The SME were also required to provide a rationale behind the score. For some harm factors, such as “opposing forces”, qualitative assessments for many operations already existed in an approved Military Threat Assessment (MTA), so if the SME were satisfied that the assessment was both current and relevant that could convert the qualitative assessment into a 0-10 score and use it as the rationale. Overall, a consensus figure was reached for each harm factor for each operation.

There are a number of important aspects of the SME groups used for the scoring which are worth expanding upon. Firstly, the SME were familiar with issues at hand, and usually the operations, since this was part of the normal day to day job. Secondly, these SME will be involved with the process on a continuing basis as the process called for regular review of the operations.

4b. The top-down analysis (qualitative using expert judgement)

The approach undertaken was to utilise the experience of a group of 20 ADF personnel, taken from the three services. Each member of this group had at least ten years experience in military service, all had been deployed, many to operations (both past and current) under consideration. None of these personnel were involved in the previous steps. In addition, none had prior experience of the DORA either in terms of the process or the aims of the new procedure.

A workshop was conducted where the group was requested to classify the operations solely based on their qualitative assessment of where the operations fitted into a ten point scale based on the category and subcategory shown in Table 4. For this, each category was split in two, for example, 2 HIGH and 2 LOW and the workshop attendees were asked to place a particular operation in one of these ten groups. Participants were also given the opportunity to rate the operation as “peacetime” if they did not believe it should be subject to special conditions. For the 7 current and 15 recent operations considered in the workshop, information was provided in the form of a one page summary of the operation, and attendees were allowed to seek further clarification or contribute any additional information. In most cases the information included the specified area, timeframe under consideration, mission, a summary of the MTA, force elements involved, rules of engagement and the degree of support from local authorities. The participants were then requested to indicate the classification of the operation on the form provided. After the final operation was considered, the participants were requested to check their assessments and make any adjustments to achieve consistency in their individual assessments.

Averaging of the scores allowed top-down rating using the ten point scale (ie 5 five categories each split into 2). At this stage there is no reason to suggest that this is the same as the ten point scale for the bottom-up analysis but is the basis for calibration between the two methods. One operation, based on the provision of security for the Sydney Olympics, was not scored since ~90% of workshop attendees considered it to be a peacetime operation.

5. Comparison and refinement of the scales

The results from the top-down and bottom-up (DORA) assessments are shown in Figure 1 for the remaining 14 previous and 7 current operations that were scored by both methods. The solid diagonal line represents identical scores, whilst the two parallel dashed lines bound a region which was considered to represent a reasonable agreement, noting that the scoring mechanism used for the top-down assessment was arbitrary. With the limited number of data points it is not possible to comment if there are differences between the scoring of current and previous operations, though the initial indications suggest not.

Four trends are of note. First, the agreement for the clearly warlike operations (categories 4 and 5) was very good. The only exception was for operation G where this was close to the boundary anyway. The second is that there are no clear break points between categories and that the scoring (by both methods) implies a continuous distribution. Third, there is the suggestion that the bottom-up method scores hazardous operations (up to about 4) higher than the top-down method (typically less than 3) with some outliers apparent (B, C and F). Last, there is scatter where the results of one method are compared to the other, with the spread of the top-down derived ratings being considerably more than for the DORA scores. For instance for a DORA score of about 6, the same operations scored in the range of about 2 to 7 (C and G). This represents a difference of 3 categories. In contrast, the spread of the DORA scores is much lower. The maximum difference between categories is only one (compare E and B) for the same top-down scores.

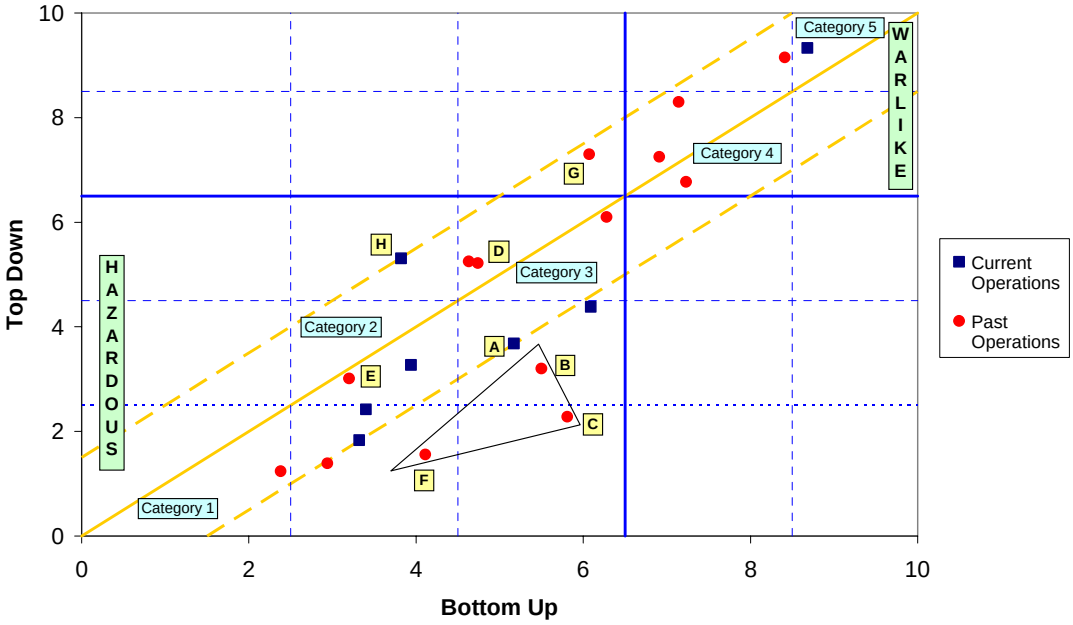

Figure 1: Comparison of top-down and bottom-up scores (categories separated by dotted lines refer to the DORA model)

The first refinement introduced by the comparison was in the boundaries between the hazardous and warlike categories. The model was susceptible to category creep on the basis of a single marginally higher harm factor score. For example, a hypothetical operation scoring all 6s except for a 7 for the lowest weighted harm factor would produce a score above 6 and hence fall in the warlike region. Three operations were in fact scored marginally over 6 by the bottom up method but only one of these was deemed warlike by the top-down assessment. Thus the boundary was raised to 6.5 thus implying that most of the categories would have to be at least 7 to rate as warlike. Second we were concerned that a peacetime operation would score above the category one region when assessed by the DORA model. The past operation

based on the provision of security for the Sydney Olympics, which is not shown in Figure 1, was the lowest scoring operation using the DORA model but still was given a score of nearly two (1.84). Thus the lower end of the hazardous scale was raised to 2.5. Other bands are also raised by 0.5 in keeping with these changes (Table 5).

Table 5: Revised DORA Category Boundaries

Type of Operation	*Operational Category*	*Modified Boundaries*
Warlike	5	8.51 - 10
	4	6.51 – 8.5
Hazardous	3	4.51 – 6.5
	2	2.51 – 4.5
	1	0 – 2.5
Peacetime	N/A	N/A

6. Lessons learned from the case studies (Operations A to H)

The operations labelled A to H ere examined in detail to see if they provided insights into the differences in top-down perception to the bottom-up analysis. Possible reasons for these differences, in particular, the poorer correlation of the three past operations shown in the shaded triangular area in Figure 1, are discussed below and summarised in Table 6.

Table 6: Features to consider in scoring of future operations

Feature	*Description*	*Issue*	*Implication*	*Example operations*
Personal experience	Scorers may have been deployed on previous phase of an operation or a similar action, or may have little exposure to the more hazardous zones	Non-typical conditions existed at the time	Top-down scoring too low	A, C
Long term operations	The operation may have run for many years with peaks and troughs of risk	Need to judge likely maximum risk	Top-down scoring too low	A
Job labelling	Deployment many been described as “military observers”, “peace keepers” or “humanitarian relief”	Words used may prejudge actual risk and imply an absence of threat	Top-down scoring too low	A, C, F
Armed/non-armed	Deployments may have been specifically non-armed	Assumption that this implies reduced risk	Top-down scoring too low	A, B
Few details	Little familiarity of the scorers to the type of operation	Wide variation in perception and scoring	Unreliable score	D, E
Short notice or duration operation	Not enough information available at the time	Pre-operational assessment may be difficult and needs to be revised later	Unreliable score	D, E
Routine operation	Operation seen to be similar to being in barracks or a training exercise	Operation may be seen as normal and not requiring any special treatment	Contentious score	E
Follow-on operation	The operation was post a higher risk activity	Tendency to maintain the higher level of risk despite a changed environment	Top-down scoring too high	H

Operation A. This refers to the UNTSO operation in the Middle East, briefly mentioned in Section 2. For this operation, the bottom-up scores were around 1.5 higher than the top-down assessments. As previously noted, this operation has been long-standing and had several spikes in activity, resulting in the loss of life, and the fact that the threat remained was made clear in the material presented. During the workshop there were several comments made about this operation by attendees, many of which suggested that the conditions encountered were similar to those in Australia. The only attendee who had deployed on that operation was

the only one to assess the operation as peacetime. Another issue which may have influenced the top-down assessment for operation A was that fact that the role of the deployed personnel was described as military observers rather than as warfighters. Throughout the development and assessment of the model several people expressed the opinion that unless deployed were armed and their rules of engagement allowed them to engage the elements outside of self defence then the operation should not be considered as warlike, regardless of the nature of the enemy force present.

Operation B. This refers to the evacuation of Australians and foreign nationals from an area in conflict. The scores were very different for the two methods. The deployed personnel were unarmed, with limited rules of engagement. Over 90% of the attendees during the top-down assessment considered the operation to be hazardous, but of these, half considered the operation to be category 2 whilst three-eighths considered it to be only category 1. Based on the DORA scores the operation fell in the middle of the category 3 region despite a low score for the "mission" harm factor which incorporates the rules of engagement and role of the personnel.

The role of the deployed personnel is a key factor in the current process when determining whether an operation is warlike or not, whereas for the DORA model the operational risk matrix is one of four matrices, with the rules of engagement captured in one of its three harm factors. There was clearly a disparity between the top-down and bottom-up assessments regarding the emphasis placed on the role of the deployed personnel. The dilemma can be expressed in terms of: is this operation more or less risky if the deployees are unarmed?

Operation C. This is a humanitarian operation to a region devastated by a natural disaster, where dissident elements were known to be present that were assessed to have the capacity and intent to harm ADF personnel (who were unarmed). This operation scored close to 6 using the DORA model, with every single harm factor scored as at least five, including several 8's. During the top-down assessment over a third of attendees considered the operation to be "peacetime" and of the rest 70% considered it to be a category 1 operation, resulting in a score which was around 2.5 lower than the DORA score. It is interesting to note that the only workshop attendee who deployed to this operation classified it as peacetime, noting that he/she was a member of the Air Force and may have had limited exposure to the full circumstances facing ADF personnel.

As previously noted the decision to classify an operation as "peacetime" is largely a political decision, and as in the case above, operations can be classified as peacetime in spite of clearly indentified threats to deployed personnel.

Operations D and E. While these two lie on the diagonal line in Figure 1 they provide some interesting insights to the top-down assessments. These are discussed below in the context of two past operations; a one week maritime operation in the Red Sea post the First Gulf War (labelled D in Figure 1), and the much publicised *Pong Su* incident in April 2003 when Special Forces elements intercepted and boarded the North Korean freighter (labelled E in Figure 1).

For these operations, a larger spread of assessments occupied during the top-down assessment, with a 59/41 split between warlike and hazardous for the first (D) and 6/53/41 between warlike, hazardous and peacetime for the second (E). Whilst there was a large spread of opinion for the above operations the top-down and bottom-up scores were very similar (within 0.5, Figure 1).

However, both these operations were amongst the least known operations considered and the amount of information available was limited. So, did the lower degree of familiarity and pertinent information impact on the top-down and, to a lesser degree, bottom-up assessments?

Furthermore, both of these operations were conducted over a very short time period, which should be irrelevant but may have influenced the decision making process during the top-down assessments. Operation D was reclassified post its completion, whilst operation E is classified as "peacetime". Finally, there were several comments made by workshop attendees that the activities undertaken during both operations could be considered as "routine".

Operations F and H. These operations achieved very similar DORA overall scores, and the also the four risk matrix scores were similar for both operations. Yet the top-down assessments were dramatically different, differing by nearly 4. Whilst over 90% of the top-down workshop attendees considered each operation as hazardous, around 80% considered F to be category 1 whilst a similar percentage considered H to be category 3. In this case a number of factors may have influenced the outcomes, including the fact that H was a stabilisation operation which was a follow-on to a warlike operation whilst F was a peace monitoring operation.

Whilst the impact on the DORA scores was most likely less significant, given the nature of the operations it is suggested that the assessments for both operations were lower than what could have been expected.

Operation G. This example was not a true discrepancy as the scores were similar. G itself placed in category 3 (and hence hazardous) region based on the bottom-up assessment but in the category 4 region (and hence warlike) based on the top-down assessment. The differences in raw score were however small and this shows the problem of trying to design a hard boundary between warlike and hazardous when there is a continuous distribution. Analysis of the operation was complicated by a number of factors, including that the ADF personnel were part of a third country deployment.

7. Discussion

A number of issues have been discussed in Section 6 which we believe have influenced the top-down assessments and, to a lesser extent, the DORA scoring. The common thread through these issues is whether or not the results have been biased.

In terms of the DORA scoring, there are sufficient mechanisms in the process, including that each harm factor score is based on a "group consensus" from the responsible SME group and that several of the harm factors are linked to the assessments from the corresponding MTA, to suggest that the DORA model is not susceptible to individual bias. Institutional biases could still occur however. This will be accentuated by some of the factors in Table 6 even if only subconsciously. For instance, the nature of the exercise, whereby a qualitative label was assigned to the operations, would have skewed the results somewhat simply based on the different perspectives on what constitutes a warlike operation, for example. Secondly, the different experiences of the attendees, including which service they were from, may have impacted on the classifications, especially where the attendee had been deployed on the operation being classified. Thirdly, during the discussion before the operation was classified some attendees were quite vocal on the merits of some operations, sometime based on personal experience, and how this impacted on the overall results is unknown.

The general sources of bias of the top-down assessment are as follows. Firstly, many of the previous operations already had a warlike/non-warlike classification. Next there were problems in differentiating "what actually happened" versus "what could have happened" including through personal experience. There are also cultural issues between different environments of the ADF that might lead to alternative perceptions of risks. There is still a history of expectation of "kinetic" casualties as being the prime reason for the classification and thus anything labelled as "peacekeeping" (for instance) is liable to be downplayed in risk.

It is probable that combination of these features led to the marked difference in interpretation of those operations that had similar DORA scores (the H to F and G to C groups in Figure 1). Thus F was scored lower owing to being labelled as peace monitoring as opposed to H being a stabilisation operation. There may also have been a bias in H which was a continuation of a previous higher categorisation operation but under a changed environment. Similarly, operations B and C were scored lower than G because of the labelling applied to the operations.

It is also pertinent to note that could be other biases that were not picked up in this process. For instance, for a follow-on operation such as peace keeping after peace making there may be an inclination to treat this as *reduced* risk rather than realising that it is actually a *changed* risk. Using the DORA method allows a proper examination of all the features to see if it is indeed an overall reduced risk or if some other harm factors rise in importance.

Finally, we do not favour a revision of classification once an operation has been completed, unless in exceptional circumstances (i.e. an unforseen spike in activity). The essence of the DORA model is that of risk, looking to see what could happen. Instead, the operations should be examined so that SMEs are better equipped to score similar operations in the future. Extensive data will also allow the assessment groups to examine the weightings and descriptions of the harm factors. However, regular reviews of ongoing operations are an important aspect of the process to unsure that changes in circumstances and potential changes in risks are incorporated in the decision making process.

8. Conclusions

Considerable progress has been made on this topic in the last few years. The arithmetic of the DORA modelling scores has now been checked against top-down perceptions and categories changed accordingly. Similarly, the use of case studies has allowed us to identify potential sources of bias in the use of the DORA model. With time we believe that the DORA model and its usage will evolve to become an established tool to provide the basis for a new categorisation of risk to deployees that is based on a transparent, credible and auditable process. A key benefit of the DORA model is that it follows a clear process which utilises subject matter experts and well defined factors, providing evidence to the decision makers, as opposed to the current process which is based on perception based opinion. We stress, however, that the tool should be used to advise, backed by narrative comment, and should only be one input to higher level Government decision-making. This was exemplified by operation G which lay close to the border of hazardous and warlike.

References

Graves I and Curtis NJ (2005). A Rigorous Method to Scale Overseas Deployments: Support to the Nature of Service Review. In: Puri V, Filippidis D, Retter P and Kelly J (eds) *Land Warfare Conference.* DSTO: Gold Coast, Queensland. pp 417-24

Graves I and Curtis NJ (2011). Development of a Model for Asssessing Exposure to Harm in ADF Operations for the Purpose of Determining the Conditions of Service Benefits for Deployed Personnel. *Australian Defence Operations Research Review 2010* 30-46

Saaty TL (1980). *The Analytical Hierarchy Process*. New York, McGraw Hill