To appear in *Military Operations Research*

WEI-WUV for Assessing Force Effectiveness: Managing Uncertainty with Info-Gap Theory

Yakov Ben-Haim
Yitzhak Moda'i Chair in Technology and Economics
Technion – Israel Institute of Technology
Haifa, Israel

ABSTRACT

The WEI-WUV algorithm purports to assess the effectiveness of a system of weapon systems based on a linear function of the quantities of each weapon type. The WEI-WUV algorithm has been widely criticized and indeed fallen into disrepute. The foremost limitations are (1) the need for subjective choice of numerical coefficients to represent complex and subtle situations, and (2) the linearity of the WEI/WUV function not accounting for synergy or competition between force elements. Is WEI-WUV really passé, out of date, of little or no use? In this paper we employ info-gap decision theory to model and manage the very real uncertainties that accompany a WEI-WUV assessment. We demonstrate how the impact of parameter and functional uncertainties can be assessed, and how the WEI-WUV algorithm can be used in supporting responsible decision making. We also explore the value of additional information when confronting dispute among experts about the values of WEI-WUV parameters.

INTRODUCTION

Military effectiveness is obtained, in part, by acquiring the proper combination of weapon systems, munitions, logistics, and other military capabilities. The choices involved are, ultimately, quantitative: how much of this, and how much of that, subject to constraints on money and time. The challenge facing quantitative decision support for force buildup is that military effectiveness is the result of many factors in addition to systems and munitions: morale, training, organization, intelligence, future adversary capabilities, goals of the political echelon, and more. In order for quantitative assessment to be useful in selecting weapon systems and munitions or in assessing force balance, it must account for the vast uncertainties that are inherent in the wider context of military effectiveness.

The WEI-WUV methodology provides an algorithm (whose mathematics we will specify later) for assessing the effectiveness of a system of weapon systems, with *N* different weapon categories (e.g. tanks, fighter aircraft, etc.) and *J* different specific weapon types in each category (e.g. M1 tanks, F16 fighters, etc.). Expert judgment is elicited to evaluate the Weighted Unit Value (WUV) of each weapon category, and the Weapon Effectiveness Index (WEI) of each weapon type in each category (Krepinevich and Watts, 2015, pp.142-148). The WEI-WUV algorithm assesses the overall effectiveness as a linear function of the quantities of each weapon type, weighted by the WEI and WUV coefficients. The WEI-WUV approach has a long history in the US Army. Raymond (1991) discusses detailed examples, and the US Army Concepts Analysis Agency (1991) cites US Army sources as far back as 1974. The WEI-WUV algorithm is easy to implement and precise: one gets a single number to assess military effectiveness of a complex configuration of systems. The challenge is two-fold: is the linear equation realistic and relevant or are non-linear functions needed, and are the coefficients correct?

It is plausible that, if we knew the broader context in detail, then we could calibrate the WEI-WUV algorithm for reliable assessment of military effectiveness in that context. The problem is that we cannot know the unknown future context in which the systems will operate. We face an **info-gap:** a deep disparity between what we *do know* and what we *need to know* in order to make a responsible decision. We thus cannot reliably calibrate the WEI-WUV algorithm (or any other algorithm). The WEI-WUV algorithm can err either in its calibration or in its mathematical form. What we can do, however, is to model and manage this uncertainty in the quantitative assessment of system configurations. This begins by distinguishing between two different questions. One question – that we cannot answer – is: how wrong is the quantitative algorithm for assessment of effectiveness? The other question – that we can answer – is: how much error in the quantitative algorithm can we tolerate without jeopardizing the quality of the assessment?

This second question addresses the robustness-to-uncertainty of the algorithm, and is the focus of this article, directed specifically at the WEI-WUV algorithm. An algorithm is highly robust to uncertainty if the algorithm can err substantially and yet the assessment is still reliable. Conversely, an algorithm has low robustness if the quality of the assessment is jeopardized when the algorithm errs even a little. The robustness question takes a different form when facing a choice between two (or more) alternatives: how much error in the quantitative algorithm can we tolerate without the decision switching from one alternative to another? In this paper we develop a methodology for assessing robustness to uncertainty, based on info-gap decision theory, and apply it to the WEI-WUV algorithm. We will consider robustness to uncertainty both in the parameters and in the mathematical form of the WEI-WUV algorithm.

It has long been recognized that quantitative decision support for assessing military effectiveness is both important and subject to limitations. Speight, Rowland and Keys (1997) bring "analytical methods to bear in order to: identify the factors most closely associated with past success; chart the manner in which these factors interact; and ascertain the hallmarks of successful historical campaigns" (p.31). They note, however, that the widely used WEI-WUV (Weapon Effectiveness Index-Weighted Unit Value) "does not take proper account of the complexity of battlefield interactions" (p.32). Lussier (U.S. Congressional Budget Office, 1988) notes that the WEI-WUV method

> "ignores many attributes of a military unit – such as quality and training of personnel, support equipment, logistic capability, and the interplay of various weapons – that can determine the outcome of a particular battle. Despite their importance, however, these factors often do not lend themselves to easy translation into numerical values." (p.16)
>
> "Static comparisons like those using the WEI/WUV method also ignore other decisive variables, such as strategy, maneuver, terrain, and combat attrition, that determine the conduct of war. … Finally, the WEI/WUV method assumes that the added benefit of additional weapons is linear – that is, more weapons of any kind continue to provide the same additional capability as the first such weapon." (p.17)

The WEI-WUV algorithm has rightly been criticized, and indeed fallen into disrepute in the US Army and elsewhere, due to very real limitations. The foremost limitations are (1) the need for subjective choice of numerical coefficients to represent complex and subtle situations, and (2) the linearity of the WEI/WUV function not accounting for synergy or competition between force elements. In this paper we will employ info-gap decision theory to model and manage the very real uncertainties that accompany a WEI-WUV assessment. We will demonstrate how parameter and functional uncertainties can be ameliorated. We also explore the value of additional information when confronting dispute among experts about the values of WEI-WUV parameters. Is WEI-WUV passé? Not necessarily, as we will see.

Info-gap robustness analyses have been performed in formulating and quantitatively evaluating a wide range of decisions in engineering, biological conservation, economics, medicine and other areas (see info-gap.com and Ben-Haim 2006, 2010). Info-gap robustness analyses have also been performed for various quantitative decision problems in security and military affairs. Moffitt, Stranlund and Field (2005) use info-gap robustness to develop inspection protocols for detecting terrorist activity. Sisso, Shima and Ben-Haim (2010) apply info-gap robustness in designing search strategies when probabilistic information about potential targets is highly uncertain. Davidovitch and Ben-Haim (2011) use info-gap theory in designing profiling strategies given severe uncertainty about criminal response to profiling. Info-gap theory has also been applied to the analysis of robustness to uncertainty in a range of qualitative decisions in military affairs (Ben-Haim 2014, 2015, 2016).

DECISION STABILITY AND ROBUSTNESS TO UNCERTAINTY

Our analysis of the WEI-WUV algorithm is based on the idea of robustness to uncertainty as evaluated with info-gap theory. An algorithm for assessing military effectiveness is highly robust to uncertainty if the decisions that it induces yield acceptable effectiveness despite substantial error in the parameters or structure of the algorithm. Equivalently, an algorithm is highly robust to uncertainty if widely different parameter values or functional forms would lead to the same decision. An algorithm that is robust to uncertainty will yield decisions that are stable in the sense that the same decision would be made even with quite different realizations of the uncertain quantities. A robust decision is desirable because it is consistent with the best available data and knowledge, as well as with a wide range of other possible data and knowledge. Stated differently, a robust decision is equivalent to a consensus decision of people whose data and knowledge is highly diverse.

It is important to emphasize that the decision-stability of a robust decision does not mean that the decision is rigid, unwavering or indifferent to existing knowledge. A robust decision is responsive to existing data and understanding, but is also consistent with a wide range of alternative knowledge that cannot be excluded on the basis of present understanding. This realism of the robustness is a result of the way the uncertainty model is formulated, based on info-gap decision theory, as we will explain.

The concept of robustness highlights two distinct aspects of a WEI-WUV algorithm to support decision making. First, the WEI-WUV assessment establishes a ranking of preference among available options. Second, we want to know if this preference ranking would stay the same even if our algorithm were substantially different in its parameter values or its functional form. That is, we want to know if the putative preference ranking is robust to uncertainty. That one option is preferred over another option by our WEI-WUV assessment does not imply that this putative preference is robust to uncertainty. Preference-ranking, and robustness-to-uncertainty of the preference-ranking, are two distinct attributes of a decision algorithm, as we will explain.

The putative WEI-WUV assessment – the numerical value of the force effectiveness – is a prediction in which we can have little faith. The point of the robustness analysis is to support the prioritization of force configurations without depending on explicit predictions of force effectiveness. When a specific force configuration has adequate predicted effectiveness and high robustness to uncertainty, then we have reason to adopt that configuration even though we cannot predict its actual force effectiveness. The info-gap robustness analysis can be seen as a response to Colin Gray's appeal (2010) for managing the unknown future:

> "You cannot know today what choices in defense planning you should make that will be judged correct in ten or 20 years' time. Why? Because one cannot know what is unknowable. Rather than accept a challenge that is impossible to meet, however, pick one that can be met well enough. Specifically, develop policy-makers, defense planners, and military executives so that they are intellectually equipped to find good enough solutions to the problems that emerge or even erupt unpredictably years from now. (p.6)
>
> "The gold standard for good enough defense planning is to get the biggest decisions correct enough so that one's successors will lament 'if only …' solely with regard to past errors that are distinctly survivable." (p.9)

Info-gap robustness analysis, when applied to the WEI-WUV algorithm, supports force planning without depending on prediction of the future. It provides a tool to assist in managing "problems that emerge … unpredictably years from now."

We will present an explicit methodology for evaluating the info-gap robustness to uncertainty. Three examples will demonstrate how robustness is evaluated and used in decision making based on the WEI-WUV algorithm.

FIRST EXAMPLE: UNCERTAIN MATHEMATICAL FORM

Introduction

Linear mathematical functions are widely used in quantitative assessment, despite their acknowledged deficiencies as mentioned earlier. The deficiencies of linear functions can be substantially ameliorated, and their advantages enjoyed, by managing their uncertainty as we now explain by examining the WEI-WUV algorithm.

The WEI-WUV algorithm assesses a system of systems, with N different weapon categories (e.g. tanks, attack helicopters, etc.) and J different specific weapon types in each category (e.g. M60A3 and M1 tanks, AH-1S and AH-64 helicopters, etc.). The concept of 'weapon categories' should be construed broadly to include logistical and support systems, intelligence capabilities, command and communication technologies etc., as well as explicit weapon systems. A specific force composition is specified by a matrix Q whose elements q_{nj} specify the number of weapons of type *j* in category *n*. The WEI-WUV algorithm assesses the effectiveness of a specific force composition with the function:

$$E(Q) = \sum_{n=1}^{N} v_n \sum_{j=1}^{J} q_{nj} w_{nj} \tag{1}$$

v_n is the Weighted Unit Value (WUV) for weapons category *n,* assessing the relative importance of this weapon category. w_{nj} is the Weighted Effectiveness Index (WEI) for weapon type *j* in category *n,* assessing the relative importance of the *j*th weapon type in category *n*.

It is clear, however, as many critics have acknowledged, that eq.(1) ignores important nonlinear effects. In other words, the function we should use to assess effectiveness of the force composition is:

$$E(Q, f) = \sum_{n=1}^{N} v_n \sum_{j=1}^{J} q_{nj} w_{nj} + f(Q) \tag{2}$$

where the function $f(Q)$ is unknown or at least highly uncertain. For example, $f(Q)$ might be a quadratic function of the weapon quantities q_{nj} expressing the diminishing marginal utility of individual weapon types, or expressing positive synergistic interactions, or negative competitive interactions, of different weapon types. Or $f(Q)$ might represent unknown higher-order polynomial relations. Or $f(Q)$ might represent discontinuous variation of effectiveness reflecting abrupt change due to highly nonlinear properties of the system of systems.

Info-Gap Model of Uncertainty

Let E_i and E_j denote the estimated effectiveness of two force compositions, Q_i and Q_j respectively, as assessed by eq.(1). Let $\bar{E}_{ij}$ denote the average estimated effectiveness: $\bar{E}_{ij} = (E_i + E_j)/2$. The uncertainty in this example is that we don't know the form of the non-linear function, $f(Q)$ in eq.(2), or its value relative to the estimated effectiveness values based on eq.(1) for any specific choice of the quantity matrix Q. This uncertainty is expressed as an unknown bound, h, on the absolute value of the ratio of $f(Q)$ to the average effectiveness $\bar{E}_{ij}$. That is:

$$\left| \frac{f(Q)}{\bar{E}_{ij}} \right| \le h \quad \text{and} \quad h \ge 0 \tag{3}$$

The left hand inequality in eq.(3) asserts that, while the form of the function $f(Q)$ is unknown, its fractional error is bounded by h for any specific choice of the quantity matrix Q. The right hand inequality asserts that the value of this bound is unknown. The two relations in eq.(3) assert that we have no knowledge that can restrict the form of the function $f(Q)$, and that we have no basis for fixing an upper bound on the relative error. We are acknowledging and accounting for very severe uncertainty in the functional form of the WEI-WUV algorithm. In order to decide whether or not the WEI-WUV recommendation is reliable, despite this uncertainty, we ask: how robust is the algorithm to this uncertainty? We will elaborate the answer shortly.

Let $U(h)$ denote the set of all functions, $f(Q)$, obeying the left hand condition in eq.(3). An info-gap model for uncertain fractional error in the function $f(Q)$, for comparing the specific compositions Q_i and Q_j, is the following unbounded family of nested sets of functions:

$$U(h) = \left\{ f(Q): \quad \left| \frac{f(Q)}{\bar{E}_{ij}} \right| \leq h \right\}, \quad h \geq 0 \tag{4}$$

$U(h)$ is a set of functions, $f(Q)$, defined on the domain of possible quantity matrices, Q. These sets become more inclusive as h increases. Thus h is called the 'horizon of uncertainty'. Because h is unknown, the info-gap model is an unbounded family of nested sets, $U(h)$, of possible realizations of the uncertain entity, the function $f(Q)$ in this case. The unboundedness of the info-gap model implies that there is no known worst case. Info-gap models can be defined in many different ways to represent different partial knowledge about the unknown entity. As in this info-gap model and others that we will encounter, bounded uncertainty in the shape of a function is readily represented. The robustness analysis based on an info-gap model of uncertainty is different from the usual sensitivity analysis in several ways. First, it considers unbounded variation rather than small or plausible estimated errors. Second, it deals with unbounded uncertainty in the shape of functions as well as in parameter values. Finally, an info-gap model of uncertainty can readily incorporate information about the uncertain variability, especially for functional uncertainty, as we now explain.

The info-gap model in eq.(4) can be modified to express additional information, if it is available to the analyst. For example, we might confidently assert that the effectiveness increases with increasing quantity of each weapon type, but that the marginal utility decreases. Thus two additional constraints would be added:

$$\frac{\partial f(Q)}{\partial q_{nj}} > 0 \quad \text{and} \quad \frac{\partial^2 f(Q)}{\partial q_{nj}^2} < 0 \tag{5}$$

The info-gap model for comparing compositions Q_i and Q_j now becomes:

$$U(h) = \left\{ f(Q): \quad \frac{\partial f(Q)}{\partial q_{nj}} > 0, \quad \frac{\partial^2 f(Q)}{\partial q_{nj}^2} < 0, \text{for all } n, j. \quad \left| \frac{f(Q)}{\bar{E}_{ij}} \right| \leq h \right\}, \quad h \geq 0 \tag{6}$$

Further additional information can be added, when available, such as second-derivative constraints representing unknown positive synergistic interactions, or unknown negative competitive interactions, between different weapon types. There is a vast array of mathematical forms for info-gap models of uncertainty, suitable to the non-probabilistic representation of uncertainty in many forms (Ben-Haim 2006, 2010).

Info-Gap Robustness

Define Q_i, Q_j, E_i, E_j and $\bar{E}_{ij}$ as before. We would like to use the algorithm in eq.(2) for comparing configurations Q_i and Q_j, but we don't know the form of the function $f(Q)$ so we can't evaluate $E(Q_i, f)$ or $E(Q_j, f)$. If we use eq.(1) instead, the robustness question (discussed earlier) is: how wrong can eq.(1) be, without unduly jeopardizing the quality of the decision?

Suppose that eq.(1) indicates that Q_i is more effective, and hence putatively preferred over Q_j. The robustness question is: how much error in eq.(1) can we tolerate without altering this preference ranking? That is, how large could $f(Q)$ be, relative to the linear term, without changing the decision based on ignoring $f(Q)$ altogether and using eq.(1)? If large error can be tolerated without altering the decision, then we are confident in using eq.(1) even though we know it is wrong. We won't be able to predict the effectiveness, but we can confidently prioritize the force configurations.

Implementation of the analysis of robustness proceeds as follows. Define the effectiveness-margin, Δ, of Q_i over Q_j as $\Delta = E(Q_i, f) - E(Q_j, f)$. For any required effectiveness-margin, the robustness, denoted $\hat{h}(\Delta)$, is the greatest tolerable absolute fractional error of the unknown function $f(Q)$. Large robustness implies that the algorithm in eq.(1) can be confidently used despite ignorance of the non-linear interactions. In contrast, small robustness implies that we cannot have confidence in eq.(1). The robustness function is formally defined as:

$$\hat{h}(\Delta) = \max\left\{ h : \left(\min_{f(Q) \in U(h)} [E(Q_i, f) - E(Q_j, f)] \right) \ \geq \ \Delta \right\} \tag{7}$$

Reading this expression from left to right: The robustness, $\hat{h}(\Delta)$, for comparing force compositions Q_i and Q_j, is the greatest horizon of uncertainty, h, up to which the lowest margin of effectiveness, $E(Q_i, f) - E(Q_j, f)$, exceeds the required critical value Δ, for all realizations of the unknown function, $f(Q)$, in the uncertainty set $U(h)$. The inner minimum is the worst case at horizon of uncertainty h, and the robustness, $\hat{h}(\Delta)$, is the greatest h at which the worst case still satisfies the required effectiveness margin as specified by Δ. We don't know the value of the horizon of uncertainty, h, so we can't calculate the true worst case. Nonetheless, we can calculate the greatest tolerable uncertainty, which is $\hat{h}(\Delta)$.

The robustness function is readily derived for the info-gap model of eq.(4). As before, define $E_i = E(Q_i)$ in eq.(1) and similarly for E_j. Thus eq.(2) becomes $E(Q_i, f) = E_i + f(Q_i)$ and similarly for j. Let $m(h)$ denote the inner minimum in eq.(7), which occurs when $f(Q)$ takes the following values: $f(Q_i) = -h\overline{E}_{ij}$ and $f(Q_j) = +h\overline{E}_{ij}$ (we are assuming that $\overline{E}_{ij}$ is positive). Thus, because $m(h)$ must be no less than Δ, we find the following expression for the robustness:

$$m(h) = E_i - E_j - 2h\overline{E}_{ij} \geq \Delta \quad \Rightarrow \quad \hat{h}(\Delta) = \frac{E_i - E_j - \Delta}{E_i + E_j} \tag{8}$$

or zero if this expression for $\hat{h}(\Delta)$ is negative (recall that $\overline{E}_{ij} = (E_i + E_j)/2$).

Data

The Weapon Effectiveness Indices, w_{nj}, and the Weighted Unit Values, v_n, are taken from the report of the U.S. Congressional Budget Office (1988, p.15). These data are also used by Krepinevich and Watts (2015, p.143). The CBO WUV and WEI values are:

$$v = (94,\ 109,\ 56,\ 71,\ 73,\ 99,\ 55,\ 30,\ 4) \tag{9}$$

$$W = \begin{pmatrix} 1.11 & 1.31 & 0 \\ 1.00 & 1.77 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0.79 & 0.69 & 0.2 \\ 1.02 & 0.98 & 1.16 \\ 0.97 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1.77 & 0 \end{pmatrix} \tag{10}$$

We will consider 3 force compositions: a standard composition Q_1, an innovative composition Q_2, and a conservative composition Q_3. The standard force composition, Q_1, is the one specified in the U.S. CBO report. The innovative and conservative force compositions, Q_2 and Q_3, modify the composition of three weapon categories: tanks (category 1), attack helicopters (category 2), and anti-tank missiles (category 5). The higher WEI value for the M1 tank, for the AH-64 helicopter, and for the TOW and the Dragon anti-tank missiles, suggest that these weapon types can be treated as innovative options compared with the alternatives (M60A3 tank, AH-1S helicopter, and LAW anti-tank launcher). Force compositions for these categories are shown in tables 1, 2 and 3.

Force composition	$q_{1,1}$	$q_{1,2}$	$q_{1,3}$
	M60A3	M1	-
Standard (1)	150	150	0
Innovative (2)	0	300	0
Conservative (3)	300	0	0

Table 1. Force compositions for tanks.

Force composition	$q_{2,1}$	$q_{2,2}$	$q_{2,3}$
	AH-1S	AH-64	-
Standard (1)	21	18	0
Innovative (2)	0	39	0
Conservative (3)	39	0	0

Table 2. Force compositions for attack helicopters.

Force composition	$q_{5,1}$	$q_{5,2}$	$q_{5,3}$
	TOW	Dragon	LAW
Standard (1)	150	240	300
Innovative (2)	540	150	0
Conservative (3)	0	0	690

Table 3. Force compositions for anti-tank missiles.

Results

We illustrate the info-gap robustness analysis by considering these 3 force compositions, Q_1, Q_2 and Q_3. Based on eq.(1), the putative effectiveness values, E_1, E_2 and E_3, for the three options are 1.22×10^5, 1.40×10^5 and 1.03×10^5 respectively. The average effectiveness $\bar{E}_{ij}$, the effectiveness-margin $E_i - E_j$, and their ratio, are shown in table 4 for the three comparisons of these three compositions. The effectiveness margin of the innovative option (2) over the standard option (1) is the smallest of the three comparisons, and it is also the smallest fraction, 0.138, of the average effectiveness. The standard (1) compared with the conservative option (3) is the next, while the innovative (2) compared with the conservative option (3) is the most definitive comparison, showing a 30.9% effectiveness margin relative to the average effectiveness of these two options. While a 30.9% advantage of the innovative over the conservative force composition may (or may not) seem substantial based on contextual understanding, how robust is this to error in the functional form of the WEI-WUV algorithm in eq.(1)?

i	j	$\bar{E}_{ij}$	$E_i - E_j$	$(E_i - E_j)/\bar{E}_{ij}$
2	1	1.31×10^5	1.82×10^4	0.138
1	3	1.13×10^5	1.94×10^4	0.172
2	3	1.22×10^5	3.75×10^4	0.309

Table 4. Effectiveness margins for example 1 with 3 compositions: standard (1), innovative (2) and conservative (3).

The robustness functions for the comparisons of the three force compositions are plotted in fig. 1, whose details we now explain. We will stress the concepts of trade off and zeroing, and explain how they are used in responsible decision making.

The solid curve in fig. 1 is the robustness, $\hat{h}(\Delta)$, vs. the required effectiveness margin, Δ, for comparing the innovative force composition, Q_2, against the conservative composition Q_3. The negative slope of the curve reflects the **trade off** between the required margin of effectiveness and the robustness-against-uncertainty. Greater margin Δ (which is desirable) can be required but only by reducing the robustness, $\hat{h}(\Delta)$ (which is not desirable). This is a universal and unavoidable trade off (sometimes called the pessimist's theorem): a more demanding outcome (greater required effectiveness margin Δ) is more vulnerable to failure (hence the robustness, $\hat{h}(\Delta)$, is lower).

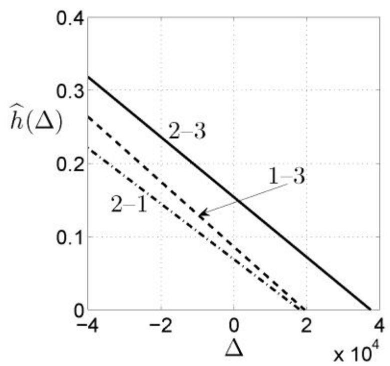

Figure 1. Robustness functions for example 1.

The solid robustness curve in fig. 1 reaches the horizontal axis when Δ precisely equals the estimated value of the effectiveness margin, 3.75×10^4. This means that the estimated margin has no robustness against uncertainty. This universal property of **zeroing** means that best-estimated values are not a good basis for decision making because they have zero robustness against uncertainty. This is not surprising to the skeptic who hastens to point out that the 'best estimate' is based on eq.(1) which is known to be wrong.

Combining the trade off and zeroing properties of the robustness curve enables the responsible use of eq.(1). On the one hand the analyst treats the estimated margin of effectiveness of the innovative over the conservative option, 3.75×10^4, with considerable caution because its robustness is zero. On the other hand, one notes from the solid curve in fig. 1 that the robustness is 0.15 when the required effectiveness margin, Δ, is zero: $\hat{h}(0)=0.15$. This means that Q_2 is preferred over Q_3 provided that the unknown nonlinear function $f(Q)$ is less in absolute magnitude than 15% of the average effectiveness of these two compositions. We don't know the shape of the function $f(Q)$, and extensive data and knowledge would be required in order to specify it. However, much less extensive contextual understanding might be sufficient in order to assert that $f(Q)$ is (or is not) plausibly greater than 15% of the linear model. The analyst might make the judgment that, for these very different force compositions, the nonlinear interactions of weapon systems could plausibly be very important, and that 15% robustness is in fact rather low. This would imply that one would not have confidence in preferring Q_2 over Q_3 based on the putative WEI-WUV recommendation. On the other hand, the analyst might make the judgment that, despite the names of these two compositions, they are in fact very much in the same 'ballpark' and that linear interactions capture most of the relevant factors. In that case, a 15% robustness of the preference for Q_2 over Q_3 is quite definitive.

Similar conclusions result from the comparison of the standard vs the conservative compositions (1–3, dashed line in fig. 1) and for the innovative vs the standard option (2–1, dot-dashed line). Both curves display the trade off between robustness and required effectiveness margin, and the zeroing property that the estimated effectiveness margin has no robustness against error in eq.(1). Q_1 is putatively preferred over Q_3 with an estimated effectiveness margin of 1.94×10^4. Likewise, Q_2 is putatively preferred over Q_1 with an estimated effectiveness margin of 1.82×10^4. However, the robustness to uncertainty of these estimated margins is zero, so these putative preferences must be treated with great caution. Furthermore, the robustness in general is low for these comparisons. For instance, at zero margin, the robustness values are $\hat{h}(0)=0.086$ and $\hat{h}(0)=0.069$ for comparisons 1–3 and 2–1, respectively. The analyst may make the judgment that 8.6% or 6.9% robustness is too low to warrant support for the putative WEI-WUV preference of Q_1 over Q_3 or for Q_2 over Q_1, respectively.

SECOND EXAMPLE: PARAMETRIC UNCERTAINTY

Introduction

In the previous section we considered uncertainty in the mathematical form of the WEI-WUV algorithm for assessing military effectiveness and for prioritizing alternative force compositions. The example considered the linear WEI-WUV algorithm, eq.(1), to which an unknown nonlinear function was appended, resulting in eq.(2). We now return to the strictly linear form, eq.(1), and consider uncertainty in the values of the parameters: the WUV's v_n assess the relative importance of the *n*th weapon category, and the WEI's w_{nj} assess the relative importance of the *j*th weapon type in category *n*.

Judgment is needed to choose the values of these WEI-WUV parameters, and considerable uncertainty can surround the choices of the parameter values. Why, for instance, is the Weapon Effectiveness Index of the AH-64 attack helicopter 1.77 while the WEI index for the AH-1S is 1.00? Why is the Weighted Unit Value of the Vulcan air defense system 56, while the WUV of the Bradley fighting vehicle is 71 (U.S. CBO, 1988, p.15)? Different professional analysts might plausibly choose different values. Likewise, the same experts might choose different parameter values when considering different conflict scenarios. We will demonstrate that the info-gap robustness analysis, used in the previous section to evaluate the impact of functional uncertainty, is directly applicable to this parameter-uncertainty analysis. (The two analyses can be combined, though we will not pursue that extension here.)

We will compare the same three force compositions considered earlier: the standard composition Q_1, the innovative composition Q_2, and the conservative composition Q_3. We will use the linear WEI-WUV algorithm in eq.(1), with the parameter values recommended in the study by Lussier (U.S. CBO, 1988, p.15). However, we are concerned that different analysts might well provide very different (though not negative) values of the WEI-WUV coefficients. More specifically, consider two force compositions, Q_i and Q_j, where the putative effectiveness, according to eq.(1) with the CBO WEI-WUV values, is greater for Q_i than for Q_j, that is: $E_i - E_j$ exceeds a required value of the effectiveness margin Δ. The robustness question, in the context of parameter uncertainty, is: by what fraction can the WEI-WUV coefficients change without reducing the effectiveness margin below the required value Δ? In other words, how robust to parameter uncertainty is the preference for Q_i over Q_j?

Formulation and evaluation of the robustness

Let $\tilde{v}$ and $\tilde{W}$ denote the WEI-WUV values in the CBO report (U.S. CBO, 1988, p.15), and let v and W denote uncertain true values. The following info-gap model represents unknown and unbounded fractional error of the nominal values, while also requiring all values to be non-negative:

$$U(h) = \left\{ v, W: \quad v_n \geq 0, \left| \frac{v_n - \tilde{v}_n}{\tilde{v}_n} \right| \leq h, \ \forall n. \ \ w_{nj} \geq 0, \left| \frac{w_{nj} - \tilde{w}_{nj}}{\tilde{w}_{nj}} \right| \leq h, \ \forall\, j, n \right\}, \quad h \geq 0 \qquad (11)$$

Like all info-gap models of uncertainty, this is an unbounded family of nested sets for which a worst case is unknown. We note that the same horizon of uncertainty, h, applies to all of the fractional errors in this info-gap model. That is, at horizon of uncertainty h, the fractional errors of all the parameters, v_n and w_{nj}, are bounded by the same value of h. The value of h, however, is unknown.

The robustness is defined analogously to eq.(7), though with respect to parameter uncertainty. The robustness for comparing force compositions Q_i and Q_j is the greatest horizon of uncertainty, h, up to which the margin of effectiveness, $E(Q_i,v,W)-E(Q_j,v,W)$ based on eq.(1), exceeds the required value Δ, for all realizations of the uncertain parameters, v and W, in the uncertainty set $U(h)$. The robustness function is formally defined as:

$$\hat{h}(\Delta)=\max\left\{h:\left(\min_{v,W\in U(h)}[E(Q_i,v,W)-E(Q_j,v,W)]\right)\geq\Delta\right\} \quad (12)$$

Let $m(h)$ denote the inner minimum in eq.(12). The uncertainty sets $U(h)$ become more inclusive as the horizon of uncertainty, h, increases. Thus $m(h)$, which is a minimum on $U(h)$, decreases as h increases. From eq.(12) we see that the robustness, $\hat{h}(\Delta)$, is the greatest value of h at which $m(h)\geq\Delta$. From this we see that a plot of h vs. $m(h)$ is equivalent to a plot of $\hat{h}(\Delta)$ vs. Δ. In other words, $m(h)$ is the inverse function of $\hat{h}(\Delta)$. In short, it is sufficient to evaluate $m(h)$, which we do as follows.

Using eq.(1) we can write:

$$E(Q_i,v,W)-E(Q_j,v,W)=\sum_{n=1}^{N}v_n\sum_{j=1}^{J}\left(q_{nj}^{(i)}-q_{nj}^{(j)}\right)w_{nj} \quad (13)$$

Define a truncation function: $x^+=x$ if x is positive, and $x^+=0$ otherwise.

It is evident from the info-gap model of eq.(11) that the inner minimum in eq.(12) is obtained if each w_{nj} is chosen as $(1+h)\tilde{w}_{nj}$ if the corresponding $q_{nj}^{(i)}-q_{nj}^{(j)}$ is negative, and chosen as $(1-h)^+\tilde{w}_{nj}$ otherwise.

Similarly, it is evident from the info-gap model of eq.(11) that the inner minimum in eq.(12) is obtained if each v_n is chosen as $(1+h)\tilde{v}_n$ if the corresponding $\sum_{j=1}^{J}\left(q_{nj}^{(i)}-q_{nj}^{(j)}\right)w_{nj}$ is negative, and chosen as $(1-h)^+\tilde{v}_n$ otherwise.

The inner minimum, $m(h)$, is obtained from eq.(13) with these choices of w_{nj} and v_n. We now discuss the numerical results.

Results

The putative effectiveness values and margins for the 3 compositions are the same as in the previous example (see table 4). As before, the innovative-conservative comparison (2-3) has the greatest estimated effectiveness margin, at 30.9% relative to the average effectiveness of these two options. How robust to uncertainty in the WEI-WUV parameters is this comparison? Robustness curves are shown in fig. 2 for the three comparisons.

The zeroing and trade off properties of robustness curves, discussed earlier in connection with fig. 1, hold in the present case as well.

The zeroing property is that putative estimates have no robustness to uncertainty in the data upon which the estimates are based. This is demonstrated by the horizontal intercepts of the 3 robustness curves in fig. 2, at which the robustness is zero. These intercepts are the same as in fig. 1, because in both figures they correspond to the putative estimates of the effectiveness margins presented in table 4.

The robustness to uncertainty, $\hat{h}(\Delta)$, trades off against the required margin of effectiveness, Δ: one can require greater margin only at the cost of accepting lower robustness to uncertainty. The innovative-conservative comparison (2–3, solid curve) has the greatest putative effectiveness margin (horizontal intercept furthest to the right), but also has the strongest trade off between robustness and effectiveness margin (lowest slope). The "cost of robustness" is greatest for the solid curve in fig. 2, meaning that a given positive increment of robustness, $\hat{h}(\Delta)$, is obtained with the largest decrement of required effectiveness margin, Δ. Interestingly, the robustness at zero margin of effectiveness, $\hat{h}(0)$, is very nearly the same for all three comparisons, equaling about 0.22. This means that, in each of the three comparisons, the preference ranking is stable for all variations of the WEI-WUV coefficients up to about 22% of their nominal values. Larger parameter variations can lead to reversal of the sign of the effectiveness margin and hence reversal of the preference ranking of the corresponding force compositions.

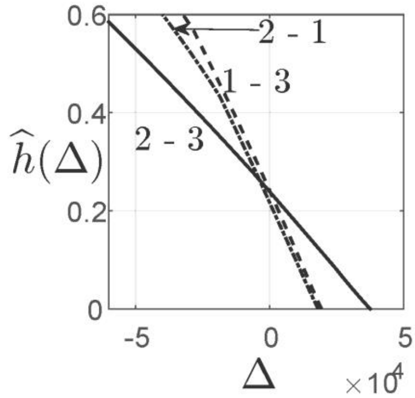

Figure 2. Robustness functions for example 2.

The nominal preference ranking of the three compositions is based on evaluating the effectiveness margins using the nominal estimates of the WEI-WUV coefficients. These nominal preferences are that the innovative composition (2) is preferred over the standard composition (1), which in turn is preferred over the conservative composition (3). In light of the robustness analysis, these preferences will be the same among analysts who agree on the values of the WEI-WUV coefficients within plus or minus 22%.

A decision maker who judges, based on experience or contextual understanding or consultation with diverse experts (see the next example), that the coefficients could vary more than 22% will not be confident in this putative preference ranking. If, for example, experience shows that WEI-WUV coefficients could vary by tens of percent, then fig. 2 shows that the margins of effectiveness can be strongly negative for each of the comparisons, indicating the possibility of reversal of the preferences. For example, if parameters can vary up to 40% , then the dashed curve at $\hat{h}(\Delta) = 0.4$ shows that the margin of effectiveness for the standard over the conservative composition can be negative and as small as $E_1 - E_3 = -1.3 \times 10^4$. This indicates possible strong preference for the conservative over the standard option. The comparison of the innovative and conservative options is even more dramatic, as seen from the solid curve in fig. 2. At 40% robustness the effectiveness margin for the innovative composition can be as negative as $E_2 - E_3 = -2.6 \times 10^4$, indicating a potential for strong preference for the conservative over the innovative option.

A negative value of Δ, the required margin of effectiveness of configuration *i* over configuration *j,* would usually imply that *j* is preferred. For instance, the innovative (*i*=2) vs. the standard (*j*=1) comparison has $\Delta = -1.3 \times 10^4$ at robustness of 0.4, suggesting that the standard option should be preferred. However, exogenous considerations – economic, organizational or political – may militate against returning to the standard configuration unless it is very substantially more effective.

THIRD EXAMPLE: EXPERTS AND THEIR TRIBULATIONS

The knowledge and judgment of experts (and sometimes of non-experts) is used in choosing the parameter values of the WEI-WUV algorithm. Experts will often disagree in their judgments, indicating the need for collecting diverse opinions in attempting to converge on realistic parameter values. The degree of dispute among experts can be reduced by research and by various methods for eliciting expert judgment. What extent of agreement is needed in order to support responsible decision making? In this section we use the info-gap robustness function to address the question: how much dispute among experts is consistent with reliable use of the WEI-WUV algorithm?

We suppose that our expert recommendations of WEI-WUV parameter values have been distilled as follows. Each WUV parameter, v_n, has been assigned an estimated value $\tilde{v}_n$ with an associated positive error estimate, s_n, that assesses the degree of dispute about $\tilde{v}_n$. Recognizing that experts can change their minds and that other experts may have yet other judgments of the value of v_n, we acknowledge that the true of value v_n may deviate from $\tilde{v}_n$ by more than s_n. Similarly, each WEI parameter, w_{nj}, has been assigned an estimate, $\tilde{w}_{nj}$, and a positive error estimate t_{nj}, recognizing that the true value w_{nj} may deviate from $\tilde{w}_{nj}$ by more than t_{nj}. We assume, as before, that the v_n's and the w_{nj}'s are non-negative. The question we address is: how much must the error estimates, s_n and t_{nj}, be reduced (by legitimately reducing dispute among the experts) in order for the WEI-WUV algorithm to be a reliable basis for decision making? (The actual dispute-reduction process may also lead to revised values of the estimates, $\tilde{v}_n$ and $\tilde{w}_{nj}$, but we won't consider their modification in this example.)

We continue the example of the previous section, only modifying the info-gap model of eq.(11) as follows:

$$U(h)=\left\{v,W: \ v_n \ge 0, \left|\frac{v_n-\tilde{v}_n}{s_n}\right| \le h, \forall n. \ w_{nj}\ge 0, \left|\frac{w_{nj}-\tilde{w}_{nj}}{t_{nj}}\right| \le h, \forall j,n\right\}, \quad h\ge 0 \qquad (14)$$

The only difference from eq.(11) is that the fractional errors of the parameters are defined with respect to their error estimates rather than with respect to their estimated values. Eq.(14) is a generalization of eq.(11) because s_n and t_{nj} could be chosen as $\tilde{v}_n$ and $\tilde{w}_{nj}$, respectively, thus reverting to eq.(11).

We consider the same three force compositions, Q_1 (standard), Q_2 (innovative) and Q_3 (conservative). The estimated WEI-WUV values, $\tilde{v}_n$ and $\tilde{w}_{nj}$, are taken from the report of the U.S. Congressional Budget Office (1988, p.15) as before.

The definition of the robustness function in eq.(12) is still valid, where now $U(h)$ is the info-gap model of eq.(14). Likewise, eq.(13) is the effectiveness margin for comparing option i with option j.

In analogy to the discussion following eq.(13), the info-gap model of eq.(14) implies that the inner minimum in eq.(12) is obtained if each w_{nj} is chosen as $\tilde{w}_{nj}+t_{nj}h$ if the corresponding $q^{(i)}_{nj}-q^{(j)}_{nj}$ is negative, and chosen as $\left(\tilde{w}_{nj}-t_{nj}h\right)^+$ otherwise.

Similarly, it is evident from the info-gap model of eq.(14) that the inner minimum in eq.(12) is obtained if each v_n is chosen as $\tilde{v}_n+s_nh$ if the corresponding $\sum_{j=1}^{J}\left(q^{(i)}_{nj}-q^{(j)}_{nj}\right)w_{nj}$ is negative, and chosen as $\left(\tilde{v}_n-s_nh\right)^+$ otherwise.

The inner minimum in eq.(12), $m(h)$, is obtained from eq.(13) with these choices of w_{nj} and v_n. We now discuss numerical results.

The results in the previous section, based on the info-gap model in eq.(11), can be understood in terms of the info-gap model of eq.(14) by defining the uncertainty estimates as $s_n = \tilde{v}_n$ for all n, and $t_{nj}=\tilde{w}_{nj}$ for all n and j. This definition of the uncertainty estimates means, in the context of the current example, that the expert dispute on each WEI-WUV parameter is on the order of 100% of the estimated value of that parameter. In other words, truly reliable estimates may deviate from the known estimated values by 100% or more. And this implies that the choice between any two options, Q_i vs. Q_j, requires robustness to uncertainty of several hundred percent: a value of $\hat{h}(\Delta)$ in the range of perhaps 3 to 4. The robustness curves in fig. 2 show that positive efficiency margins, Δ, have much lower robustness for all comparisons. In other words, the expert dispute must be substantially reduced in order to support confident choice based on the WEI-WUV algorithm.

By how much must the expert dispute be reduced, as expressed by smaller uncertainty estimates s_n and w_{nj}, in order to achieve acceptable robustness and thereby achieve reliable WEI-WUV assessment? We now illustrate the answer.

Let $s_n=\varepsilon\tilde{v}_n$ for all n and $t_{nj}=\varepsilon\tilde{w}_{nj}$ for all n and j for some positive choice of ε. This means that the expert dispute over the WEI-WUV parameter values is about a fraction ε of each estimated value. (One can consider different degrees of dispute for different parameters if one has relevant information, but, for simplicity, we don't consider that.)

Fig. 3 shows robustness curves for $\varepsilon = 0.5$ (left frame) and $\varepsilon = 0.25$ (right frame). When $\varepsilon = 0.5$ the range of expert dispute is about $\pm 50\%$ for each parameter value, implying the need for robustness several times this: $\hat{h}(\Delta)$ in the range of 1.5 to 2.0. From the left frame in fig. 3 we see positive efficiency margins only for much lower robustness. Evidently, $\pm 50\%$ dispute is too large to warrant a confident choice based on the WEI-WUV algorithm.

The right hand frame in fig. 3 shows robustness curves for $\pm 25\%$ expert dispute, implying that reliable use of the WEI-WUV recommendation requires a robustness value on the order of 0.75 to 1.00. We see positive efficiency margins with robustness in this range: $\hat{h}(\Delta)$ nearly equal to 1. For example, the innovative option, Q_2, has positive efficiency margin over the conservative option Q_3, with robustness nearly 4 times the uncertainty estimate. This might reasonably be judged substantial robustness to uncertainty in the parameters, and this preference – based on the WEI-WUV analysis – might confidently be accepted. A similar judgment applies to the comparison of the standard vs the conservative options (1 – 3), and the standard vs the innovative options (1 – 2), though with lower efficiency margins than for the innovative-conservation comparison (2 – 3). We *don't* have confidence in the putative predictions of force effectiveness, but we *do* have confidence in the prioritizations of these force configurations.

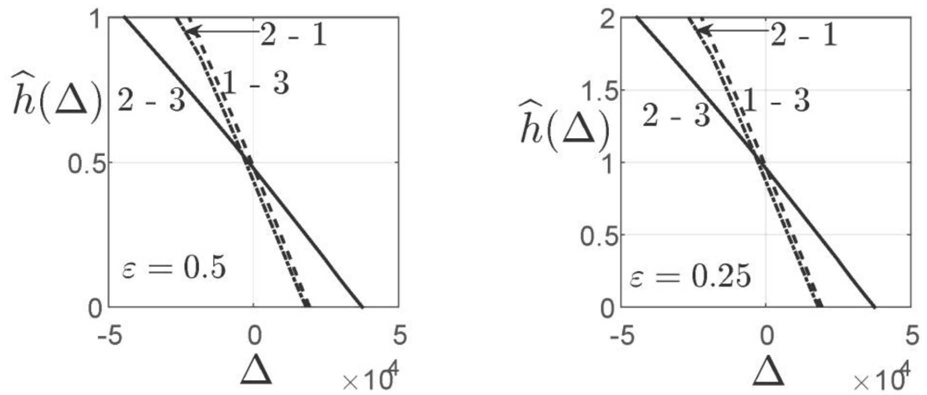

Figure 3. Robustness functions for example 3.

DISCUSSION

Robustness analysis adds a dimension of responsible skepticism to the interpretation and use of quantitative assessments of military effectiveness. We have focused on the WEI-WUV algorithm, though info-gap robustness analysis is relevant to other methods as well. Preferences among force buildup options, based on best estimates of quantitative measures of effectiveness, are the starting point, not the conclusion, of the analysis. After establishing the nominal preferences, we ask the robustness question: how much error (in parameter values or in mathematical form) can we tolerate without altering the preference ranking? As illustrated in the first two examples, if the answer is that the ranking would remain unchanged even with great error, then the analyst can be confident in the preference ranking despite substantial uncertainty. On the other hand, if small changes in parameters or functional forms could alter the preference ranking, then additional analysis, information, or options are needed before a confident decision can be made. The third example illustrated how the info-gap robustness analysis can be used to evaluate how much additional information or analysis might be needed.

The WEI-WUV algorithm is conventionally used to predict the effectiveness of alternative force configurations, leading to prioritization of the configurations based on these predictions. We have shown that these predictions are unreliable, which is the main limitation of the conventional use of WEI-WUV. The robustness analysis of the WEI-WUV algorithm supports prioritization of the force configurations without attempting to predict the true effectiveness values. The prioritization of configurations in based on the degree of robustness of putatively acceptable options.

Info-gap analysis of robustness provides the basis for deliberation and assessment of confidence in a decision, when facing substantial uncertainty. Robustness curves such as those in figs. 1 to 3 focus attention on two salient aspects of the decision.

First, the property of zeroing asserts that putative evaluations have no robustness against error in the data and mathematical models upon which the evaluations are based. This means that preferences based on putative evaluations cannot be confidently accepted when facing severe uncertainty. The putative evaluations are "best estimates" in the sense that they use accepted or normative data and models. But these estimates must be treated cautiously when facing severe uncertainty because "the best" may be substantially wrong.

Second, robustness-to-uncertainty trades off against the effectiveness margin that is required for decision. The putative effectiveness margin has zero robustness, and only lower (less decisive) margin has positive robustness. The robustness curve focuses attention on this irrevocable trade off, and supports deliberation and decision by quantifying the trade off. Robustness analysis is not inherently conservative. It does, however, indicate how large an effectiveness-margin can be confidently required.

The robustness analysis does not presume knowledge of the maximum possible error or uncertainty. One must distinguish between two distinct questions. One is the scientific question about truth and falsity: how wrong are we? The other is the robustness question: how much error can we tolerate in making a decision? The first question cannot be answered without improving our knowledge and understanding. The second question can be answered with our current state of imperfect knowledge, as we have demonstrated with three examples. The final decision will probably employ some qualitative judgment about the answer to the first question, but this does not depend on a full and explicit answer to that question. The info-gap robustness analysis is fundamentally different from a min-max or worst-case analysis, which require knowledge of the maximum error, as explained in detail elsewhere (Ben-Haim *et al.,* 2009, section 7).

Acknowledgement This paper was written while the author was at the University of Oxford as a Distinguished Visiting Scholar, an Oxford-Martin Visiting Fellow and a Leverhulme Visiting Professor. The author is grateful for the support of these foundations.

REFERENCES

Ben-Haim, Yakov, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty,* 2nd edition, Academic Press, London.

Ben-Haim, Yakov, 2010, *Info-Gap Economics: An Operational Introduction,* Palgrave-Macmillan.

Ben-Haim, Yakov, 2014, Strategy selection: An info-gap methodology, *Defense & Security Analysis,* 30(2): 106-119.

Ben-Haim, Yakov, 2015, Facing the unknown: Dealing with uncertainty in defense policy and strategic decisions, *Parameters,* US Army War College Quarterly, Vol, 45, No. 3, Autumn 2015.

Ben-Haim, Yakov, 2016, Policy neutrality and uncertainty: An info-gap perspective, to appear in *Intelligence and National Security.*

Ben-Haim, Yakov, Clifford C. Dacso, Jonathon Carrasco and Nithin Rajan, 2009, Heterogeneous Uncertainties in Cholesterol Management, *International Journal of Approximate Reasoning,* Vol. 50: pp.1046-1065.

Davidovitch, Lior and Yakov Ben-Haim, 2011, Is your profiling strategy robust? *Law, Probability and Risk,* 10: 59-76.

Gray, Colin S., 2010, War–continuity in change, and change in continuity, *Parameters,* Summer, 2010, pp.5-13.

Krepinevich, Andrew F. and Barry D. Watts, 2015, *The Last Warrior: Andrew Marshall and the Shaping of Modern American Defense Strategy,* Basic Books, New York.

Moffitt, L. Joe, John K. Stranlund, and Barry C. Field, 2005, Inspections to Avert Terrorism: Robustness Under Severe Uncertainty, *Journal of Homeland Security and Emergency Management,* Vol. 2, No. 3. http://www.bepress.com/jhsem/vol2/iss3/3.

Raymond, Major Allen D., 1991, *Assessing Combat Power: A Methodology For Tactical Battle Staffs,* School of Advanced Military Studies, United States Army Command and General Staff College, Fort Leavenworth, Kansas.

Sisso, Itay, Tal Shima, and Yakov Ben-Haim, 2010, Info-gap approach to multi agent search under severe uncertainty, *IEEE Transactions on Robotics,* Vol. 26, No. 6, pp.1032-1041.

Speight, L. R., D. Rowland and M. C. Keys, 1997, Manoeuvre Warfare: Force Balance in Relation To Other Factors and to Operational Success, *Military Operations Research,* V3 N3, pp.31-46.

U.S. Army Concepts Analysis Agency, 1991, *FY 91 Annual Report,* Bethesda, Maryland.

U.S. Congressional Budget Office, *U.S. Ground Forces and the Conventional Balance in Europe,* U.S. Government Printing Office, June 1988. https://www.cbo.gov/sites/default/files/100th-congress-1987-1988/reports/doc01b-entire.pdf, accessed 9.2.2016.