SMALL-UNIT WATER PURIFIERS IN U.S. ARMY SPECIAL OPERATIONS: A MULTI-ATTRIBUTE EVALUATION

Lauren A. Koban

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in Environmental Sciences in the Gillings School of Global Public Health.

Chapel Hill
2014

Approved by:

Jacqueline MacDonald Gibson

Jill Stewart

Eric Kelly

ABSTRACT

Lauren A. Koban: Small-Unit Water Purifiers in U.S. Army Special Operations: A Multi-Attribute Evaluation
(Under the direction of Dr. Jacqueline MacDonald Gibson)

Due to the austere and isolated locations of their missions, U.S. Army special operations forces units need to be self-sufficient in sustaining their potable water supply needs for survival. Current equipment used in the conventional Army is too heavy and operationally complex to meet size, mobility, and maintenance requirements. Therefore, special forces purchase most of their water purification equipment off-the-shelf; these systems are not designed with special forces in mind. This research applies multi-attribute decision analysis methods to identify a preferred commercial off-the-shelf water purification system for use in a special operations forces environment. Using feedback from seven public health professionals and end users in the Army, four water purification systems were identified to evaluate against nine performance criteria. The results illustrate the utility of multi-attribute decision processes in selecting technologies when there are multiple performance objectives and no single technology best meets any single objective.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3rd SFG	Unites States Army 3rd Special Forces Group
AHP	Analytic Hierarchy Process
ANSI	American National Standards Institute
CASCOM	United States Army Combined Arms Support Command
CEM	Certainty Equivalent Method
CHHPPM	United States Army Center for Health Promotion and Preventative Medicine
COTS	Commercial-off-the-shelf
DAT	Decision Analysis Team
ECBC	Edgewood Chemical Biological Center
MADA	Multi-attribute Decision Analysis
MAUT	Multi-Attribute Utility Theory
PEM	Probability Equivalent Method
PWD	United States Army Petroleum and Water Department
ROC	Rank Order Centroid
SMART	Simple Multi-attribute Rating Technique
SMARTS	SMART using Swings
SMARTER	SMART Exploiting Ranks
SOF	Special Operations Forces
TOPSIS	Technique for Order Preference by Similarity
USAPHC	United States Army Public Health Command
USASFC	United States Army Special Forces Command
USASOC	United States Army Special Operations Command
WPM	Weighted Product Model
WSM	Weighted Sum Model

1. INTRODUCTION

According to the United States Army Combined Arms Support Command (CASCOM) Water Planning Guide, safe water is “essential not only for the sustainment of life, but critical to the combat effectiveness of a military force"(CASCOM, 2008). Within the US Armed Forces, debilitating illnesses and injuries due to medical threats have caused more casualties than battlefield injuries throughout American history(U.S. Dept of the Army, 2002). These medical threats include waterborne diseases that are commonly transmitted through contaminated water. Therefore, achieving the ultimate objective of a military force requires soldiers to maintain a constant state of good health to maintain mission readiness. This thesis assesses alternative technologies for providing safe drinking water to United States Army Special Operations Forces (SOF) soldiers stationed in remote areas of Afghanistan in order to maintain soldiers’ health.

Due to their unique mission requirements, SOF units face different capability gaps in water purification technologies compared to the Army’s conventional force. Current equipment used in the conventional Army does not meet size, mobility, manpower, maintenance, or water production requirements of SOF operations. As a result, SOF units rely on commercial-off-the-shelf water purification systems. These systems are not purpose-built for the military and hence also may not be optimized for SOF needs. Furthermore, there is a lack of guidance or criteria for evaluating commercial-off-the-shelf products. Many off-the-shelf systems lack necessary treatment methods required to eliminate all of the possible contamination threats that SOF units may encounter in operational environments (Lundquist, White, Bonilla, Richards, & Richards, 2011). Instead of following a specific standard, SOF medics rely on research from the United States Army Public Health Command (USAPHC).

However, USAPHC research focuses on conventional Army needs, rather than specifically on the requirements of SOF units (USAPHC, 2010).

In Afghanistan, SOF units are currently engaged in village stability operations, where they are employing a "bottom-up" approach to fostering stability for the populace (Connett and Cassidy, 2011). Village stability operations are conducted by small teams in strategically important rural areas and work with the local populations. A typical embedded team ranges between eight to twelve men, including a medic responsible for field sanitation duties, which include water purification. These teams are trained to maintain a low profile and minimal footprint in their area to avoid detection. Due to the austere locations of the rural villages and the long to larger support areas of operation, internal logistics are a challenge. Consequently, SOF units need to be self-sufficient for survival and operations, requiring them to meet food, water, medical, and personal needs with little to no logistical support (Army FM 3-05, 2010).

Currently, in order to obtain drinking water, embedded teams use bottled water, village wells, or local surface water-sources, or they establish contracts with the local population. The teams purify the latter three (local) sources using commercial-off-the-shelf purification units, or construct their own on site gravity-fed sand filters. In extreme conditions, teams must boil their water or treat it in small batches by adding a coagulant and a disinfectant. These current options often are sub-optimal. The mountainous terrain in Afghanistan hinders resupply, making air drop missions of bottle water or replacement parts for off-the-shelf systems difficult due to lack of security and unpredictable weather. Additionally, some embedded units are operating in close proximity to the enemy, and maintaining their low signature profile is paramount to mission success. There are also difficulties in confirming the quality or source of locally contracted water.

In 2010, the USAPHC, formerly the United States Army Center for Health Promotion and Preventive Medicine (USACHPPM), conducted a technical evaluation of nineteen small-unit water purifiers to expand their knowledge base, and improve support to units wishing to purchase commercial-off-the-shelf purifiers. The goals of this evaluation included expanding the limited knowledge base on small-unit water purifiers, improving support to units wishing to purchase purifiers, and assisting in future procurement and use of commercial-off-the-shelf purifiers when Army-provided water sources are not adequate (USAPHC, 2010). The evaluation focused on small-unit water purifiers to sustain 5-50 personnel requiring 30-425 gallons per day over a period of ten days to six months (CASCOM, 2008). However, this evaluation did not address the specific needs in an SOF environment. The results concluded that no single system performed optimally across all of the different performance criteria the USAPHC had identified and that as a result tradeoffs would be required to select a system. For example, systems with lower sizes and weights produce smaller volumes of water, hence requiring trade-offs between the weight and water production performance criteria.

CASCOM is currently conducting a small-unit water purifier study focused on supporting 40-45 personnel with a minimum of 160 gallons per day. Although initiated in 2012, results are not anticipated until 2015. Additionally, the target populations for this study are platoon and company size elements which are significantly larger than SOF teams. Due to the limitations in current and past research specific to the unique environment, team size, and time sensitive requirements for missions, there is a need for a SOF-specific commercial-off-the-shelf water purifier evaluation model.

Research Objective

The main objective of this research is to help the U.S. Army Special Operations Command select the best existing commercial-off-the-shelf system based on mission requirements, taking into account the trade-offs that must be made in performance criteria. This research also aims to establish an evaluation model that accounts for multiple, sometimes conflicting technology performance attributes and can be adapted to future mission requirements.

Significance

Although CHPPM conducted a thorough small-unit-water purifier study in 2009, attributes specific to SOF environments were not included. The research provided in this thesis can help the SOF community not only select among currently available off-the-shelf systems but also develop a SOF-specific protocol to provide deploying units with the optimal choice for their specific location or mission. This research addresses the lack of adequate small-unit water purifier evaluations pertinent to the SOF community by establishing an evaluation model that aids in identifying the best equipment available, ensuring improved soldier sustainment on the battlefield.

2. BACKGROUND

Multi-attribute decision analysis (MADA) methods can help decision-makers choose among alternative options when no single option dominates all others in meeting all of the decision-maker's objectives(Huang, Keisler, & Linkov, 2011).

MADA has been used across many disciplines; examples include public transportation projects (Site & Filippi, 2009), real estate evaluation (L. F. A. M. Gomes & Rangel, 2009), sustainable development planning (Kain & Söderberg, 2008), renewable technologies selection (Afgan & Carvalho, 2002), and NASA missions(Tavana & Hatami-Marbini, 2011).

Within environmental decision making, MADA has been used for water resources and planning (Karjalainen et al., 2013), selection of remediation techniques of contaminated sites, optimization of coastal and water resources (Linkov et al., 2006), protecting aquatic ecosystems, and forest management and planning (Ananda & Herath, 2009).

A recent study by Huang et al. (2011) showed that the use of MADA tools in environmental decision-making has grown significantly over last two decades. Huang et al. hypothesized that this growth can be attributed to increased decision complexity as knowledge of environmental processes becomes more sophisticated and increased stakeholder demands for transparency in the environmental decision-making process (Huang et. al 2011).

There is strong precedent within the military for using MADA methods. The Army used MADA methods to prioritize military bases for closure or realignment under the 2005 Base Realignment and Closure program (Ewing Jr., Tarantino, & Parnell, 2006). Von Winterfeldt used MADA to help NATO trade off weapon weight and range in selecting rifles (Von Winterfeldt, 1986). Yoon & Hwang (1995) illustrated the use of MADA during officer promotion boards, where selection members have to identify the best-qualified officers for promotion based on military education, civilian education, physical readiness, duty performance, and potential.

The U.S. Army Corps of Engineers "Trade-Off Analysis Planning and Procedures Guidebook" describes MADA methods for use within the Army Corps of Engineers Civil Works planning process (Yoe, 2002). Most importantly, the Army also has a decision analysis team that works on chemical and biological research at the Edgewood Chemical Biological Center, but can be consulted for other projects. The decision analysis team assisted with the 2010 Small-Unit Water Purifier Study initiated by Army Public Health Command (USAPHC) and used decision analysis methods to evaluate potential commercial-off-the-shelf systems for use by military personnel supporting medium- and large-sized units (USAPHC, 2010). Although two previous off-the-shelf purifier studies have been conducted, both failed to look at the specific requirements of an eight to twelve man team. The current CASCOM study is researching purifiers for units of 40-45 personnel, which require over four times more water production than that of an SOF team. Additionally, SOF teams operate in austere environments requiring them to set-up, maintain, and transport purifiers on their own, compared to larger units that have multiple soldiers trained to solely operate the water equipment. This research includes specific requirements to address the ease of maintenance and maneuverability for an SOF team with minimal purifier training.

MADA Techniques

There are many different methods within MADA and the method selected varies based on the context of each decision. Table 1 shows seven categories of methods. .

Table 1. Methods for multi-attribute decision analysis, utility and weight determination

Multi Attribute Decision Analysis Methods	Method Includes Weights	Method Requires Determination of Weights
Multi-Attribute Utility Theory (MAUT)		x
SMART, SMARTS, SMARTER	x	
Analytical Hierarchy Process (AHP)	x	
Outranking	x	
Weighted Sum Model (WSM)		x
Weighted Product Model (WPM)		x
Technique for Order Preference by Similarity (TOPSIS)	x	

All MADA techniques consist of the same components: a set of alternatives, a set of attributes (meaning features of each option of importance to the decision-maker), weights for the attributes (describing the relative importance of the different attributes to the decision-maker) , and a trade-off algorithm (Yoe, 2002). One notable difference among methods is in the approach for weighting attributes (that is, for assigning relative priority to one attribute over another). As Table 1 shows, some methods have built-in algorithms for determining weights, while others require that weights be determined specifically for the decision at hand. There are other differences among methods, as well, including different protocols for eliciting inputs, modeling preferences, combining inputs and preferences, and analyzing the results (Huang et al., 2011). With so many potential variants to MADA, Triantaphyllou (1989) noted the MADA paradox "What decision making method should be used to choose the best decision making method?" (p.303).

Multi-Attribute Utility Models

Multi-attribute utility theory (MAUT) score each alternative against individual criteria (e.g., weight, production rate) and then use a mathematical function to aggregate individual attribute scores into an overall score for each alternative. The aggregation function may be linear or multiplication in the attributes depending on the decision-maker's preferences.

Keeney and Raffia (1976) developed the MAUT process based on earlier utility theory work by von Neumann and Morgenstern in 1947 (Huang et al., 2011). The MAUT approach assumes the decision-maker is rational, has perfect knowledge, and is consistent in judgment (Linkov et al., 2006).

The most common form of aggregation function assumes that the decision-maker's preferences can be modeled with a function that is linear in each attribute. This model is expressed as:

$$(x_1, x_2, \dots, x_n) = \sum_{i=1}^{n} k_i U_i(x_i) \tag{1}$$

Where

$$n = \mathit{number\ of\ attributes}$$

$$k_i = \mathit{weight\ of\ ith\ attribute}$$

$$U_i(x_i) = \mathit{utility\ score\ for\ ith\ attribute}$$

$$0 \leq k_i \leq 1$$

$$\sum_{i=1}^{n} k_i = 1$$

In order to accurately model preferences with such an additive utility function, conditions known as "mutual utility independence" and "additive independence" must be satisfied. According to Clemen (2001), "An attribute Y is considered utility independent of attribute X if preferences for uncertain choices involving different levels of Y are independent of the value of X" (p. 648). Additive independence is a similar but stricter condition requiring that preferences among alternatives hold even when the outcome of a decision is uncertain (in other words, levels of each attribute are not guaranteed) (Clemen, 2001).

Although using an additive utility function requires conformance with mutual utility and additive independence conditions, Clemen (2001) notes “in extremely complicated situations with many attributes, the additive model may be a useful rough-cut approximation” (p. 539).

SMART, SMARTS, SMARTER

The Simple Multi-attribute Rating Technique (SMART) was presented by Edwards in 1977 to provide a direct assessment method that was easier than the indifference methods required by Keeney and Raiffa’s (1976) approach to deriving MAUT functions. SMART eliminated judgment of preferences or hypothetical indifference between entities, making it easier to teach and use (Edwards, 1977). The method consists of two stages: first attributes are ranked based on importance of the best performance. Next, attributes are scored based on their importance compared to the worst attribute, which is scored a 10. All scores are then normalized to one (M. Gomes, Alberto, Rangel, & Leal, 2011). A criticism of SMART is that it does not consider the range of the each attribute (Edwards & Barron, 1994).

Edwards and Barron (1994) corrected the lack of range by proposing SMARTS (SMART using Swings), which added a hypothetical alternative based on the worst level of each attribute used as a comparison or benchmark. The swing is the changing of an attribute score from its worst value to its best (0 to 100). For example, in a car buying scenario with three cars and four value attributes, the hypothetical benchmark would be a fourth option that scored a 0 in all attributes. Once all attributes are chosen in the order they would be improved, the swings in attribute scores are compared against each other. This method addresses the range of each attribute but has been criticized because of the time consuming nature of the weight elicitation process was and the potential for difficulty in judgments for decision-makers inexperienced with the swing weighting method (Edwards & Barron, 1994).

To further refine the SMARTS process, Edwards and Barron (1994) introduced SMARTER (SMART Exploiting Ranks) which used rank weights to remove the weight elicitation step in SMARTS. This model was based on the rank order centroid (ROC) weights, developed by Barron and Barrett (1996), which are calculated by

$$w_k = (\frac{1}{K}) \sum_{i=k}^{K} (\frac{1}{i}) \tag{2}$$

$$where\ K = number\ of\ attributes$$

$$w_k = the\ weight\ of\ the\ kth\ attribute$$

$$i = the\ rank\ of\ the\ kth\ attribute$$

SMARTER was considered a significant improvement on SMARTS because it no longer required interviews and appealed to researchers because mailed surveys could be used. In comparison studies, Barron ((Edwards & Barron, 1994) found ROC weights to gain 98 to 99% of the utility in full weight elicitation methods and to identify the best option 74 to 87% of the time. In situations where the best option wasn't selected, the second best option was selected by SMARTER. However, two concerns of SMARTER are the lack of insight occurs during the swing weighting process from the decision-maker since he or she cannot applying specific weights and limited research on its effectiveness in supporting decision-making (Edwards & Barron, 1994).

Analytical Hierarchy Process (AHP)

The Analytical Hierarchy Process (AHP) was developed by Saaty (1980) and is a group of approaches that uses a hierarchical model and pairwise comparisons to determine the importance of one attribute over another. This method establishes a hierarchy of objectives, attributes, sub attributes, and alternatives. Pairwise comparisons made by asking "How important is attribute A_i relative to A_j?" are used to assess the relative importance of attributes using a number scale of 1 to 9 f (Table 2) (Fülöp, 2001, p.7). The available values for each comparison are members of the set: {9,8,7,6,5,4,3,2,1, ½, $^1/_3$, ¼, $^1/_5$, $^1/_6$, $^1/_7$, $^1/_8$, $^1/_9$} where the reciprocal of the Table 2 values are used if A_j is favored over A_i.

Comparison values are organized in a matrix, and matrix algebra is used to determine weights as the elements in the eigenvector associated with the maximum eigenvector of the matrix.

Pairwise comparisons are then conducted between alternatives on each attribute using the same scale in Table 2. Performance scores for each alternative are calculated the same way as the weights across each attribute using the question "How important is system A relative to system B? ". Once weights and performance scores are calculated, they are combined using MAUT aggregation techniques, which provide an overall ranking for each alternative. (Fülöp, 2001; Triantaphyllou & Mann, 1995).

Table 2. Analytical Hierarchy Process Scale of Relative Importance according to Saaty (1980)

Intensity of Importance	Definition	Explanation
1	Equal Importance	Two activities contribute equally to the objective
3	Moderate Importance	Experience and judgement strongly favor one activity over another
5	Strong Importance	Experience and judgement strongly favor one activity over another
7	Very strong or demonstrated importance	An activity is strong favored and its dominance demonstrated in pracctice
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2,4,6,8	Intermediate values between the two adjacent judgements	When compromise is needed
Reciprocals of above nonzero	If activity *i* has one of the above nonzero numbers assigned to it when compared with activity *j*, then j has the reciprocal value when compared with *i.*	

When compared to MAUT and outranking methods described below, AHP historically has been used the most in environmental decision making due to the wide availability of software packages and support for user groups (Huang et al., 2011). This process is also considered simple and flexible when involving multiple stakeholders, and can be used with relative values for each attribute instead of actual values (Triantaphyllou & Mann, 1989). Sharma (2013) stated when dealing with multiple stakeholders, AHP provides “a useful mechanism for checking the consistency of the evaluation measures and alternative suggested by the team thus reducing bias in decision making” (p.51).

Conversely, Belton and Gear (1986) found “The limitation of the scale to 1-9 imposes unnatural restrictions on judgments” (p. 11) and proposed a revised AHP version. The revised version recommends dividing each relative value by the maximum value of the relative values instead of the earlier version ensuring the relative values of the alternatives A_1, A_2, A_3, …, A_M sum to one (Triantaphyllou & Mann, 1989). One critique of both versions of AHP is the complexity of comparisons as the number of attributes increases. Using four alternatives and nine attributes would require a decision-maker to make 198 comparisons. Additionally, without a software package, the AHP method is difficult to apply.

Dyer (1990) has also critiqued the AHP because, mathematically, it can be shown that rank reversals are possible when employing this method, meaning that a decision-maker may change his or her preference for option A over option B, if a third alternative is added. This tendency for preference reversal violates the axioms required to satisfy the Von Neumann-Morgenstern utility theorem and hence calls into question whether the method accurately reveals the preferences of a ration decision-maker. Dyer (1990) further concluded that as a result of the lack of consistency with the axioms of utility theory “the rankings provided by the [AHP] are arbitrary.” (Dyer, 1990, p. 252).

Outranking Methods

Outranking methods are based on the principle that one option may have a degree of dominance over another. Dominance between two alternatives occurs when one performs better than another on at least one attribute and does not perform worse than the other option on any attributes. Outranking models compare the performance of two or more alternatives at a time, initially in terms of each attribute, to identify the extent of one preference over another. Preference information is then aggregated across all relevant criteria to establish the strength of evidence favoring one alternative over others (Linkov et al., 2006).

One flaw of outranking techniques is that they do not always identify the single best alternative. Outranking allows lesser performance on some attributes to be compensated for by superior performance on other attributes, leading to the alternative that performs the highest on the most attributes being favored when it may not be the best option (Linkov et al., 2006). Although outranking methods are successful for initiating a dialogue between multiple stakeholders, they do not provide a single solution to problems; Instead they drive a deliberative process between multiple stakeholders (Huang et al., 2011). Additionally, algorithms used in outranking are complex, and are not easily understood by decision-makers (Linkov et al., 2006). Outranking techniques are best suited when attribute metrics are not easily aggregated, measurement scales vary over large ranges, and units are disproportionate or incomparable (Seager, 2004).

Weighted Sum Model and Weighted Product Model (WSM, WPM)

The Weighted Sum Model is the most commonly used approach in single dimension problems, where all units are the same (dollars, feet, seconds). With M alternatives and N attributes, the best alternative is the one that satisfies (in the maximization case):

$$A^*_{WSM} = \max_i \sum_{j=1}^{N} a_{ij} w_j \quad for\ i = 1, 2, 3, \ldots, M \tag{3}$$

Where A(WSM score) = the WSM score of the best alternative

a_{ij}= the actual value of the ith alternatives in the jth attribute

w_j = weight of importance of jth attribute

This method is one of the most widely used; however, it is difficult to apply to multi-dimensional decision making problems that involve combining different units (Triantaphyllou & Mann, 1989).

The weighted multiplication method is similar to WSM but uses a multiplicative model, where each alternative is compared to others by multiplying a number of ratios. Ratios are raised to the power of the attribute's relative weight. To compare alternatives A_K and A_L, the ratio, $R(A_K/A_L)$ is calculated using the following equation:

$$R\left(\frac{A_K}{A_L}\right) = \prod_{j=1}^{N} \left(\frac{a_{K_j}}{a_{L_j}}\right)^{w_j} for\ K, L = 1, 2, 3, \ldots, M \tag{4}$$

If $R(A_K/A_L)$ is greater than or equal to one, the alternative in the numerator or A_K is preferred over the alternative in the denominator. This method is effective with both single and multidimensional problems and is dimensionless. Another advantage is it can also use relative values instead of actual values(Triantaphyllou & Mann, 1989). However, as the number of attributes and alternatives increase, this method becomes overly complicated for both the decision-maker and the analyst. Additionally, in Triantaphyllou and Mann's (1989) comparison of MADA method that included the AHP, revised ASHP, WSM, and WPM, they found the revised AHP to be by far the most accurate method of the four.

Technique for Order Preference by Similarity (TOPSIS)

The TOPSIS method was developed by Yoon and Hwang in 1981 and refined in 1987 (Yoon & Hwang, 1995). It is based on the idea that the best solution should have the shortest geometric distance from the ideal solution and the farthest distance from the least optimal solution. Alternatives are compared by establishing weights and normalized scores for each dimension. Next the distance is calculated between each alternative and the ideal alternative (best on each dimension), and the negative ideal alternative (worst) across the weighted dimensions. A ratio is then calculated between the negative ideal distance and the sum distance of the ideal and negative ideal for each alternative (Huang et al., 2011).

Two benefits identified by Huang et al. (2011) are that the only judgments needed are for the weights, and smoother tradeoffs are established due to the non-linear relationship between single dimension scores and distance ratios. However, compared to other MADA techniques, there are limited studies published using the TOPSIS method, with only five published articles in the fields of strategy and manufacturing identified in Huang et al.'s (2011) analysis of multi-attribute decision analysis. One limitation of TOPSIS is the need for complete or deterministic values, which are difficult to obtain in real world problems. Because of limited deterministic data, recent research has extended the TOPSIS method to address decision making problems with fuzzy data with positive results (Jahanshahloo, Lotfi, & Izadikhah, 2006).

Summary of Techniques

The selection of an MADA technique depends on multiple factors, including the scope of the problem, number of alternatives, number of attributes, nature of the attribute, and involvement of stakeholders. Numerous comparisons between the different MADA techniques have been conducted through the years. In an analysis of MADA used in environmental sciences over an eight year period, Huang et. al. (2011) concluded that regardless of the method, recommendations from decision-makers did not vary significantly.

In his comparison between the AHP, revised AHP, WSM, and WPM, Triantaphyllou (1989) found the revised AHP performed the best in a scenario with two criteria and varied weights; however, he also acknowledged that as the number of attributes increases, AHP may become too complicated for the decision-maker. Another critique of the AHP and outranking methods is that they lack a sufficiently strong axiomatic basis. The MAUT is based on von Neuman and Morgenstern's 1947 paper, "Theory of Games and Economic Behavior" that established a set of axioms for choice behavior that leads to maximization of expected utility, which guarantees to results in a choice that reflects the best option as would be viewed by a rational decision-maker (Clemen, 2001). Although Saaty established axioms for the AHP in 1986, his axioms are criticized for a lack of a testable description of behavior (Dyer, 1990).

Although many MADA techniques have been used when involving stakeholders, the MAUT approach was the selected method for this research due to its strong axiomatic foundation.

Utility Determination Techniques

In order to evaluate each alternative on each attribute, utility functions must be established to accurately portray a decision-maker's preferences. These functions translate quantitative or qualitative data for each alternative into quantitative scores on the same scale, so that attributes can be compared directly. The following methods can be used to establish utilities:

a. Indifference Methods
b. Direct Rating
c. Proportional scores

Indifference Methods (Certainty Equivalent and Probability Equivalent Techniques)

Indifference methods incorporate risk attitudes and consist of adjusting pairs of options until a decision-maker is indifferent between options. Two indifference techniques include the certainty equivalent (CE) and the probability-equivalent (PE) methods. Both methods involve a reference lottery and varying outcomes until a decision-maker is indifferent between the gamble and certain outcome. The certainty method provides a lottery between a gamble and a guaranteed return, such as the lottery illustrated in Figure 1.

In the example in Figure 1, the decision-maker's utility for option B can be determined by asking him or her to specify the amount of money he or she would accept to trade a guarantee of outcome B for a gamble between winning $100 and winning nothing. For example, if he or she would accept $40 in place of the chance to play the lottery shown in the figure but would not accept $39 to give up the lottery, then his or her indifference point—also called the certainty equivalent— is $40.

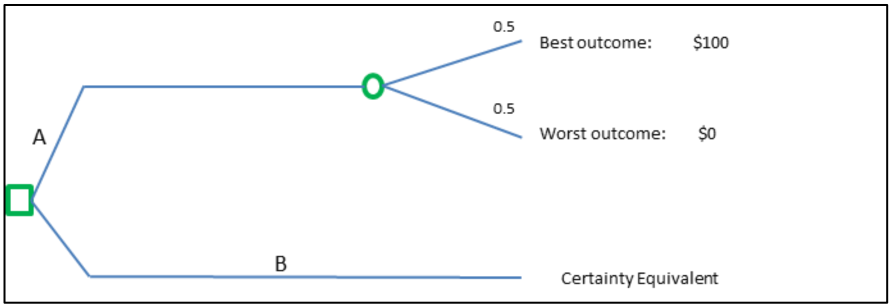


Figure 1. Certainty equivalent reference lottery

Similar to the previous method, the probability equivalent method also uses a reference lottery, but this time the probability is directly assessed between the lottery of best and worst outcomes and another given alternative. For example, using the same gamble values as above, and a guaranteed $65, a decision-maker would be given the options shown in the top branch of Figure 2 (winning $100 with a probability *p* and losing $100 with a probability of *1-p*).

The decision-maker adjusts the probability of winning between the best and worst until indifference is met between the lottery (option A) and $65 (option B), establishing the probability equivalent (Clemen, 2001).

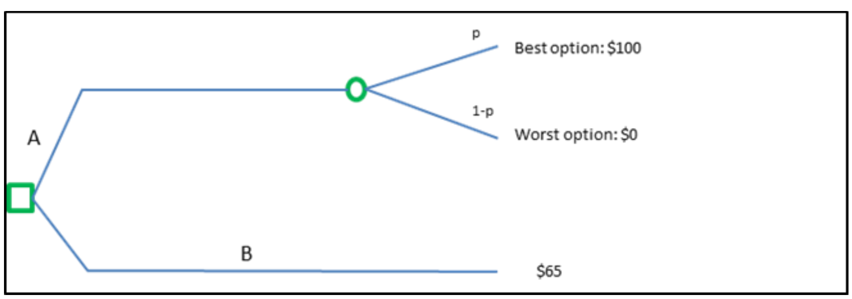


Figure 2. Probability equivalent reference lottery

Direct Rating

This method establishes numerical values for qualitative information by asking the decision-maker to select the best and worst alternatives within each attribute and using them as anchor points. All remaining alternatives are rated between the anchors, with higher scores showing stronger preference. Consistency checks are conducted by comparing alternatives against each other to ensure alternatives are ranked properly between anchors. For example, in a car buying scenario with red, blue, green, and yellow cars, a decision-maker prefers a red car the most and the blue car the least. Therefore, using the values of 0-100, he assigns scores of 100 to the most desirable option (red car) and 0 points to his least desirable option (blue car) making the red car and blue car his anchor points. In this scenario, the decision-maker would rank the yellow and green cars between the two anchor points. A consistency check would confirm that if the voter preferred yellow to green, then the yellow car should have a higher score than the green car.

Ratios

Ratios are another method to transform qualitative data into quantitative values. Instead of using the best and worst options as anchors in direct rating, this method establishes values using ratio comparisons. For example, in the above car buying scenario, if the decision-maker decides that a yellow car is twice as nice as a blue car, and that red is three times as nice as blue, then using scale of 0 to 100, 90 points are assigned to the red car, 60 points to the yellow car, and 30 points to the blue car (Clemen, 2001). Then, the resulting scores are scaled between 0 and 100, so that the red car is assigned a score of 100, the blue a score of 0, and the yellow car a score that is half-way between zero and 100 (i.e. 50).

Weighting Techniques

Weighting reflects the relative importance of attributes according to the decision-maker. This step enables decision-makers to create tradeoffs between attributes. Weights can be ordinal or cardinal. Ordinal values focus only on the numerical order, where cardinal weights address the order and magnitude between values (Ananda & Herath, 2009).

The MADA methods AHP, SMART, SMARTS, and SMARTER include weight determination within the process; however, MAUT, WSM, and WPM require the decision-maker to determine the weights.

Hence, in addition to choosing an MADA technique, for these latter three MADA approaches a method for choosing weights must also be selected. Weights can be elicited through the following processes:

a. Rank Weights
b. Fixed Point Scoring
c. Swing Weights
d. Tradeoff Weights

Rank Weights

This is the simplest concept and only requires decision-makers to rank attributes against each other in order of importance. Scales of 1 to 10 or 1 to 100 are commonly used. To calculate weights, ordinal rankings are reversed to determine importance points, so the most desirable attribute has the highest score and the least desirable attribute has the lower score. Weights for each attribute are calculated as the percentage of total scores for all attributes using equation 5. Although easy for the decision-maker to understand, this method does not force the decision-maker to make explicit tradeoffs between attributes or consider the range of scores between attributes (Von Winterfeldt, 1986).

$$k_i = \frac{a_i}{\sum_i a_i} \tag{5}$$

Where

$$k_i = weight\ of\ ith\ attribute$$

$$a_i = importance\ points\ of\ ith\ attribute$$

Fixed Point Scoring

Fixed point scoring gives a decision-maker a set number of points such as 100, and has him or her distribute points between all attributes. More points given to a specific attribute signify a higher preference for that attribute. This method is being used for weighting in the Army CASCOM study of small-unit water purifiers described in the introduction. This technique is simple for the decision-maker to understand; however, the decision-makers sometimes have difficulties with making tradeoffs between attributes (Yoe, 2002).

Swing Weighting

Swing weighting, published by Von Winterfeldt (1977) and Edwards (1986), is a three step process. The first step involves ranking the order of importance of the attributes based on the swing from the worst to the best level. This can be done using a hypothetical comparison that is considered the worst in all attributes. For example, consider choosing a car from a set of three models shown in Table 3, where the attributes of interest are cost, color, and gas mileage (indicated in Table 3 with utility scores).

Table 3. Utility scores for car buying scenario

	Attributes		
Car	Cost	Color	Gas Mileage
Model A	100	60	80
Model B	40	0	100
Model C	0	100	0

Values reflect utilities, where 100 is best score and 0 is worst score

To assess the order of importance of the attributes, the decision-maker is provided with a scenario in which he is required to purchase a fourth car model that scored a zero in all attributes. Then, the decision maker is asked which attribute he would select for improvement from worst to best, if he were allowed to improve only one of the attributes. If he chose cost, then he would be asked again which attribute he would choose, beyond cost. The line of questioning is continued until all attributes are selected. Attribute selection is shown in Table 4.

The second step is to elicit the relative value of the decision-maker of improving from the worst to the best outcome on each attribute, in comparison to swinging from the worst to best on other attributes. For example, suppose the decision-maker thought that swinging from 0-100 on gas mileage is 90% as valuable as swinging from worst to best on cost (the most important attribute). Then, the decision-maker would assign a value of 90 to the gas mileage attribute. This question is again repeated with color. If the decision-maker cared little about color, in comparison to cost, he may score the color swing as having a value of 10. The final step is normalizing the scores to establish the weights (Table 4) (Edwards & Barron, 1994). The weights are normalized by summing all the swing scores and then dividing the swing score for each attribute by this sum.

Table 4 Swing weights for car buying scenario

	Attributes		
Car	Cost	Color	Gas Mileage
Model A	100	60	80
Model B	40	0	100
Model C	0	100	0
Model D	0	0	0
Attribute Selection	1	3	2
Swing Score	100	10	90
Normalized Weight	0.5	0.05	0.45

Method Choice

The MADA technique selection is based on the type of decision, number of attributes, availability of a software support tool, and the number of stakeholders involved in a decision. Depending on the method selected, there may be additional decisions based on how to elicit utilities and weights.

Based on the multiple numbers of stakeholders involved in the small unit water purifier acquisition process and number of attributes with dissimilar measures, the multi-attribute utility theory was selected. Although initially only swing weighting was selected to elicit weights, the rank order centroid (ROC) method was also employed. Both techniques were selected based on ease of understanding and limited time burden for interviewees.

3. METHODOLOGY

The goal of this research is to identify the best commercial-off-the-shelf water purification system for use in the SOF environment using stakeholder input. This section describes the method to obtain stakeholder feedback and apply it using the multi-attribute utility theory. This research follows the decision analysis process steps outlined in Figure 3.

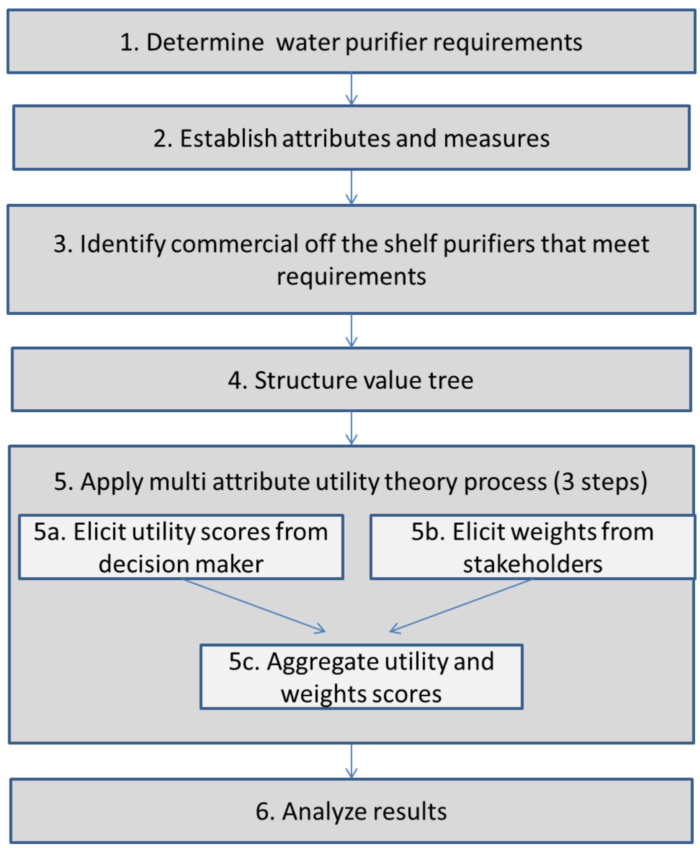


Figure 3 Decision Analysis Process for identifying a commercial-off-the-shelf water purification system

This research is based on the feedback of seven stakeholders working within the Army organizations listed in Table 5. Stakeholders were identified through communication with the United States Army's Special Operations Command Surgeon's Office. The first two stakeholders listed in Table 5 were selected based on their participation in previous small unit water purifier studies conducted by United States Army Public Health Command. The third stakeholder works within the Quartermaster Center and School's Petroleum and Water Department and has over twenty years of experience in the Army as a water treatment specialist with multiple deployments, while the last four from USASFC and 3rd SFG are active duty or retired soldiers serving within the Army's Medical Service Corps Preventive Medicine branch.

Table 5. Stakeholder organizations in descending order from higher echelons to end users, number of representatives (in parenthesis), and abbreviations used in the text

Organization	Abbreviation
United States Army Public Health Command (1)	USAPHC
United States Army Combined Arms Support Command Sustainment Division (1)	CASCOM
United States Army Quartermaster Center and School, Petroleum and Water Department (1)	PWD
United States Army Special Forces Command (2)	USASFC
United States Army 3rd Special Forces Group (2)	3rd SFG

Step 1: Determine Requirements

This research was initiated by the United States Army Special Operations Command (USASOC) Surgeon's Office, which recognized the lack of water purification devices specific to the SOF environment and the need for an evaluation process for commercial-off-the-shelf systems. Therefore, the USASOC Surgeon's Office was contacted for system requirements feedback due to its medical oversight role in preventing waterborne disease and ensuring adequate field water supplies within the SOF environment (U.S. Dept of the Army, 2010). Prior to starting any interviews or contacting vendors, the UNC Intuitional Review Board (IRB # 13-3054) approved all research methods.

Through email conversations, members of the surgeon's section provided an initial list of eleven requirements based on their preventive medicine role and experience using commercial-off-the-shelf purifiers in the SOF environment. To these requirements were added two requirements used in the 2010 water purification system study initiated by Army Public Health Command: meeting microbial and chemical reduction standards. These standards are outlined in the following three regulations: the NSF

International (NSF) / American National Standards Institute (ANSI) Protocol 248: Emergency Military Operations Microbiological Water Purifiers, NSF/ANSI Protocols 53: Drinking Water Treatment Units and NSF/ANSI Protocol 58: Reverse Osmosis Drinking Water Treatment Systems Overview. Although an important attribute to address, cost was not included in this research as an attribute due to an inability to obtain accurate quotes for the systems. Therefore, cost will not be addressed as an attribute and a cost benefit analysis will not be included in this research.

Using this initial list of thirteen requirements, phone interviews were conducted with the stakeholders listed in Table 5 to confirm or add additional requirements based on their personal experiences with water purification systems or previous water purifier studies. Based on this feedback, two more requirements were added: system filters must have the ability to treat enough water in the first 72 hours and the ability for the system to provide visual and audible warning if it fails to operate properly and water quality might be comprised. Because of the concern of filter life, the need for 30 days of accessories was added to the weight requirement. The process of identifying requirements is outlined in Figure 4 with the revised requirements listed in Table 6.

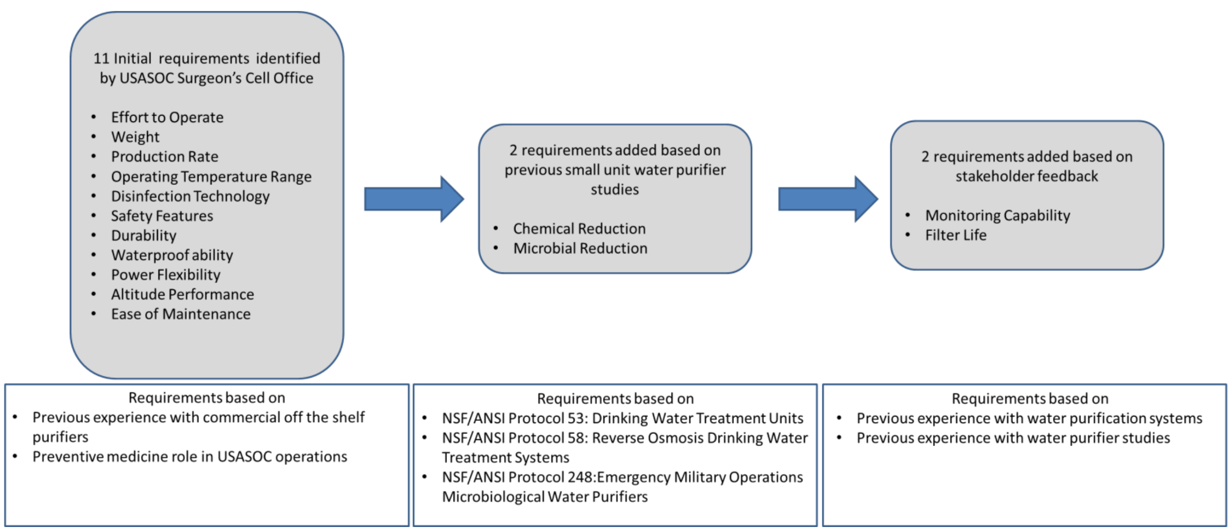


Figure 4. Requirement identification process

Table 6. Requirements list for water purification systems

Requirements
Must be operable by end user with minimal training (no more than 4 hours of instruction)
Weight of system plus 30 days of accessories must be less than one to two man lift requirements (87/174 pounds)
Must be able to deliver 72 hours of all required potable water (minimum) without resupply
Must be able to operate from 30 – 120 ^{o} Fahrenheit with 0 - 140 ^{o} Fahrenheit (optimal)
Must operate from 0 - 10,000 feet above mean sea level
Must be able to survive a drop of 3 feet onto a flat concrete surface without impairing functionality
Must not require follow-on disinfection for potablity
Must be maintainable by the end user without major vendor support (i.e. the end user can change filters, components, circuit boards, etc)
Must be waterproof to 66 feet (can be in a specialized container) for full immersion for delivery via maritime platform
Must operate on multiple power sources (50/ 60 Hz, 110 - 240V, power either via external port, generator, battery, or solar power) and have sufficient battery power to operate for 4 hours without external power supply at full production
Should have an automatic shutoff at the end of element life to prevent production of contaminated water
Must meet NSF Protocol 248 for reduction of microbiological contaminants
Must meet NSF/ANSI Protocols 53 and 58 for reduction of chemical contaminants
Should provide visual and audible feedback when system is not operating properly and water quality might be compromised
Filter having the ability to treat enough water for 20 Soldiers in first 72 hours/30 days - based off 8.5 gal/day per Soldier[a]: ***20 Soldiers for 3 days= 510 gallons*** ***30 days =5100 gallons***

[a] 8.5 gallons per day in arid environment and 6 gallons per day in temperate environment derived from CASCOM's Water Planning Guide (CASCOM, 2008)

Step 2: Establish Attributes and Measures

Attributes and measures were determined in a similar process to the requirements using feedback from the stakeholders listed in Table 5. Fifteen initial attributes were developed to assess and score the performance of each alternative against the requirements outlined in Table 6, but this list of attributes was shortened to nine based on the inability of four attributes to discriminate among the four water treatment systems and the two attributes lacking viable scoring measures (Figure 5).

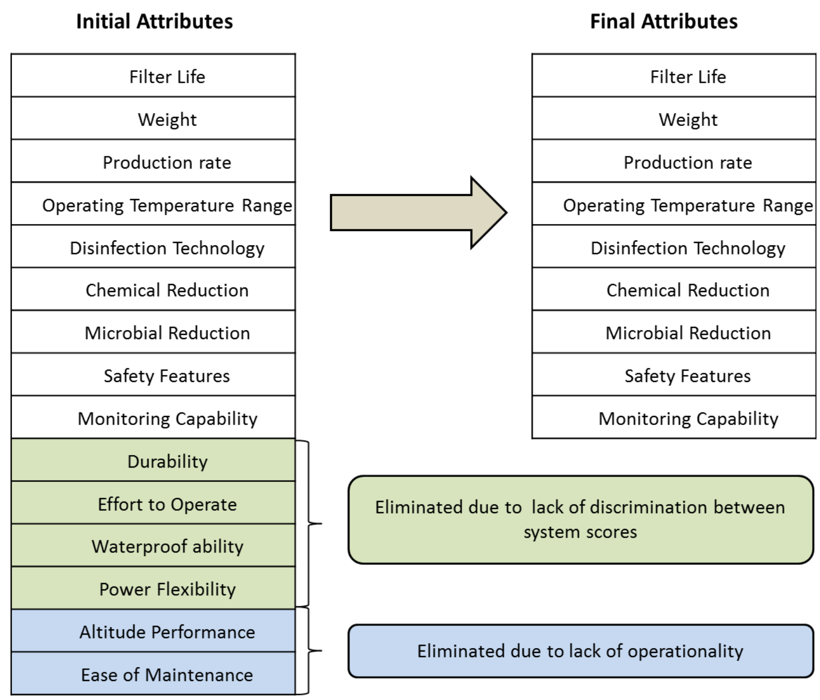


Figure 5. Elimination process from initial 15 attributes to final 9 attributes

Once attributes were finalized, measures were established for each in order to effectively score alternatives' performance for each attribute (Table 7). These measures were utilized in the following sections to elicit utility functions for each attribute.

Table 7. Established attributes, descriptions, and measures for water system evaluations

Attributes	Description	Measure
Chemical Reduction	Ability to meet NSF/ANSI Protocols 53 and 58 for reduction of chemical contaminants	Technology of each system to reduce chemical contaminants based on technologies addressed in NSF 58 and NSF 53: 1. carbon filtration + multi-stage cartridge + nanomesh filter 2. multi-media canister 3. NSF/ANSI P53 certified 4. carbon filter + sediment filter + reverse osmosis membranes
Microbial Reduction	Ability to meet NSF/ANSI Protocol 248 for reduction of microbiological contaminants	Certification (if any) by NSF in meeting NSF P 248; if not, current technology of system and testing results in meeting Log 6, Log 4, and Log 3 reductions: 1. NSF P248 certified 2. 99.99% reduction for bacteria and virus 3. technology of previous system NSF P248 certified
Weight	Total weight is < 80 pounds including all accessories for 30 days	Pounds
Filter Life[a]	Filter having the ablility to treat enough water for 20 Soldiers in first 72 hours/30 days - based off 8.5 gal/day per Soldier[a]: 20 Soldiers for 3 days = 510 gallons 30 days =5100 gallons	Life of filter by gallons purified until filter needs to be replaced. (General Water[b]) Life of filter by gallons purified until filter needs to be replaced (Challenge Water[b])
Operating Temperature Range	Ability to operate from 30 - 120 °F or 0 - 140 °F (optimal)	Temperature range of system
Production Rate	Ability to purify fresh water NLT .28 gallons per minute Calculated for 20 personnel based per 10 hour production day in arid environmenta	Production rate in gallons per minute
Disinfection Technology	Requirement of system for follow on disinfectionfor potability	Yes or No if system requires follow on disinfection. If No, list system technology: 1. UV 2. UV + optional chlorine injection 3. Membrane filtration good for 2 weeks, then requires chlorine
Safety Features	Must have an automatic shutoff at the end of element life to prevent production of contaminated water	Measured by technology in system to prevent production of contaminated water: 1. Flow rate stops when pre-filters clog 2. Two-stage protective shutdown circuit + automatic shutdown in the event of a failure of disinfection + pump stops and a valves close 3. Flow rate stops when filters clog 4. Automatic shutdown
Monitoring Capability	Ability to provide visual and audible feedback (warned when system is not operating properly and water quality might be compromised)	Measured in amount/types of audible feedback: 1. Pressure gage shows reduction in flow rate 2. Digital display shows water volume passing through system and total run time of filter canister and UV exposure unit 3. Field test for filter device 4. System status messages and water quality threshold monitor

[a] 8.5 gallons per day in arid environment and 6 gallons per day in temperate environment derived from CASCOM's Water Planning Guide (CASCOM, 2008)
[b] Without knowing exact water quality in operational setting, NSF P248 general and challenge water qualities wer

Step 3: Identify Alternatives

Three alternatives (Aspen 1800 BC™, Seldon Waterbox™, Global Water LS3 Village System™) for this model were predetermined by senior leaders within the United States Special Operations Command Surgeon Cell based on current systems in use and preliminary market research. However, after conducting an initial review of system specifications, insufficient performance data was available on the Global Water alternative, so it was removed.

In an effort to produce more varied results, we identified three additional systems through online market research. Alternatives were selected based on the ability to meet as many of the requirements listed in Table 6 as possible. Additionally, system selection was based on the ease of acquisition for future laboratory testing to validate system capabilities.

Two of these alternatives, the SLMCO FBS 180™ and SLMCO FBS 400™, were identified based on the high performance of a similar system, SLMCO 5.0™, from the same manufacturer in previous small unit water purifier studies (USAPHC, 2010). Although both FBS systems met requirements, only the SLMCO FBS 180 was evaluated based on its smaller size, as well as to reduce redundancies of having two systems from the same manufacturer. The third system, the Nephros MSU Ultra filter Water Purification System™ configured with an Aquamira DIVVY50™ water pump, was identified based on recommendations from the Combined Arms Support Command (CASCOM) small unit water purifier study (USAPHC, 2010). Although this system is being evaluated for CASCOM's current purifier study to support a larger number of soldiers (up to 50), its small size and technology also met requirements for this study.

The resulting four alternatives evaluated in this study were the Aspen 1800 BC™, Selden Waterbox™, SLMCO FBS 180™, and the Nephros MSU Ultra filter™ configured with an Aquamira DIVVY50™ water pump. All systems are considered briefcase size and meet the mobility requirements. The Aspen, Seldon and SLMCO models use multi-stage cartridge and carbon filtration systems while the

Nephros system uses ultra-filtration. A detailed description of each system according to each attribute is listed in Table 8.

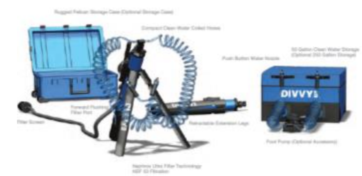
Table 8. Manufacturers' specifications for four water purification system alternatives

Small-Unit Water Purification System Specification Sheet: All Systems

Sources: http://seldonwater.com/product/waterbox-max/, http://www.aspenwater.com/id3.html, http://www.slmcopurewatersystems.com/man-portable-units/, http://www.nephros.com/military-water/, U.S. Army Public Health Command (USAPHC), Materiel Systems Directorate, CASCOM, Fort Lee, VA

Selden Waterbox

Aspen 1800 BC



Nephros MSU Ultra filter™ with Aquamira

SLMCO FBS 180

Objectives	Description	Measures	Seldon Waterbox™	Aspen 1800 BC™	Nephros MSU Ultra filter™ with Aquamira DIVVY50™ Configuration	SLMCO FBS 180™
Chemical Reduction	Ability to meet NSF/ANSI Protocols 53 and 58 for reduction of chemical contaminants	Certification (if any) by NSF in meeting NSF P53/58; if not, current technology of each system to reduce chemical contaminants based on technologies addressed in NSF 58 and NSF 53	• Multi-stage Cartridge and Carbon Filtration •2 Pre-filters (carbon core with nanomesh) • Main nanomesh filter	•Two stage filtration • Multi-media water canister	• NSF/ANSI P53 certified filter	•Sediment filters • carbon filters • NSF P53/58 certified reverse osmosis membranes
Microbial Reduction	Ability to meet NSF/ANSI Protocol 248 for reduction of microbiological contaminants: Bacteria - Escherichia coli, Raoultella terrigena, or Bacillus atrophaeus (spore form),	Certification (if any) by NSF in meeting NSF P248; if not, current technology of system and testing results in meeting Log 6, Log 4, and Log 3 reductions	Passed NSF P248 Protocol - 1 Mar 2013	99.99% reduction for bacteria and virus	Techology of previous system (UF- 40) NSF P248 certified	Techology of previous system (SLMCO 5.0) NSF P248 certified
Weight	Total weight is < 80 pounds including all accessories for initial operation - 30 days (does not include weight of storage tanks for 30 days)	Pounds (including weight of accessories for 30 days)	76 pounds	118 pounds	75 pounds	82 pounds
Filter Life	Filter having the ablility to treat enough water for 20 Soldiers in first 72 hours/30 days - based off 8.5 gal/day per Soldier[a]: ***20 Soldiers for 3 days= 510 gallons*** ***30 days =5100 gallons***	Life of filter by gallons purified until filter needs to be replaced. (General Water[b])	• Prefilter: 3,963 gallons • Nanomesh filter: 7,926 gallons	9,000 galllons	26,420 gallons	13,209 gallons
		Life of filter by gallons purified until filter needs to be replaced (Challenge Water[b])	• Prefilter:264 gallons • Nanomesh filter: 2,641 gallons	7,500 gallons	2,642 gallons	5,040 gallons
Temperature Range	Ability to operate from 30 - 120 °F or 0 - 140 °F (optimal)	Temperature range of system	41 °F -100.4 °F	-4 °F - 122 °F	-22 °F - 158 °F	4 °F - 110 °F
Production Rate	Ability to purify fresh water NLT .28 gallons per minute (to meet 170 gallons/day daily requirement for 20 soldiers) Calculated for 20 soldiers based per 10 hour production day in arid environment[a]	Production rate in gallons per minute	.6 gallons per minute/360 gallons per day	1.25 gallons per minut/750 gallons per day	1.95 gallons per minute/1,170 gallons per day	.14 gallons per minute/ 200 gallons per day[c]
Disinfection Technology	Must not require follow-on disinfection for potability	Requirement of follow-on disinfection	Yes, requires disinfection	No, Ultraviolet (UV) disinfection	No, Membrane filtration; recommend chlorine for storage > 2 weeks	No, Ultraviolet (UV) with optional chlorine injection
Safety Features	Must have an automatic shutoff at the end of element life to prevent production of contaminated water	List technology in system to prevent production of contaminated water	Flow rate stops completely when pre filters clog	• Two-stage protective shutdown circuit • Automatic shutdown in the event of a failure in the disinfecting process • Pump stops and valves close	Flow rate stops when filter clogs	Automatic shutdown
Monitoring	Ability to provide visual and audible feedback (warned when system is not operating properly and water quality might be compromised)	Amount and types of audible feedback	Pressure gage shows reduction in flow rate	Digital display shows: •water volume passing through the system • total run time of the filter canister and UV exposure unit.	Field integrity test (FIT) for filter devices; can test membrane at anytime	• System status messages • Water quality threshold monitor

[a] 8.5 gallons per day in arid environment and 6 gallons per day in temperate environment derived from CASCOM's Water Planning Guide (CASCOM, 2008)

[b] General and challenge water qualities as defined by NSF P248

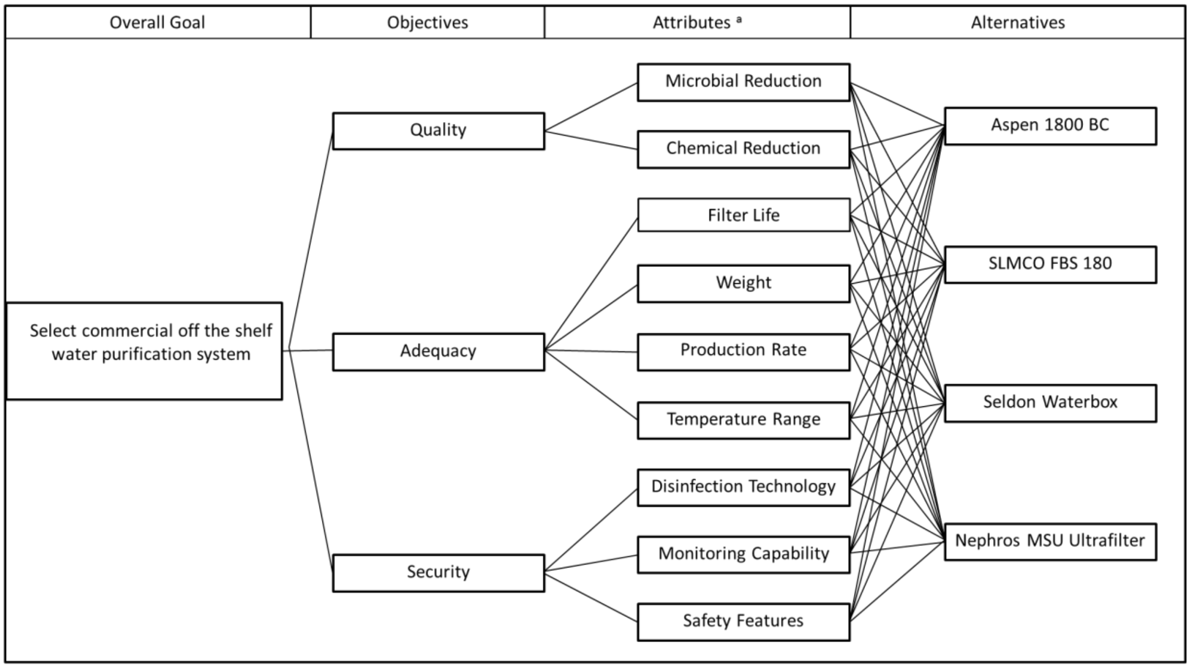
[c] This system is designed to run for 24 hours without stopping; values are based on 24 hour production day.

Step 4: Structure Value Tree

Once the requirements, attributes, and alternatives were identified, a value tree was structured that addressed the overall goal of this research. According to the Army's Technical Bulletin on Sanitary Control and Surveillance of Field Water supplies, one of the many roles of a preventive medicine officer in the Army is to ensure the "security, adequacy, and quality of field water supplies" (Technical Bulletin Med 577, Sanitary Control and Surveillance, 2010, p. 107). Therefore, when constructing the model for the small unit water purifier decision, the objectives of security, adequacy, and quality, were used to frame the value tree. Security encompasses the features of a system to safely treat water and provide an indication of compromised water quality. Adequacy addresses the system's ability to meet the demands in the SOF environment including operating in extreme temperatures, and being light enough for one soldier to transport. The objective of quality ensures that a water purification system is able to treat water to a safe level and not provide a health threat to soldiers.

The nine attributes were based on the initial requirements, feedback from the stakeholders (Table 5) collected through telephone and email conversations, and previous water purifier study reports. Initially, 15 attributes based on the 15 requirements identified in the previous section were established. Two attributes, altitude performance and ease of maintenance did not have enough data from manufacturers on the systems to establish measures. Those attributes were eliminated based on a lack of operationality, or the inability of attributes to be well-defined and viable for working and scoring (Keeney & Raiffa, 1976). During the first round of utility interviews, four additional attributes were eliminated due to a lack of discrimination because of similar features among all four systems, making it too difficult to compare the systems. The attribute elimination process is illustrated in Figure5.

Using the three objectives and final nine attributes, a value tree was constructed to structure the overall goal of selecting a commercial-off-the-shelf water purification system and is shown in Figure 6.



[a] Initially fifteen attributes were identified. Altitude performance and ease of maintenance were eliminated due to the principle of operationality (Keeney & Raiffa, 1976). Durability, effort to operate, waterproof ability, and power flexibility were excluded due to lack of discrimination during utility elicitation.

Figure 6. Final value tree for the multi-attribute decision analysis of commercial-off-the-shelf water purification systems

Step 5: Application of Multi-attribute Utility Theory Process

The linear additive model was chosen to represent decision-maker preferences because as Clemen and Reilly (2001) note "evidence has shown that the additive model is reasonable for most situations under conditions of certainty" (p.599). The following steps proposed by Von Winterfeldt (1986) were followed to apply the MAUT procedures:

1. Evaluate each alternative separately on each attribute to determine utility scores
2. From each decision-maker, elicit weights of the attributes
3. For each decision-maker, use the resulting multi-attribute function to compute the overall utility of each system using the linear additive model (Equation 1):

$$(x_1, x_2, \dots, x_n) = \sum_{i=1}^{n} k_i U_i(x_i) \tag{1}$$

Where

$$n = number\ of\ attributes$$

$$k_i = weight\ of\ ith\ attribute$$

$$U_i(x_i) = utility\ score\ for\ ith\ attribute$$

$$0 \leq k_i \leq 1$$

$$\sum_{i=1}^{n} k_i = 1$$

Each alternative was scored on each attribute using manufacturer's specifications provided through market research and phone calls with vendors. Although manufacturers' specifications have not been verified by independent lab testing, the focus of this research is on the model and assumes values are accurate. The United States Army Tank Automotive Research, Development and Engineering Center conducts testing on water purification systems, however the four alternatives selected were either not tested or have been modified since their last test. The current Combined Arms Support Command water purification study also includes some alternatives from this study, but due to funding, testing has been postponed until fiscal year 2015. Once the laboratory testing of the alternatives is completed, the MAUT model developed in this research can be updated.

Elicitation of Individual Attribute Utilities

Utilities for each attribute were elicited by one expert using two methods: the direct rating method and the probability equivalent method. Only one person was interviewed to elicit utilities for all attributes based on his experiences working in the SOF environment, multiple deployments to Afghanistan, and position as an Environmental Science and Engineering Officer. His background established his level of knowledge as expert and followed methods outlined in previous studies (Edwards, 1977). The decision to use only one expert for all nine utilities was also to reduce the time constraints on the remaining six stakeholders. Once his utility scores was collected, utility functions were established for each attribute in order to transform qualitative data into quantitative data. The utility elicitation interview consent form and questions are listed in Appendices B and C.

Both utility elicitation methods were conducted during the same two-hour interview period. Before the interview, the interviewee was provided an operational scenario document that was developed to reduce bias by identifying a particular case where a purifier would be used (Box 1).

Box 1. Operational Scenario for Use of Small Unit Water Purifier

1 Portability: During operation, the purifier would be stationary; however, based on mission requirements, the system needs to be portable when not in use. Platforms to transport the system may include mounting on a vehicle or trailer, or via sling load from a helicopter. Lifting the system onto the transportation platform will be done by soldiers. According to the Military Standard 1472-G, a two person lift should not exceed 174 pounds, and limits a one person lift to 87 pounds when placing the item on a surface not greater than 3 feet above the floor (*Department of Defense Design Criteria Standard Human Engineering*, 2012)(Bray, Rae Olmsted, Williams, Sanchez, & Hartzell, 2006).

2 Water Quality: Water sources within villages can include existing wells, contracted delivery of local water, or surface water. Water quality is assumed questionable, making it essential to remove all hazardous toxins to include microbial and chemical contaminants in order to prevent any detriment to the overall health of SOF personnel.

3 Daily Water Requirement: The minimum potable water requirement per soldier per day was approximately 8.5 gallons (32 Liters) and included considerations for drinking, personal hygiene, field feeding, heat injury treatment, vehicle maintenance, and medical treatment. Based on a VSO team of eight to twelve men, the daily demand from a water purifier ranges between 68 and 102 gallons (258 to 387L) (CASCOM, 2008).

4 Length of Mission: Village stability operations (VSO) usually range from 30 days to six months. The water purification system is assumed to be the only source of potable water with limited resupply of system components. Therefore each system and accessories must adequately provide all water requirements for initial 30 days of mission with resupply anticipated for the remaining five months.

5 Durability: Based on the climate of the operational environment, the water purification system must be durable enough to withstand extremely high or low temperatures, as well as rough handling or transport via multiple platforms such as air, water, or vehicle.

6 End User: Based on the small size of a VSO team, each member of the team may be required to set up, operate, or monitor the SUWP at any given time, regardless of his knowledge of water purification processes. Prior to deployment, training opportunities may be limited based off time and resources, making the ease of system use important. During missions, knowing when to shut down if water quality or system is compromised is essential for all team members.

7 Location: VSO teams are usually embedded within rural communities. Based on location of larger support bases, logistics are challenging. Terrain may limit vehicular traffic, and weather, altitude, or security may impact air resupply. Power sources may also be limited.

The direct rating process consisted of showing the interviewee one notecard for each of the four alternatives' performance on specific attributes. Figure 7 shows the notecards and questions used to determine the utility score for the attribute of weight. Systems were randomly labeled, and the expert had to first rank the four options from best to worst. Once cards were lined up from best to worst, the expert then scored the systems using the best and worst systems as the anchors. The best system was given a score of 100, and the worst system was scored 0. After the worst system was scored, the intermediate two systems were scored between the anchor scores.

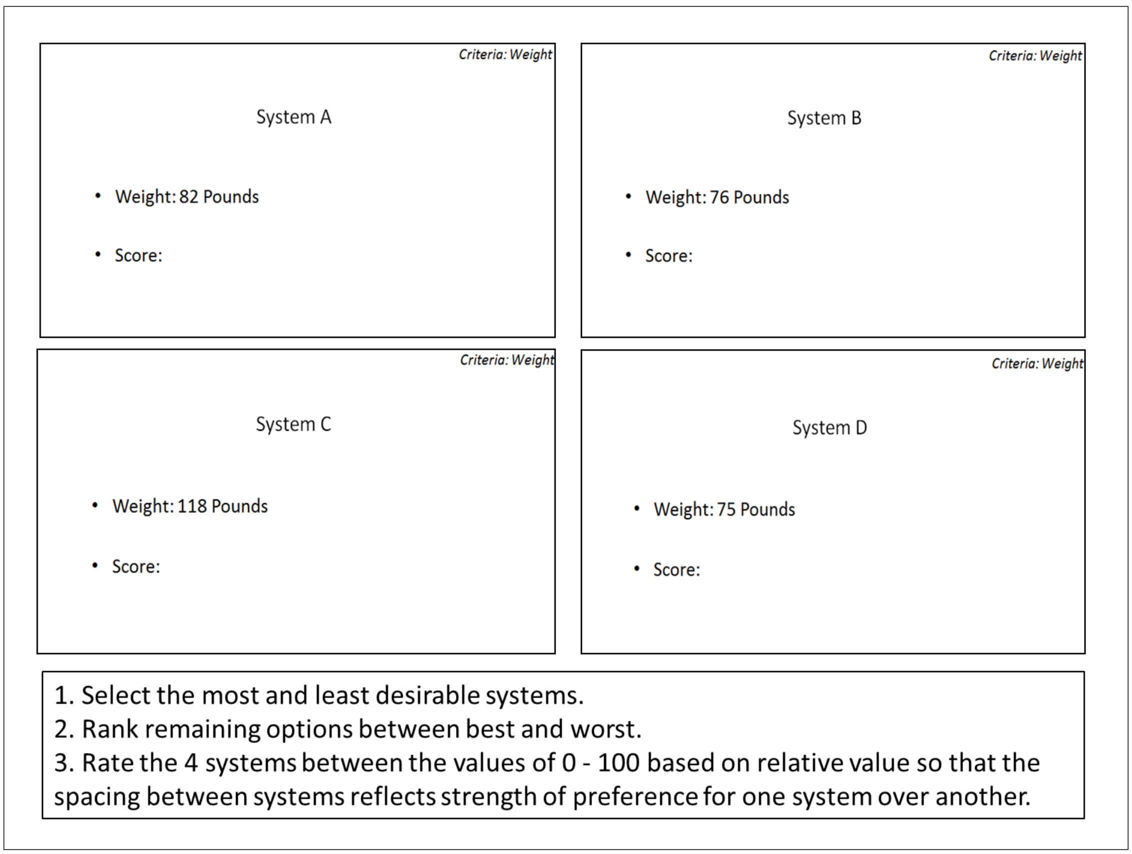


Figure 7. Notecards and questions used during direct rating utility elicitation for the attribute of weight

After the initial scoring process, a consistency check was conducted to ensure all intermediate values were accurately scored. The consistency check was a series of questions confirming the placement of alternatives between the two anchor systems and is listed below:

1. Read back and confirm each system scores and rankings with interviewee.
2. Ask the expert “You rated system B halfway between system C and D. Is this correct?” (this is to confirm the difference in points accurately portray expert’s preference)
3. Ask the expert “Given a new system that weighs 90 pounds, where would you rank and score it?” (this is to confirm it would be ranked between the systems A and C above)

This process was performed for all nine attributes. One attribute, filter life, required two rounds of scoring to address both general and challenge water. The scores were then averaged to provide only one score for the filter life attribute per alternative.

The probability equivalent method (PEM) interview was conducted after a short break. In this method, the expert was given a choice of either a lottery between the highest and lowest scoring system in each attribute, or the intermediate system (Figure 8).

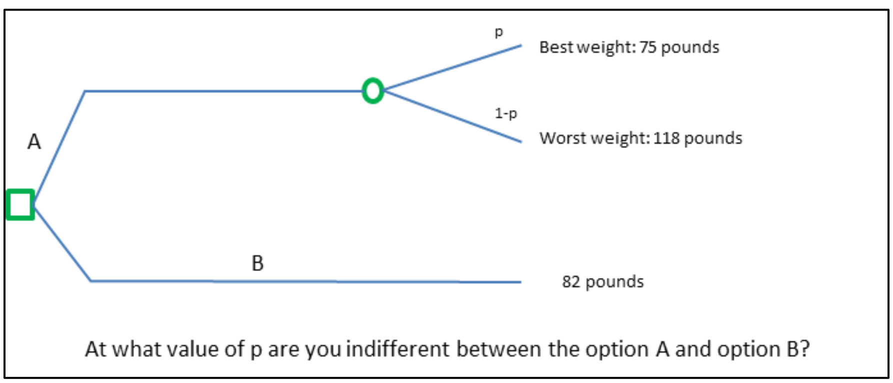


Figure 8. Probability equivalent method used to elicit utility score for the attribute of weight

The lottery probabilities (*p* value) were adjusted until the decision-maker was indifferent between the best and worst system and the guarantee of an 82 pound system. Similar to the first method, this process was repeated nine times (to include two iterations for filter life) with the same consistency check questions from the previous method after each round. In order to reduce bias, the order of attributes was randomized compared to the direct rating method. The full interview format is listed in Appendix C.

Elicitation of Stakeholders' Weights

Weighing was conducted to identify the range of importance for each attribute. Prior to any weighting interviews, the same operational scenario (Box 1) provided during the utility interviews was also provided for the weight interviews. This was to reduce the chance of stakeholders weighting attributes based on general importance instead of a specific situation. Weights were established using two different methods, SMARTS and SMARTER. The same experts in Table 4 were interviewed based on their knowledge and expertise in their respective disciplines and organizations.

Interviews were conducted over a five week period and were either conducted in person or over the phone depending on location of voters. Similar to the utility elicitation method, the order of attributes was randomized for each interview. Prior to starting the interviews, Edwards and Barron's 1994 swing weighting example of purchasing a car (Appendix C) was used to explain the swing weighting process to the stakeholders (Edwards & Barron, 1994).

Stakeholders were first provided a list of the nine attributes and their descriptions (Table 9) and had the opportunity to ask questions about any attribute before the interview started.

Table 9. Listing of attributes and descriptions provided to stakeholders prior to the weighting interview process

Attribute	Description
Chemical Reduction	Aequate technology to meet NSF/ANSI Protocols 53 and 58 for reduction of chemical contaminants
Microbial Reduction	Ability to meet NSF/ANSI Protocol 248 for reduction of microbiological contaminants
Weight	Total weight is < 80 pounds including all accessories for initial operation through 30 days
Filter Life	Filter has ablilty to treat enough water for 20 Soldiers in first 72 hours/30 days - based off 8.5 gal/day per Soldier: 170 gal/day/arid envir with 20 Soldiers for 3 days= 510 gallons (1931L); 30 days =5100 gallons(19310L)
Temperature Range	Ability to operate from from 30 - 120° F or 0 - 140 °F (optimal)
Disinfection Technology	Must not require follow-on disinfection for potability
Safety Features	Have an automatic shutoff at the end of element life to prevent production of contaminated water.
Monitoring	Ability to provide visual and audible monitoring feedback
Production Rate	Amount of fresh water provided measured in LPM and GPM. Based on 20 Soldiers, requirement is 120-160 GPD (455-606 LPD) using 6 GPD/8.5 GPD (23 LPD/32PLD) for arid and temperate environments.

The first step in the interview was to ask the respondents to rank each attribute compared to a hypothetical system that had the worst score in every attribute. This was illustrated using an excel spreadsheet that listed the worst scores and the best scores for each attribute. The worst scores were shaded to construct the hypothetical system (Table 10).

Table 10. Best and worst scores for each attribute illustrating hypothetical system that performed the worst in all attributes (shaded cells)

Attribute	Worst Scores	Best Scores
Weight	118 pounds	75 pounds
Safety Features (Auto shutdown)	Flow rate stops when filter is clogged	Two-stage protective shutdown circuit, automatically shuts down in the event of a failure in the disinfecting process, pump stops and a valve will close, stopping any more water from passing through the system.
Production Rate	.14 gallons per minute/ 200 gallons per day	1.95 gallons per minute/1,170 gallons per day
Operating Temperature	41 ºF -100.4 ºF	-22 ºF - 158 ºF
Monitoring	Pressure gage shows reduction in flow rate	Field integrity test (FIT) for filter devices; can test membrane at anytime
Microbial Reduction	99.99% reduction for bacteria and virus. No independent lab testing conducted.	Passed NSF P248 Protocol - 1 Mar 2013
Filter Life (measured by number of liters until filter change)	General water: 7,926 gallons Challenge water: 2,641	General water: 26,410 gal Challenge water:7,396 gal
Disinfection Technology	Does not hav UV, Requires disinfection	UV with optional chlorine injection
Chemical Reduction	• Multi-stage Cartridge and Carbon Filtration •2 Pre-filters (carbon core with nanomesh) • Main nanomesh filter	System meets NSF/ANSI 53 Standard for Health Effects
BENCHMARK (worst score in all attributes)		

Respondents were asked the following questions:

Given the above hypothetical system (shaded cells) that has the worst score in every attribute (highlights scores), if you could only choose one attribute to improve from worst to best , which one would it be (rank #1)? What is the next attribute you would improve from worst to best (rank #2)? Continue until all attributes are ranked from 1-9 (benchmark will rank 10th)

After the first step, interviewee's responses were read back to them to confirm their rankings. If at any point they wished to change their respsonse, they were given the chance. They were also asked to explain their reasoning on why they ranked attributes in the order they chose. Once the stakeholders confirmed their rankings were accurate, they moved onto the next step. If at any point during the process a stakeholder was confused, he was able to stop and ask questions to confirm his understanding of the process. The second step of the interview involved rating attributes. Stakeholders were asked to complete the following:

Rate your hypothetical system 0 and your top attribute 100. Based off the swing in each attribute from worst to best, rate each attribute against the top ranked attribute swing from step one. Rate other attributes between 0 and 100. Rating corresponds with % of value by changing each attribute.

Ex. Rating attribute 50 means improving attribute ranked from worst to best is worth 50% value from improving the #1 choice. *Comments: You may rate attributes the same if you feel the swings in attributes are equal when compared to the #1 attribute swing.*

The more indifferent you are between the worst/best score, the lower the rating would be. A higher rating signals a higher perceived significance of the attribute swing from low to high.

Similar to the previous step, once all scores were provided, scores were read back to the interviewee to confirm his ratings. Comparisons between scored attributes were made to ensure ratings accurately reflected the stakeholder's preferences. A consistency check was conducted using similar questions from the utility interviews previously listed. Once all ratings were confirmed, the interview was complete. The final step of normalizing the scores did not require stakeholder feedback. The weight interview format is listed in Appendix C.

Initially, only the SMARTS method was to be used, however, after piloting the interview, respondents found the swings in SMART difficult to understand. The interview format was revised to reduce confusion by providing Edwards and Barron's (1994) car purchasing example and the rank order centroid method (ROC) was implemented during data analysis. Because stakeholders ranked attributes in the first step of swing weighting, these rankings were used to calculate ROC scores. This additional step did not add any time to the interview and provided a consistency check on the responses, which were consistent between methods without any additional follow up required.

Throughout the decision making process, several steps were taken to reduce any misunderstandings, response mistakes, and biases that can occur through elicitation procedures, or psychological reasons(Marttunen & Hämäläinen, 2008). The value tree was structured simplistically, without multiple levels of attributes, to reduce the chance of splitting bias, which occurs when an attribute is weighted more when it is split into sub attributes(Hämäläinen & Alaja, 2008). Prior to conducting any interviews, the utility and weight questionnaires were piloted on two individuals; one familiar with water purification, and one with limited knowledge on the research topic. This was to ensure both the process and system specifications were easy to understand. Participants were provided with instructions on how the interviews would occur and a detailed operational scenario to reduce the chance of general importance weights before starting utility and weight interviews. During the utility interviews, index cards were utilized to simplify the process and help the voter visualize his options. The order of the attributes was randomized for each interview, and a consistency check was conducted by asking for voter explanations and confirming his responses. Each time an inconsistency was identified, the voter was asked to reconsider his response, which resulted in modified results.

Aggregate Utility and Weight Scores of Each Alternative

After obtaining utility and weight measurement scores, they were aggregated using the linear additive model (Equation 1).

$$U(alternative) = U(x_1, x_{2,} \dots x_n,) = \sum_{i=1}^{n} k_i U_i(x_i) \tag{1}$$

where

$$k_i = weight\ of\ ith\ attribute$$

$$U_i(x_i) = utility\ score\ for\ ith\ attribute$$

$$U(x_i) = overall\ utility\ of\ ith\ alternative$$

Inserting individual attribute utility functions for the attributes defined in Table 7 into Equation 1 yields:

$$U(alternative) = k_{chem} * U_{chem}(x_{chem}) + k_{micro} * U_{micro}(x_{micro}) + k_{weight} * U_{weight}(x_{weight}) + k_{filter} * U_{filter}(x_{filter}) + k_{temp} * U_{temp}(x_{temp}) + k_{production} * U_{production}(x_{production}) + k_{disinf.} * U_{disinf.}(x_{disinf.}) + k_{monitor} * U_{monitor}(x_{monitor}) + k_{safety} * U_{safety}(x_{safety})$$

This aggregation provides each alternative with a score, with the highest score corresponding to the recommended alternative. All scores are shown in the results section and Appendix F.

4. RESULTS

As described in the methods section, this research compared two different methods for eliciting individual attribute utility scores--that is, the relative value to the decision-maker of one level of an attribute (for example, as weight of 75 pound) as compared to another level (for example, a weight of 118 pounds). The research also compared two different methods for estimating the weights for each attribute—that is, the willingness of the decision-maker to trade an option that scores high along attribute *X* but low on attribute *Y* for an option that scores high on *Y* but low on *X*. If different elicitation methods lead to different individual attribute utility scores and different attribute weights, then the different methods potentially could lead to conflicting conclusions about which option best reflects the decision-maker's preferences. This section first describes the results of the two different methods for eliciting individual attribute utilities, then presents weights as elicited using two different methods, and finally computes total utility scores for the four water treatment technologies using different combinations of the different weights and utilities. The key finding is that regardless of method choice, the most preferred technology for six of the seven stakeholders interviewed in this research is the Nephros MSU ultra filter™ configured with an Aquamira DIVVY50™ water pump. For the seventh decision-maker, changing from one weight elicitation method to another resulted in a slight change in preference ordering of the technologies, with the Nephros system scoring second highest, after the Aspen system.

Utility Scores Using Two Different Methods

The direct rating and probability equivalent methods for eliciting utility functions for the individual attributes yielded very similar results (Table 11). On average, the difference between individual utility scores for each attribute between methods was eleven points. The largest difference between methods occurred for the filter life (30 points), microbial reduction (25 points), and disinfection technology (25 points) attributes; the higher scores were from the probability equivalent method (PEM), which takes risk into consideration. The higher scores identified the decision-maker's aversion to the gamble between best and worst. The expert valued the intermediate outcomes more in the scenario with uncertainty compared to directly rating the systems under certain conditions. Differences in utility scores for all other attributes were minimal.

Table 11. Utility scores using direct rating method and probability equivalent method (PEM)

Utility scores by system and method								
Attributes	**Seldon Waterbox**		**SLMCO FBS 180**		**Aspen 1800 BC**		**Nephros MSU**	
	Direct	PEM	Direct	PEM	Direct	PEM	Direct	PEM
Chemical Reduction	0	0	40	52	60	60	100	100
Microbial Reduction	100	100	75	50	0	0	100	100
Weight	80	60	40	52	0	0	100	100
Filter Life	0	0	75	88	100	100	58	88
Operating Temperature Range	0	0	29	29	71	71	100	100
Production Rate	29	43	0	0	57	57	100	100
Disinfection Technology	0	0	100	100	50	75	25	50
Safety Features	17	33	33	50	100	100	0	0
Monitoring Capability	0	0	50	50	33	33	100	100
Total	225	236	442	470	472	497	683	738
System Ranking	4	4	3	3	2	2	1	1

Weighting Scores

Swing Weighting Method

On average, stakeholders weighted each attribute almost equally, as shown by the small and nonsignificant differences among median values of weights for attributes in Figure 9.

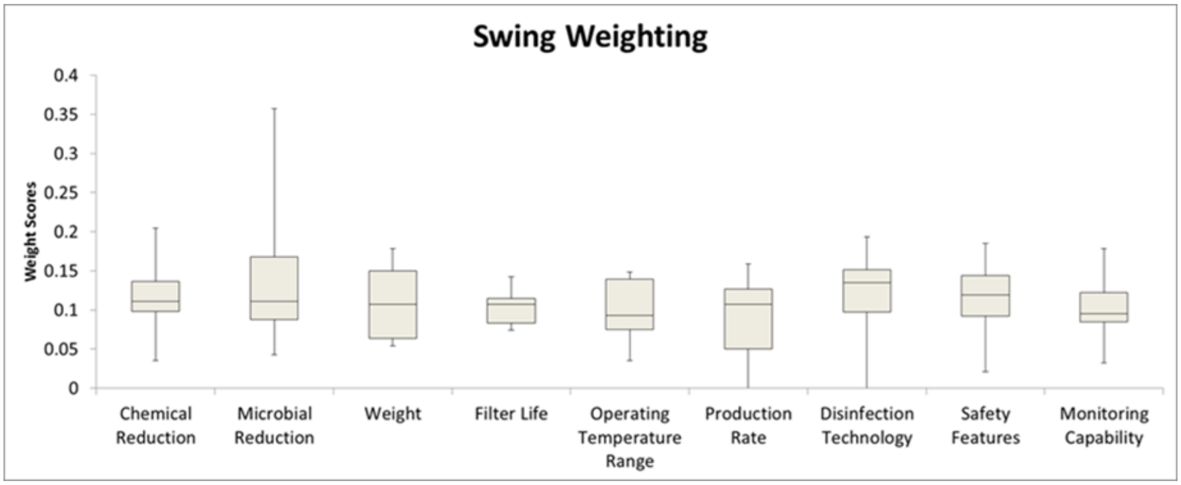


Figure 9. Weights for the nine attributes important in selecting water purification technologies as elicited from seven stakeholders using the swing weighting method. The horizontal bars show the median across all seven stakeholders; the shaded boxes show the interquartile range; and the end points of the vertical bars indicate the minimum and maximum weights.

There is only a 4 percentage point difference between the median weights of all nine attributes. On the other hand, the large ranges in weights for the microbial reduction and disinfection technology attributes (illustrated by the length of the vertical bars) shows disagreement among stakeholders. The difference in weights assigned to microbial reduction could reflect skepticism among some stakeholders of manufacturers' claims about microbial reduction capabilities. Stakeholder agreement was highest in weighting the importance of filter life, followed by operating temperature range, and weight.

The minimum attribute weights for production rate and disinfection technology were zero, signifying that some stakeholders viewed these attributes as unimportant in selecting among the four candidate technologies. One stakeholder selected a weight of zero for production rate because even the least desirable system still met daily requirements. Other stakeholders said that production rate was more significant because they valued the speed at which daily water requirements were met and anticipated additional water requirements in the future. Some stakeholders assigned zero weight to the disinfection technology attribute because Army doctrine dictates, that regardless of disinfection technology in commercial-off-the-shelf systems, commanders should require additional disinfection chemicals, (U.S. Dept of the Army, 2010). Other voters assigned a higher weight to disinfection technology due to their preference for multiple barriers against microbial contamination.

Rank Order Centroid Weighting Method

The rank order centroid (ROC) weighting method produced more variation in median weights across attributes and in individual stakeholder weighting within single attributes (Figure 10). On average, stakeholders assigned the highest weight to disinfection technology, followed by safety features. The range in weights among stakeholders was largest for the safety features attribute. High variation among stakeholders was also observed for the microbial reduction, weight, and production rate attributes. For each of these four attributes, at least one voter assigned the highest weight among attributes while another assigned it the lowest weight.

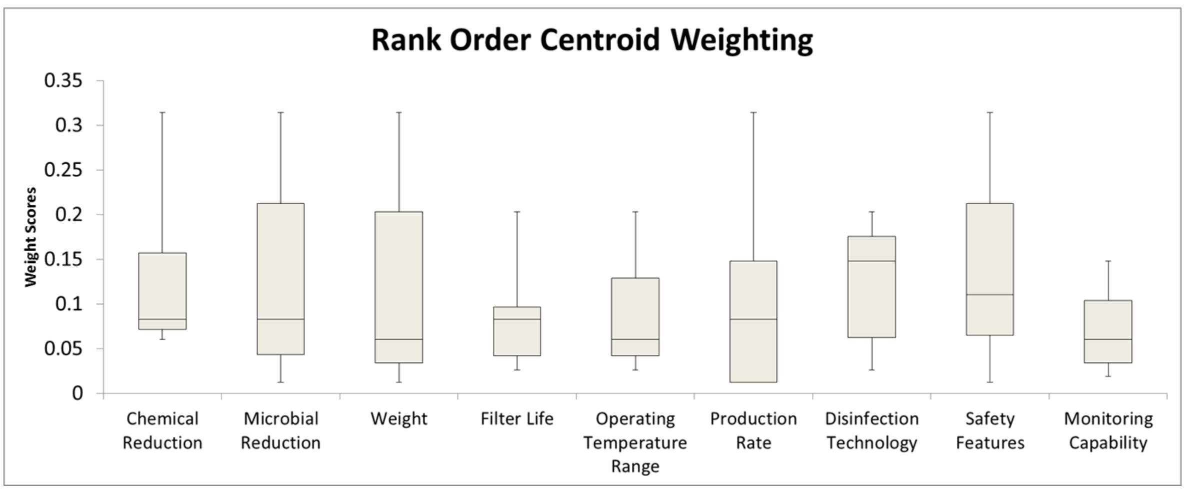


Figure 10. Weights for the nine attributes important in selecting water purification technologies as elicited from seven stakeholders using the rank ordered centroid weighting method. The horizontal bars show the median across all seven stakeholders; the shaded boxes show the interquartile range; and the end points of the vertical bars indicate the minimum and maximum weights

As the lengths of the vertical bars in the above figures show, in terms of weighting, there is not a strong consensus among stakeholders. The swing method identified disinfection technology as the most important attribute when considering the median rating across stakeholders, but there is only a four percentage-point difference between the median values of the most and least important attributes.

The ROC method also resulted in disinfection technology having the highest median weight, with a smaller range between maximum and minimum weights. However, in general, the ROC method produced more disagreement among stakeholders than the swing weighting method, as illustrated by the width of the vertical bars in Figure 10. Appendix D provides a detailed list of weights by voter and attribute.

Aggregated Multi-Attribute Utility Scores

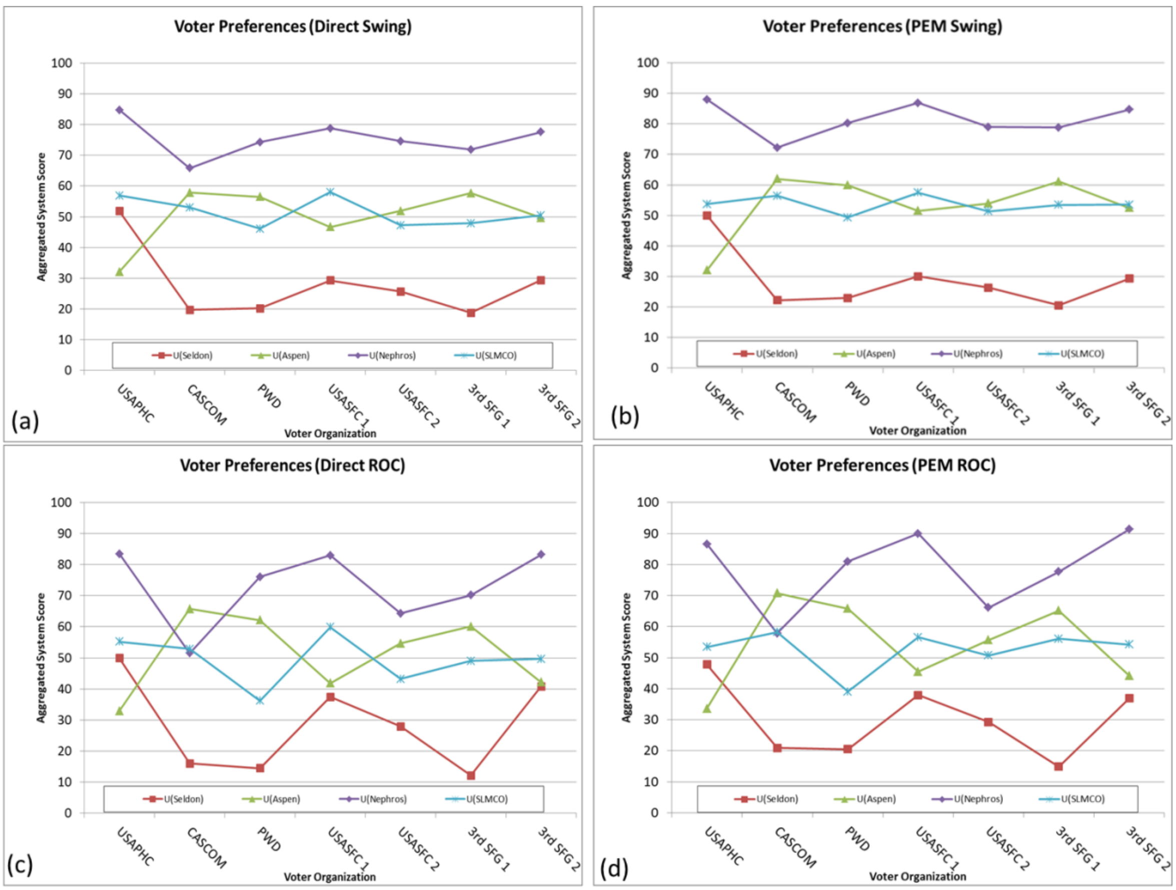


Figure 11. Overall values of the alternatives for each interviewed stakeholder based on utility elicitation methods of direct rating (a,c) and probability equivalent (b,d) and weighting methods of rank order centroid (c-d) and swing (a-b).

The Nephros MSU Ultra filter system had the highest score from all voters in all methods with the exception of the multi-attribute utility function that used the CASCOM voter's weights as elicited using the ROC method. This latter result was caused by a higher prioritization of safety features by the

CASCOM voter compared to other voters. However, all voters preferred the Nephros system when the swing weighting method was used to elicit weights among attributes, regardless of the individual attribute utilities (Figures 11a and 11b).

Utility scores between the PEM and direct rating method were similar and did not cause significant difference in overall scores. The weighting methods caused the most disparity between aggregated scores due to the voters' ability to assign weights in the swing method compared to already established ROC weights based on ranking (Figures 11c and 11d). The drop in score for the Nephros system when using ROC weights elicited from the CASCOM voter was due to the CASCOM voter's high preference for safety features and the comparatively low utility score of the Nephros MSU ultra filter system along this attribute.

Similar to the top alternative, the Seldon system was the least preferred system from all voters with the exception of one. The higher score for the USAPHC voter was due to the relatively high weight the USAPHC voter assigned to the microbial reduction attribute, an attribute on which the Seldon system performed comparatively well. The Seldon system's lower utility scores in production rate and safety features—priorities for PWD, CASCOM, and 3rd SFG 1 voters—kept overall scores for this system lower than the alternatives. Appendix E provides a detailed list of aggregated system scores by voter and method.

5. DISCUSSION

In summary, regardless of method or voter background, the Nephros MSU Ultra filter system was preferred in almost all scenarios, establishing a consistency between methods. These results identify not only a clear leading candidate among the four technologies but also provide evidence that the MADA process used in this research is robust against methods for eliciting individual attribute utility functions and attribute weights.

Comparison with Previous MADA Studies

The results of this research are consistent with other MADA studies finding that decision-makers' preferences generally do not vary significantly when elicited using different MADA techniques (Huang et al., 2011). In their recent systematic review of MADA applications between 1990 and 2010, Huang et al. found 20 papers (out of 312 total) that compared different MADA methods. From this review, they concluded that regardless of which multi-attribute decision analysis method was used that the top alternatives were the same. Furthermore, they noted that in the few cases where different methods yielded different rankings of decision options, there were still significant overlaps in the top few alternatives. Huang et al. recommended further research to determine which MADA approaches are most appropriate for different kinds of decision problems. The results in this thesis lend further support to Huang et al.'s conclusion that the selection of a preferred alternative is generally robust across MADA methods(Huang et al., 2011).

Furthermore, as in many previous applications of MADA techniques, implementation of the MADA process revealed agreement on the best decision option, even when there are a large number of stakeholders with different backgrounds (Karjalainen et al., 2013; Marttunen & Hämäläinen, 2008). In a case study on the use of formal decision analysis methods in an aquifer land use problem, Karjalainen et al. confirmed that MADA processes are beneficial in learning and collaboration when different interests are represented. Karjalainen et al. used a participatory method with nineteen stakeholders from diverse backgrounds to structure the value tree and an interactive interview process to elicit attribute weights for a MAUT function. All stakeholders agreed on the most important criteria during the weighting process, and all preferred the same alternative when weights and utilities were aggregated. This study revealed that even with various interests and priorities, stakeholders were still able to agree on many critical issues. The agreement among seven stakeholders representing diverse backgrounds in this thesis confirms the findings in Karjalainen et al.'s study and furthers support for the usefulness of MADA in facilitating stakeholder involvement through MADA methods.

Similar agreement among stakeholders with varying backgrounds was also identified in a study conducted by Marttunen and Hämäläinen (2006). They conducted decision analysis interviews with twenty stakeholders in a water course regulation project. Again, stakeholders with different backgrounds were able to agree on the priority of objectives and all stakeholders were willing to approve the outcome, even though the final recommendation was not the initial alternative that some stakeholders had wanted. Marttunen and Hämäläinen concluded that decision analysis interviews improve the quality and efficiency of the planning process, and collaborative meetings involving all stakeholders lead to a consensus in a group with strong interests. Although this thesis used a similar interview process, it did not conduct collaborative meetings with all stakeholders due to the locations and schedules of the stakeholders. However, this thesis did have similar results in terms of stakeholder agreement on the same outcome.

Limitations

The major limitations of this study are the reliance on one expert for eliciting individual attribute utility scores for each technology and the exclusion of cost information.

Time and resource constraints necessitated relying on one expert (the SOF officer who served multiple tours in Afghanistan) for the elicitation of individual attribute utility functions. Elicitation of these functions took two hours, and the other stakeholders faced time constraints that prohibited elicitation of their single-attribute utility functions. Weighting interviews lasted between one to two hours for each stakeholder; therefore eliciting all utilities from a single expert reduced the time requirements of stakeholders. Nonetheless, the SOF officer from whom individual attribute scores were elicited had the most relevant experience among all stakeholders, so his preferences along individual attributes likely best reflect the preferences of the SOF units for whom the water purification technologies are intended. Future studies could elicit individual attribute utility functions from multiple stakeholders, in order to gauge whether the identification of a preferred technology would change as a result. Given the consistency of the results presented here across utility and weight elicitation methods, it is unlikely that the preferred alternative would change if individual attribute utility scores were elicited from all stakeholders.

The second limitation was the lack of information on costs for military purchase of the water purification units assessed in this research. Unfortunately, the manufacturers of these units were unable to provide information on costs that would be charged to the military for procurement of these systems. Furthermore, unit costs would depend on the number of units purchased. Once cost information becomes available, if the Nephros system is more costly than other alternatives, the elicitation of weights for the attributes could be re-done, and cost information could be included in the analysis.

6. CONCLUSION

This research applied multi-attribute decision-making methods to successfully identify a preferred commercial-off-the-shelf water purification system for use in a Special Operations Forces environment. Using the multi-attribute utility theory, with two utility and two weight elicitation methods, consistently identified the same top alternative for six out of the seven stakeholders interviewed. Furthermore, the seventh stakeholder ranked the preferred technology identified by the other six stakeholders as best under one weighting scheme and as second-best under the alternative weighting scheme. Hence, the multi-attribute utility method was robust against elicitation method in identifying stakeholder preferences and also revealed strong stakeholder agreement on the best water purification technology from among the four evaluated, despite differences in stakeholder opinions about which attributes were more or less important than others. The framework applied in this work was simple for stakeholders to understand and can be applied to future water purification system decisions by adjusting the operational scenario or attributes.

This framework used two methods (the direct rating and probability equivalent methods) for eliciting utility scores for each attribute and two methods (swing weighting and the ROC method) for eliciting attribute weights. Both utility elicitation methods yielded similar results, were easy to implement, and were well understood by the participants. The only major difference between the methods was that the probability equivalent method better reflected the decision-maker's risk aversion, eliciting utility scores that were slightly higher for intermediate outcomes than the direct weighting method and hence revealing the decision-maker's preference for avoiding gambles between the best and worst systems. However, this slight score increase was not enough to affect the overall results between methods.

Similarly, the elicitation of weights from seven voters with various levels of experience and knowledge provided similar results regardless of the elicitation method. Disinfection technology was the most important attribute using both swing and ROC weighing. The ROC method provided a larger range in weight scores compared to the swing method. The swing method more accurately portrayed voters' preferences by enabling voters to provide insight to determine weights for each attribute instead of using pre-calculated weights based on the voter's ordinal ranking of attributes.

Due to the similarities between methods, future use of this model does not need to include two methods for both utility and weight elicitation. Each round of utility interviews is time consuming; therefore, if time is limited, the use of only one utility scoring method will suffice. The selected method should be based on the decision problem, type of data, and level of risk or uncertainty in the decision outcome. Generally speaking due to the lack of insight from the ROC method, the swing method should be used to ensure voter risk tolerances are accurately weighted. However, the ROC method may provide a consistency check on weights determined with the swing weighting method.

This research addressed specific SOF requirements identified by key personnel, was easy to understand for voters, and can be implemented for future small-unit water purification decisions. This framework can be easily adapted to the variety of missions that SOF units must undertake and the variety of environments in which they must operate. Furthermore, it is less time consuming than current studies, which last over twelve months. This method provides a complementary approach to current technology selection approaches, which require substantial investments of time to test performance of purification systems in Army laboratories. The MADA approach could narrow the list of candidate technologies for testing, hence saving on costs of testing multiple different technologies. In addition, although this research focused on the SOF environment, this model can be applied in various scenarios with conflicting objectives and requirements. It provides a rational approach that can be used regardless of military background or familiarity with MADA techniques.

APPENDIX A: INTERVIEW CONSENT FORM/ UTILITY AND WEIGHT INTERVIEW FORM

Interview Consent Form

You have been asked to participate in a study that compares commercial-off-the-shelf water purification units for use in U.S. Army special operations. This research project will comprise my Master's paper, written to fulfill requirements for the Master's degree in Environmental Science at the University of North Carolina at Chapel Hill. The purpose of this study is to establish an evaluation protocol in selecting a small unit water purifiers. Your participation will consist of one interview, approximately an hour in length. If you are willing, your participation could also include follow-up questions by phone or email. If you agree, I will make notes of this interview. I plan to interview seven subject matter experts during this study.

I am the only person who will have access to data associated with your name. I am not aware of any risks that would result from your participation in this study. You are free to withdraw from the study at any time without penalty. You may also choose to not answer specific question and still continue to participate.

Please feel free to contact me, Lauren Koban (724.312.7703; Koban@live.unc.edu) or Dr. Jacqueline MacDonald Gibson (919-966-7892 or jackie.macdonald@unc.edu), my faculty advisor, at any time if you have questions about this study.

Please contact the UNC-CH Academic Affairs Institutional Review Board at (919) 962-7761 or aa-irb@unc.edu if you have any questions about your rights as a research participant.

Please sign and date this form to indicate that you agree to participate in this study, and keep one copy for your records.

Participant's signature　　Date　　Participant's printed name

APPENDIX B: UTILITY INTERVIEW FORMAT

Decision-maker Interview
Eliciting Single Attribute Utilities

The purpose of this interview is to elicit a utility function for each attribute in this evaluation model. Each utility will be determined based on scores you provide for each system within each attribute, meaning you will rank each of the 5 systems 12 times (one for each attribute).

Two methods will be used, the direct rating scale method and the probability equivalent method. We will work through each attribute one at a time. I will provide flashcards for each system for each iteration. You can write your scores on the flashcards. After you initially rank and score the systems, I will ask you a series of questions as a consistency check to ensure you feel comfortable with your rankings. At the end I will review your scores for attribute again to ensure your preferences are accurately reflected in the utility functions.

For any attribute that all 4 systems score in a way there is no discrimination between systems, that attribute can be eliminated from the model.

For the direct rating scale, you will always give the most desirable system a score of 100 and the least desirable system a score of 0. The 2 intermediate systems will then be scored between 0-100.

For the probability equivalent method, I will ask you your preference between having a definite system trait or a gamble between the most and least desirable trait within each attribute. You will be able to modify the probability in the reference gamble until you are indifferent between the sure bet or a gamble.

The first attribute will be the weight of the system. The weight of the system can be defined as the total weight to include all accessories required to operate for 30 days.

Rating Scale
High/most desirable Score: 100
Lowest Score/least desirable: 0

v(72 lbs) = 100
v(76 lbs) =
v(79) =
v(92) =
V (118 lbs) = 0

1. Select the most and least desirable systems.
2. Rank remaining options between best and worst.
3. Rate the 4 options between the values of 0 - 100 based on relative value so that the relative spacing between systems reflects strength of preference for one system over another.

Consistency Check:
1. Confirm whether value steps between each system are equal (if the difference between any 2 systems are close to equal.

2. Is the difference between 1 and 2 truly larger/smaller than 2 and 3, 3 and 4, or 4 and 5? (wording will be adjusted based on rankings set by decision-maker)
    a. Example question: You have rated system B halfway between system C and D. Is this correct?
3. Add in additional values for attributes (ex. Weight, 80, 95, 110 lb) and see where new values are ranked. Confirm they are ranked between other similar weight values, and the difference between weights accurately reflects strength of preference.

* The above questions will be reiterated 9 more times for each attribute. Each will be defined, and then how each attribute is measured. Flashcards for each system will be laid on the table to provide a visual aid for the decision-maker for him to rank and score.

Probability Equivalent Method

I am now going ask you a different question for the middle scores for each of the attributes addressed in the direct rating method. Based on your high/low scores for each system, we are now going to compare the middle scores against the probabilty of either the best or worst outcome.

I will ask you your preference between having a definite system trait or a gamble between the most and least desirable trait within each attribute. This process will be repeated for each attribute.

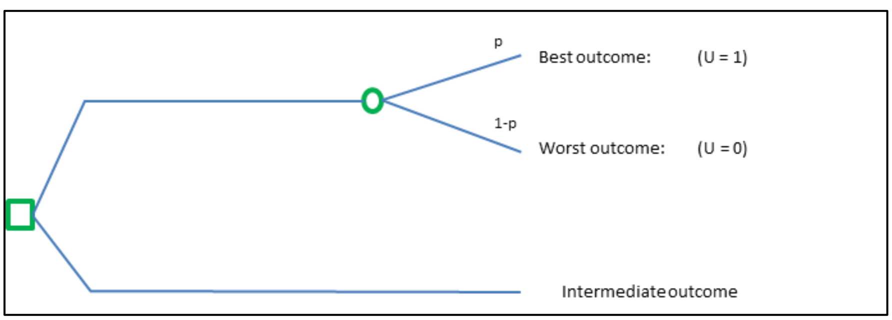


Figure 12. Probability equivalent reference lottery

** Order of attributes will be different than order during the direct rating method*

APPENDIX C: WEIGHT INTERVIEW FORMAT

Subject Matter Expert (SME) Interview
Accessing Weights

The purpose of this interview is to assess weights for each attribute to evaluate small unit water purifiers. The two methods used will be direct weighting and the swing weighting approach where I will ask you to compare each attribute directly through hypothetical outcomes. The benchmark system for comparison in this scenario is a SUWP that ranks the worst in all attributes. Each row in the table below swings a different attribute from worst to best (Comparison values are subject to change based on responses in decision-maker's utility elicitation interview).

The following steps will be followed using the table to assess weights.

Swing Weighting

1. **Rank Attributes**: Given a hypothetical system (benchmark) that has the worst score in every attribute (highlights scores), if you could only choose one attribute to improve from worst to best (or swing), which one would it be (rank #1)? What is the next attribute you would improve from worst to best (rank #2) ? Continue until all attributes are ranked from 1-9 (benchmark will rank 10th) Example: Weight swing is the perceived value of reducing weight from 118 to 75 pounds
2. **Rate Attributes**: Rate your benchmark 0 and your top attribute 100. Based off the swing in each attribute from worst to best, rate each attribute against the top ranked attribute swing from step one. Rate other attributes between 0 and 100. Rating corresponds with % of value by changing each attribute. Ex. Rating attribute 50 means improving attribute ranked from worst to best is worth 50% value from improving the #1 choice. *Comments: You may rate attributes the same if you feel the swing in attributes is equal when compared to the #1 attribute swing. The more indifferent you are between the worst/best score, the lower the rating would be. A higher rating signals a higher perceived significance of the attribute swing from low to high.*

3. **Calculate weights**: Normalize weights

** The order of the attributes will be randomized for each interview*

Table 12. Listing of attributes and descriptions provided to stakeholders prior to the weighting interview process

Attribute	Description
Chemical Reduction	Aequate technology to meet NSF/ANSI Protocols 53 and 58 for reduction of chemical contaminants
Microbial Reduction	Ability to meet NSF/ANSI Protocol 248 for reduction of microbiological contaminants
Weight	Total weight is < 80 pounds including all accessories for initial operation through 30 days
Filter Life	Filter has ablilty to treat enough water for 20 Soldiers in first 72 hours/30 days - based off 8.5 gal/day per Soldier: 170 gal/day/arid envir with 20 Soldiers for 3 days= 510 gallons; 30 days =5100 gallons
Temperature Range	Ability to operate from from 30 - 120° F or 0 - 140 °F (optimal)
Disinfection Technology	Must not require follow-on disinfection for potability
Safety Features	Have an automatic shutoff at the end of element life to prevent production of contaminated water.
Monitoring	Ability to provide visual and audible monitoring feedback
Production Rate	Amount of fresh water provided measured in LPM and GPM. Based on 20 Soldiers, requirement is 120-160 GPD using 6 GPD/8.5 GPD for arid and temperate environments.

Given a choice between 4 cars with 4 attributes:

Example taken from SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement, (Edwards and Barron, 1994)

	Attributes			
Cars	*Power*	*Shop Trips*	*Crusable Steel*	*Styling*
Anapest	100	90	0	0
Dactyl	0	100	90	70
Iamb	70	40	100	40
Trochee	50	0	40	100

Step 1: Rank the order of weights	Step 1 - Rank the order of weights: Imagine there was a 5th car that you were required to buy. This model scored a 0 in all 4 attributes. However, you were given the option to change just one attribute from worst to best, which attribute would you chose to improve? If the respondant chose power. The respondant now has the option to improve any dimesion except power from worst to best. What would it be? They chose shop trips.
Step 2: Rate the weights	Step 2 - Rate the weights: Lets call the weight of power the most important attribute, 100. That is a swing from 0 to 100 is worth 100 points. Consider the weight of something not important like the size of an ashtray. A 100 point swing on that attribute won't matter and is rated a 0. If trips to the shop is consdiered the second most important attribute, then the question would be what is the weight of a 100 point swing on the second most important dimension? Rating it a 50 would mean that swing is worth 50% in value of the power swing from 0 to 100.

Figure 13. Swing Weighting Example from SMARTS and SMARTER: Improved Simply Methods for Multiattribute Utility Measurement (Edwards and Barron, 1994)

APPENDIX D: WEIGHT SCORES

Table 13. Elicited scores from stakeholders using the swing weighting method

Objectives	**Swing Weighting Scores**						
m_i	k_i						
	USAPHC	**CASCOM**	**PWD**	**USASFC 1**	**USASFC 2**	**3rd SFG 1**	**3rd SFG 2**
Chemical Reduction	0.0357	0.1296	0.1111	0.2043	0.0960	0.1429	0.1009
Microbial Reduction	0.3571	0.1111	0.0794	0.2151	0.0960	0.0429	0.1211
Weight	0.1786	0.0556	0.0714	0.0538	0.1440	0.1071	0.1553
Filter Life	0.1071	0.0741	0.0873	0.1075	0.0800	0.1214	0.1429
Temperature Range	0.0357	0.1481	0.1429	0.0645	0.1360	0.0929	0.0854
Production Rate	0.0000	0.0370	0.1587	0.1075	0.0640	0.1286	0.1242
Disinfection Technology	0.0000	0.1667	0.1349	0.1935	0.0800	0.1357	0.1149
Safety Features	0.1071	0.1852	0.1190	0.0215	0.1600	0.1286	0.0776
Monitoring	0.1786	0.0926	0.0952	0.0323	0.1440	0.1000	0.0776

Table 14. Elicited scores from stakeholders using the rank order centroid weighting method

Objectives	**Rank Order Centroid Weighting Scores**						
m_i	k_i						
	USAPHC	**CASCOM**	**PWD**	**USASFC 1**	**USASFC 2**	**3rd SFG 1**	**3rd SFG 2**
Chemical Reduction	0.0606	0.1106	0.0828	0.2032	0.0828	0.3143	0.0606
Microbial Reduction	0.3143	0.0828	0.0262	0.3143	0.0606	0.0123	0.1106
Weight	0.2032	0.0262	0.0123	0.0421	0.2032	0.0606	0.3143
Filter Life	0.0828	0.0421	0.0421	0.1106	0.0262	0.0828	0.2032
Temperature Range	0.0421	0.1477	0.2032	0.0606	0.1106	0.0262	0.0421
Production Rate	0.0123	0.0123	0.3143	0.0828	0.0123	0.1477	0.1477
Disinfection Technology	0.0262	0.2032	0.1477	0.1477	0.0421	0.2032	0.0828
Safety Features	0.1106	0.3143	0.1106	0.0123	0.3143	0.1106	0.0193
Monitoring	0.1477	0.0606	0.0606	0.0262	0.1477	0.0421	0.0193

APPENDIX E: AGGREGATED SYSTEM SCORES BY VOTER AND ELICITATION METHOD

Table 15. Aggregated scores by elicitation method and voter: utility elicitation methods of direct rating (a-b) and probability equivalent (c-d) and weighting methods of rank order centroid (b,d) and swing (a,c).

Direct Swing				
Voter	U(Seldon)	U(Aspen)	U(Nephros)	U(SLMCO)
USAPHC	52	32	85	57
CASCOM	20	58	66	53
PWD	20	56	74	46
USASFC 1	29	47	79	58
USASFC 2	26	52	75	47
3rd SFG 1	19	58	72	48
3rd SFG 2	29	50	78	50

PEM Swing				
Voter	U(Seldon)	U(Aspen)	U(Nephros)	U(SLMCO)
USAPHC	50	32	88	54
CASCOM	22	62	72	56
PWD	23	60	80	49
USASFC 1	30	52	87	57
USASFC 2	26	54	79	51
3rd SFG 1	21	61	79	53
3rd SFG 2	29	53	85	54

Direct ROC				
Voter	U(Seldon)	U(Aspen)	U(Nephros)	U(SLMCO)
USAPHC	50	33	83	55
CASCOM	16	66	52	53
PWD	14	62	76	36
USASFC 1	37	42	83	60
USASFC 2	28	55	64	43
3rd SFG 1	12	60	70	49
3rd SFG 2	41	42	83	50

PEM ROC				
Voter	U(Seldon)	U(Aspen)	U(Nephros)	U(SLMCO)
USAPHC	48	34	87	53
CASCOM	21	71	58	58
PWD	21	66	81	39
USASFC 1	38	45	90	57
USASFC 2	29	56	66	51
3rd SFG 1	15	65	78	56
3rd SFG 2	37	44	91	54

APPENDIX F: AGGREGATED SCORES BY METHOD, ATTRIBUTE, AND VOTER

F.1 Utility Scoring: Probability equivalent method
Weight Scoring: Rank order centroid

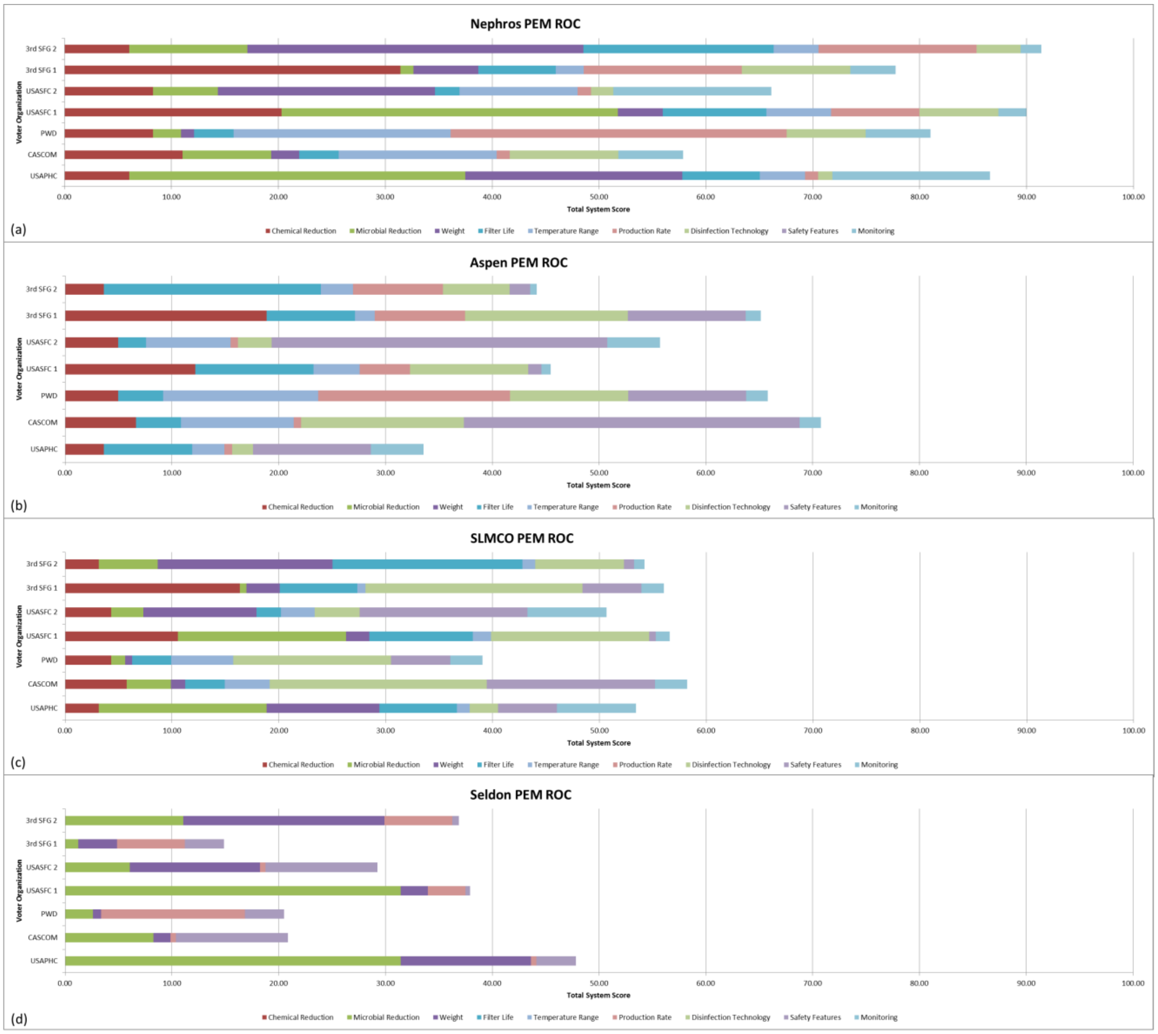


Figure 14. Bar chart illustrating total aggregated system scores by attribute for the probability equivalent and rank order centroid weighting methods for the following systems: (a) Nephros MSU ultra filter, (b) Aspen 1800 BC, (c) SLMCO FBS 180, and (d) Seldon Waterbox

F.2 Utility Scoring: Probability equivalent method
Weight Scoring: Swing method

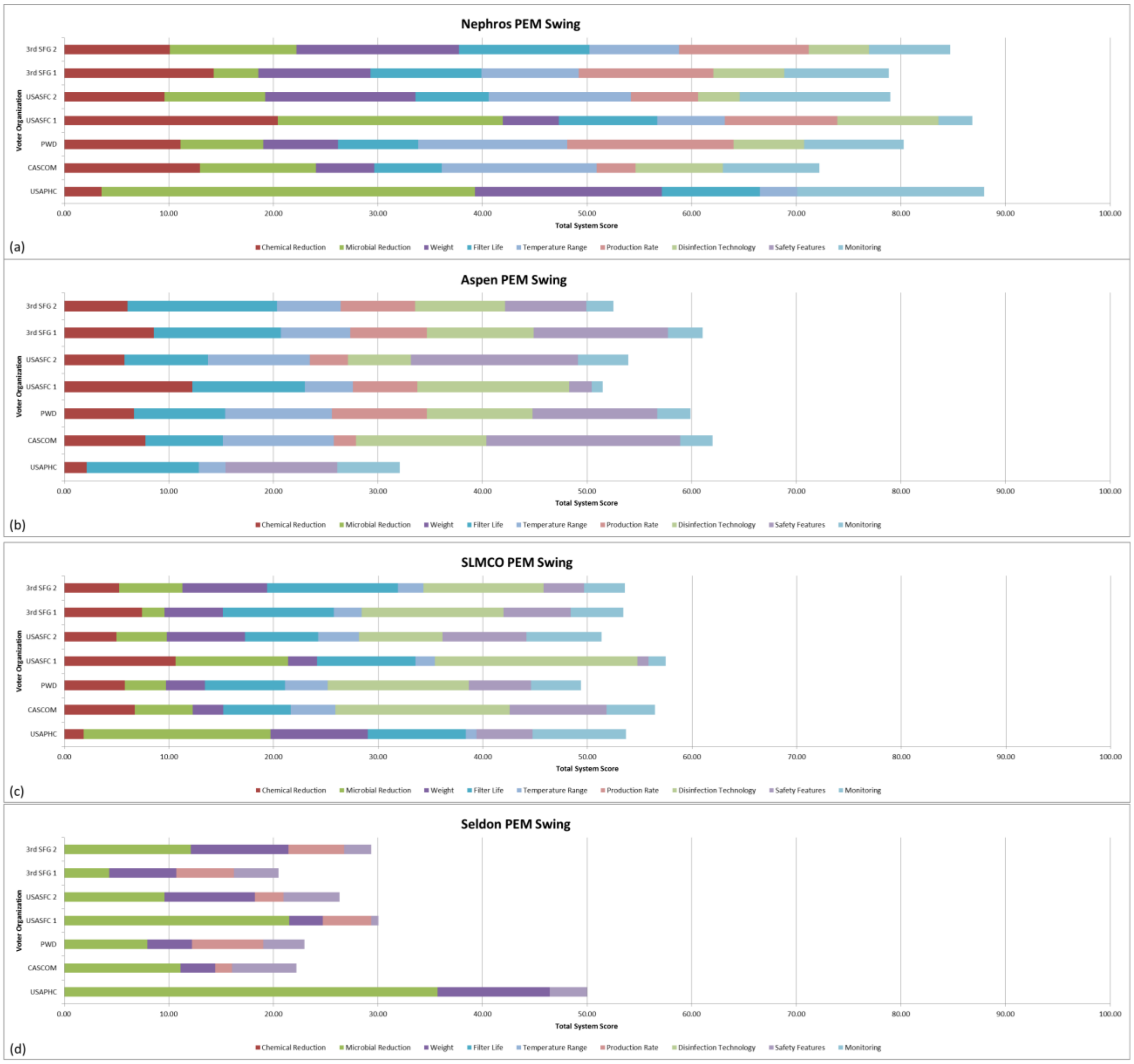


Figure 15. Bar chart illustrating total aggregated system scores by attribute for the probability equivalent and swing weighting methods for the following systems: (a) Nephros MSU ultra filter, (b) Aspen 1800 BC, (c) SLMCO FBS 180, and (d) Seldon Waterbox

F.3 Utility Scoring: Direct rating method
Weight Scoring: Swing method

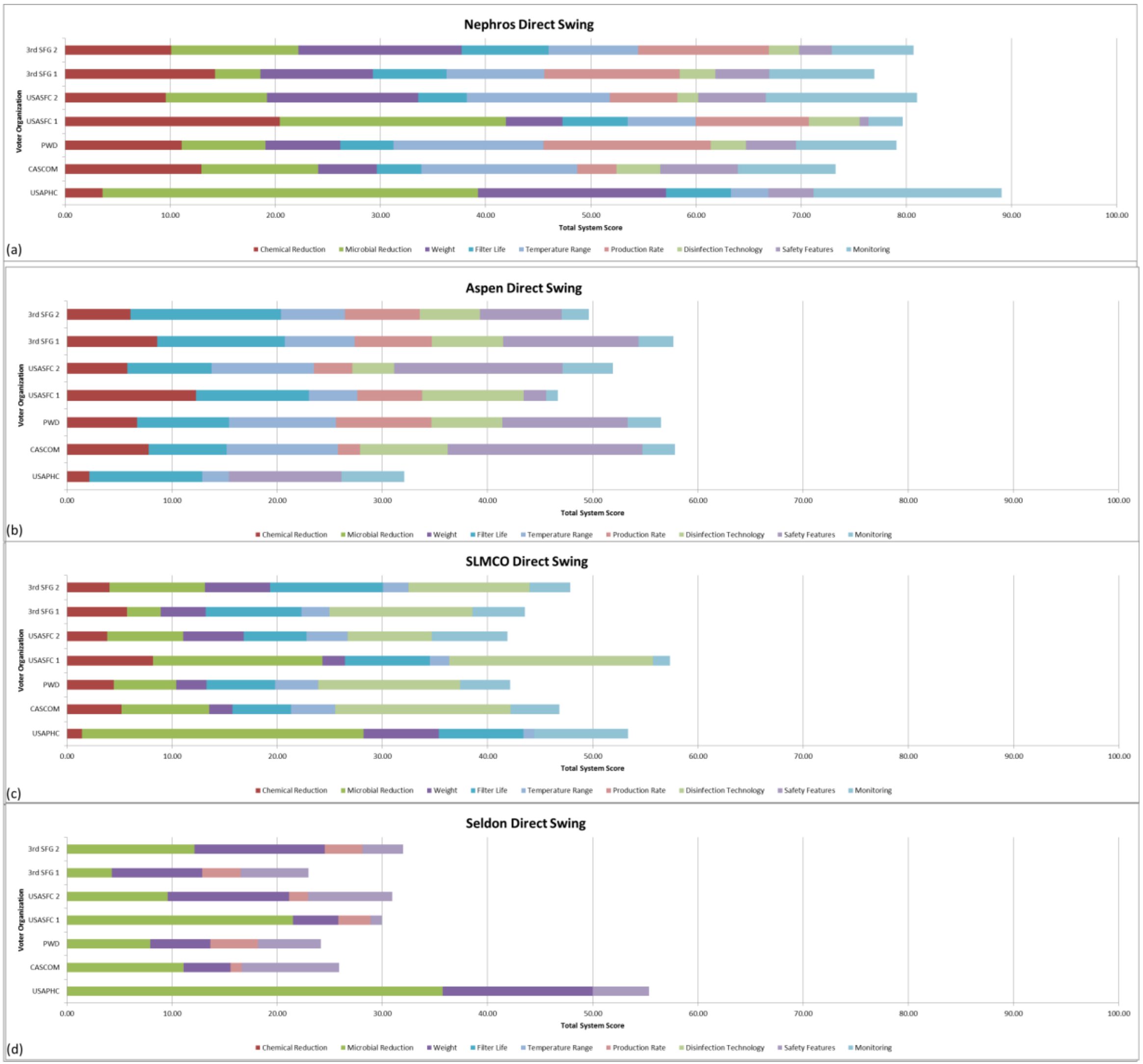


Figure 16. Bar chart illustrating total aggregated system scores by attribute for the direct rating and swing weighting methods for the following systems: (a) Nephros MSU ultra

F.4 Utility Scoring: Direct rating method
Weight Scoring: Rank order centroid method

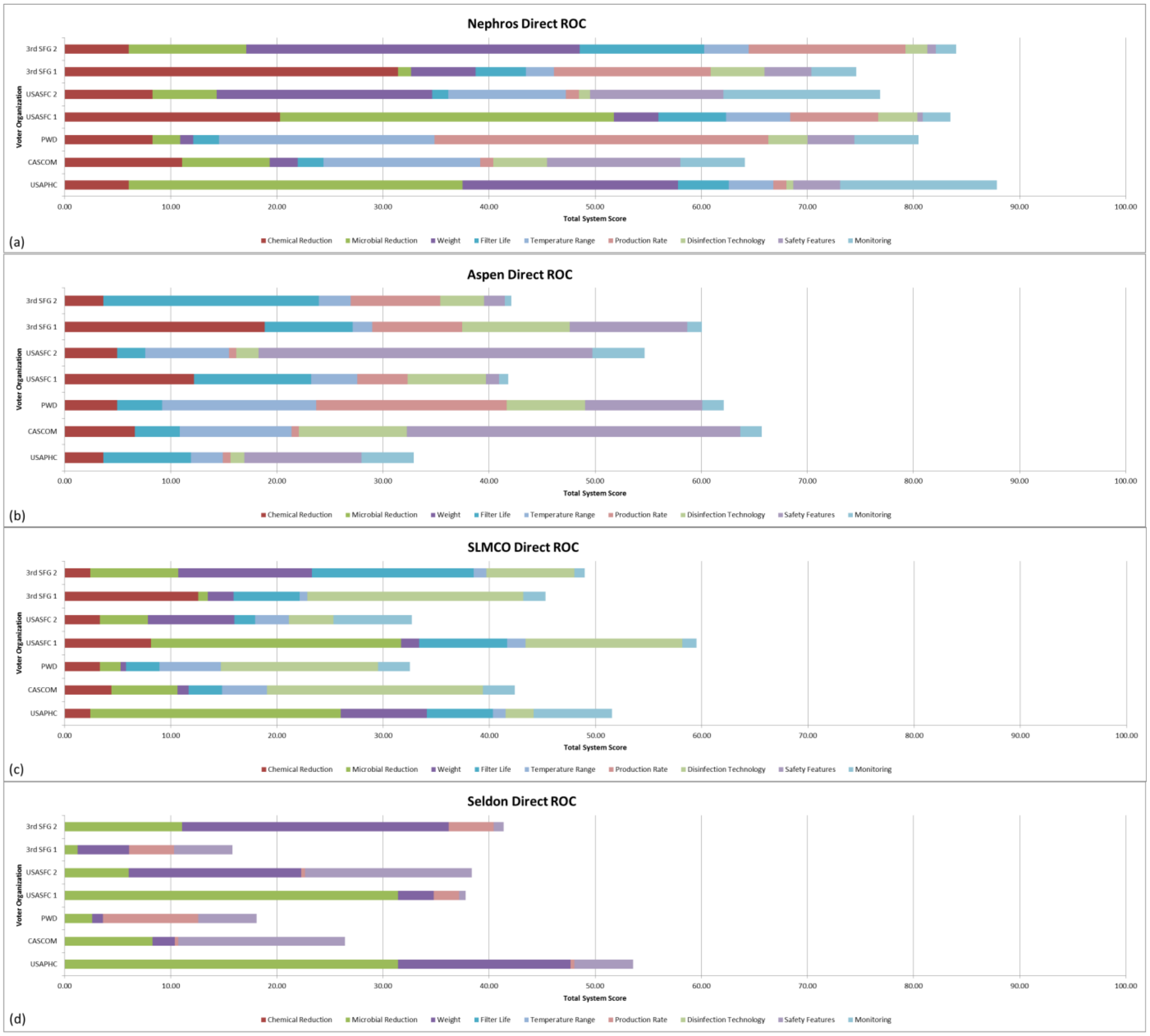


Figure 17. Bar chart illustrating total aggregated system scores by attribute for the direct rating and rank order centroid weighting methods for the following systems: (a) Nephros MSU ultra filter, (b) Aspen 1800 BC, (c) SLMCO FBS 180, and (d) Seldon Waterbox

REFERENCES

Afgan, N. H., & Carvalho, M. G. (2002). Multi-criteria assessment of new and renewable energy power plants. *Energy*, *27*(8), 739–755. doi:10.1016/S0360-5442(02)00019-1

Ananda, J., & Herath, G. (2009). A critical review of multi-criteria decision making methods with special reference to forest management and planning. *Ecological Economics*, *68*(10), 2535–2548. doi:10.1016/j.ecolecon.2009.05.010

Bray, R. M., Rae Olmsted, K. L., Williams, J., Sanchez, R. P., & Hartzell, M. (2006). Progress toward healthy people 2000 objectives among U.S. military personnel. *Preventive Medicine*, *42*(5), 390–6. doi:10.1016/j.ypmed.2006.01.009

CASCOM. (2008). *Water Planning Guide*. Fort Lee, VA. Retrieved from http://www.quartermaster.army.mil/pwd/Publications/Water/Water_Planning_Guide_rev_103008_dtd_Nov_08_(5-09).pdf

Clemen, R. T. (2001). *Making Hard Decisions with Decision Tools* (2nd rev. e.). Pacific Grove, CA: Duxbury/Thomson Learning.

Col. Ty Connett and Col. Bob Cassidy. (2011). Village Stability Operations: More than Village Defense. *Special Warfare*, *24*(3), 23–27. Retrieved from http://www.soc.mil/swcs/swmag/archive/SW2403/SW2403VillageStabilityOperations_MoreThanVillageDefense.html

Department of Defense Design Criteria Standard Human Engineering. (2012) (p. 381). Redstone Arsenal. Retrieved from http://www.everyspec.com

Dyer, J. S. (1990). Remarks on the Analytic Hierarchy Process. *Management Science*, *36*(3), 249–258.

Edwards, W. (1977). How to Use Multiattribute Utility Measurement for Social Decisionmaking. *IEEE Transactions on Systems, Man and Cybernetics*, (5), 326–340.

Edwards, W., & Barron, F. (1994). SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement.

Ewing Jr., P. L., Tarantino, W., & Parnell, G. S. (2006). Use of Decision Analysis in the Army Base Realignment and Closure (BRAC) 2005 Military Value Analysis. *Decision Analysis*, *3*(1), 33–49. doi:10.1287/deca.1060.0062

Fülöp, J. (2001). Introduction to Decision Making Methods, 1–15.

Gomes, L. F. A. M., & Rangel, L. A. D. (2009). Determining the utility functions of criteria used in the evaluation of real estate. *International Journal of Production Economics*, *117*(2), 420–426. doi:10.1016/j.ijpe.2008.12.006

Gomes, M., Alberto, L., Rangel, D., & Leal, R. (2011). Treatment of Uncertainty Through The Interval SMART/SWING Weighting Method: A Case Study, *31*, 467–485.

Hämäläinen, R. P., & Alaja, S. (2008). The threat of weighting biases in environmental decision analysis. *Ecological Economics*, *68*(1-2), 556–569. doi:10.1016/j.ecolecon.2008.05.025

Huang, I. B., Keisler, J., & Linkov, I. (2011). Multi-criteria decision analysis in environmental sciences: ten years of applications and trends. *The Science of the Total Environment*, *409*(19), 3578–94. doi:10.1016/j.scitotenv.2011.06.022

Jahanshahloo, G. R., Lotfi, F. H., & Izadikhah, M. (2006). Extension of the TOPSIS method for decision-making problems with fuzzy data. *Applied Mathematics and Computation*, *181*(2), 1544–1551. doi:10.1016/j.amc.2006.02.057

Kain, J.-H., & Söderberg, H. (2008). Management of complex knowledge in planning for sustainable development: The use of multi-criteria decision aids. *Environmental Impact Assessment Review*, *28*(1), 7–21. doi:10.1016/j.eiar.2007.03.007

Karjalainen, T. P., Rossi, P. M., Ala-aho, P., Eskelinen, R., Reinikainen, K., Kløve, B., … Yang, H. (2013). A decision analysis framework for stakeholder involvement and learning in groundwater management. *Hydrology and Earth System Sciences Discussions*, *10*(7), 8747–8780. doi:10.5194/hessd-10-8747-2013

Keeney, R. L., & Raiffa, H. (1976). *Decisions With Multiple Objectives: Preferences and Value Tradeoffs*. New York: John Wiley& Sons.

Linkov, I., Satterstrom, F. K., Kiker, G., Batchelor, C., Bridges, T., & Ferguson, E. (2006). From comparative risk assessment to multi-criteria decision analysis and adaptive management: recent developments and applications. *Environment International*, *32*(8), 1072–93. doi:10.1016/j.envint.2006.06.013

Lundquist, A. H., White, G. H., Bonilla, A., Richards, T. E., & Richards, S. C. (2011). Adapting military field water supplies to the asymmetric battlefield. *U.S. Army Medical Department Journal*, 62–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21805456

Marttunen, M., & Hämäläinen, R. P. (2008). The decision analysis interview approach in the collaborative management of a large regulated water course. *Environmental Management*, *42*(6), 1026–42. doi:10.1007/s00267-008-9200-9

Seager, T. P. (2004). Understanding Industrial Ecology and the Multiple Dimensions of Sustainability. In O'Brian (Gere Engineers) (Ed.), *Strategic Environmental Management* (pp. 17–70). New York: John Wiley& Sons.

Sharma, M. (2013). Multi Attribute Decision Making Techniques, *1*(1), 49–51.

Site, P. D., & Filippi, F. (2009). Weighting methods in multi-attribute assessment of transport projects, 199–206. doi:10.1007/s12544-009-0018-1

Tavana, M., & Hatami-Marbini, A. (2011). A group AHP-TOPSIS framework for human spaceflight mission planning at NASA. *Expert Systems with Applications*, *38*(11), 13588–13603. doi:10.1016/j.eswa.2011.04.108

Triantaphyllou, E., & Mann, S. H. (1989). An Examination of the Effectiveness of Multi-Dimensional Decision-Making Methods : A Decision-Making Paradox, 303–312.

Triantaphyllou, E., & Mann, S. H. (1995). Using the Analytic Hierarchy Process for Decision Making in Engineering Applications: Some Challenges, *2*(1), 35–44.

U.S. Dept of the Army. (2002). *Field Manual 4-25.12: Unit Field Sanitation Team* (Vol. 12). Washington, DC.

U.S. Dept of the Army. (2010). *Technical Bulletin Med 577, Sanitary Control and Surveillance*. Washington, DC.

US Dept of the Army. (2010). *Field Manual 3-05: Army Special Operations Forces*. Washington, DC. Retrieved from https://armypubs.us.army.mil/doctrine/DR_pubs/dr_c/pdf/fm3_05.pdf

USAPHC. (2010). *Technical Evaluation and Operational Needs Assessment of Commercial-Off-The-Shelf Small Unit Water Purifiers* (p. 288). Aberdeen Proving Ground, MD.

Von Winterfeldt, D. (1986). *Decision Analysis and Behavioral Research*. New York; Cambridge: University Press.

Yoe, C. (2002). Trade-Off Analysis Planning and Procedures Guidebook, (April).

Yoon, K. P., & Hwang, C.-L. (1995). *Multiple Attribute Decision Making: An Introduction*. Thousand Oaks, CA: Sage Publications.