15TH INTERNATIONAL COMMAND AND CONTROL RESEARCH AND TECHNOLOGY SYMPOSIUM

THE EVOLUTION OF C2

USING FORMATIVE EVALUATIONS TO ASSESS THE POTENTIAL VALUE OF GEOSPATIAL DECISION SUPPORT PRODUCTS DURING DEVELOPMENT

Suggested Topics:
Track 5 – Experimentation and Analysis
Track 8 – C2 Assessment Tools and Metrics
Track 3 – Information Sharing and Collaboration Processes and Behaviors

Walter A. Powell [STUDENT] - GMU*
Ryan Johnson [STUDENT] – GMU
Leonard Adelman - GMU
Kathryn Blackmond Laskey - GMU
Michael Altenau – Viecore, FSD
Andrew Goldstein – Viecore, FSD
Kenneth Braswell - ERDC
Vineet Gupta - ERDC

George Mason University (GMU)
Center of Excellence in C4I
The Volgenau School of Information Technology and Engineering
George Mason University
4400 University Drive Fairfax, VA 22030-4444 USA
wpowell@gmu.edu
(703) 993-7589

Viecore, FSD
23 Christopher Way
Suite 301
Eatontown, NJ 07724 USA

U.S. Army Engineering, Research and Development Center (ERDC)
7701 Telegraph Road Alexandria, VA 22315 USA

ABSTRACT

Geospatial Decision Support Products (GDSPs), computer-based tools that help interpret geospatial data, are ubiquitous within current military forces. Although GDSPs are used widely, and are generally perceived as useful, there is too little empirical evidence to assess the value of GDSPs to the warfighter. Achieving a better understanding of the potential value of GDSPs while they are still in development would help to prioritize scarce intellectual, physical, and monetary resources, and would result in having the most valuable GDSPs available to the warfighter. General experimental methods are well understood and have been used recently to evaluate GDSPs during development (e.g., Powell et al., 2009). However, it can take considerable time and effort to recruit the type and number of participants required to conduct experiments with adequate statistical power. This paper describes how formative evaluations emulating experiments in terms of hypotheses, scenarios, procedures, and measures, but using only a small number of subject matter experts (SMEs), can be used effectively to inform development decisions about the potential value of GDSPs. A case study in which a formative evaluation was used to inform development decisions for the Battlespace Terrain Reasoning and Awareness-Battle Command (BTRA-BC) Tactical Spatial Objects (TSOs) is presented to illustrate the method.

1. Background

The fundamental goal of developing Geospatial Decision Support Products (GDSPs) for the military is to aid the military decision-maker, at whatever level, by transforming the vast amounts of geospatial data available into geospatial information products that are relevant to the decision at hand. As researchers and developers strive to provide advanced GDSPs to process more data faster and more accurately, good management necessitates the assessment of each innovation so that key resources can be allocated to areas that yield the most “bang for the buck.” The Joint Geospatial Enterprise Service (JGES) program of the U.S. Engineering, Research and Development Center (ERDC) is designed to meet this need by evaluating the value of GDSPs to military decision-making.

GDSPs provide the context-specific model portion of a Geospatial Decision Support System (GDSS). GDSSs are Decision Support Systems (DSSs) that perform automated analyses of geospatial data, transforming that data into information, and displaying the resulting geospatial information. A GDSS can include one or more Geospatial Decision Support Products (GDSPs) which calculate and display geospatial data or geospatial information.

The specific GDSPs evaluated in this paper are part of the Battlespace Terrain reasoning and Analysis – Battle Command (BTRA-BC) GDSP tool set. BTRA-BC consists of a collection of GDSPs called Tactical Spatial Objects (TSOs) that can be embedded as modules of a GDSS. TSOs are computationally lightweight objects that generate geospatial information specific to the tactical context for which they are developed. The following TSOs were the subject of this formative evaluation:

(1) Assembly Area (AA) – Evaluated potential AA sites which met required minimum size and geospatial criteria. The suitability of the potential sites was evaluated based on geospatial data above the minimum requirements.

(2) Choke Point (CP) – Evaluated areas which would restrict the movement of brigade, battalion, company, and platoon sized units based on calculated mobility corridors and the width of the passable terrain between obstacles. The choke points were as to the size of the unit that could pass through the restriction.

(3) Engagement Area (EA) – Evaluated potential sites for EAs based on the size and suitability of geospatial considerations, such as the proximity of choke points to mobility corridors.

(4) Movement Projection (MP) – Evaluated routes based on the mobility corridors and suitability of underlying geospatial criteria for use by the selected vehicle types. The MP options were:

a) Route Weighting Options;
   i. Shortest – evaluates for the shortest route;
   ii. Fastest – evaluates for the fastest route;
   iii. Primarily on-road – gives preference to on-road segments;
   iv. Primarily off-road – gives preference to off-road segments;

b) Routing Options;
   i. One-to-One – one start point and one endpoint;
   ii. One-to-Many – one start point and multiple endpoints;
   iii. Many-to-One – multiple start points and one endpoint;
   iv. Many-to-Many – multiple start points and multiple endpoints;

c) Time-Constrained routing – evaluates route/distance combinations that can be achieved in a given time;

d) Time- and Objective-Constrained routing – evaluates route/distance combinations that can be achieved in a given time toward a specified objective

e) Force-On-Force routing – evaluates potential areas where opposing forces may meet given specified starting points and objectives ;
f) Named Area of Interest routing – evaluates times of arrival and departure from specified areas given a starting point and objective;
g) Capacity Flow routing – evaluates routes based on the aggregate size of the units moving toward the objective(s);
h) Suitability-Seeking routing – evaluates routes toward areas with user-specified geospatial characteristics;
i) Multi-Constrained routing – evaluates routes based on a user-specified combination of geospatial criteria;

The general context for which our TSOs were developed is that of military ground operations. In that context, the potentially superior situation awareness afforded by BTRA-BC TSOs opens up new possibilities for the planning of military operations. Translating these possibilities into practical decision support requires a build-test-build cycle that channels technology in spiral development to ensure results that best support the warfighter. This paper reports on the fourth in a series of experiments and evaluations designed to assess the value of GDSPs to the decision-maker, and to provide results that inform the spiral development cycle. In the case of military GDSSs and GDSPs, the ultimate decision-maker is the military commander, and the ultimate goal is to support command decisions in the most effective way.

The current paper describes the first in a series of formative evaluations designed to assess the benefits of BTRA-BC TSOs early in their development. Previous summative experiments (Laskey et al., 2007; Powell et al., 2009, 2008) on BTRA-BC TSOs concentrated on achieving statistically significant results with a small number of military subject matter experts as participants. These experiments provided to the program sponsor both quantitative and qualitative evidence of the benefits of the TSOs to the warfighter and the areas that needed further attention. These experiments were conducted near the end of development of the TSOs under evaluation, and although they achieved statistically significant results with approximately 16 participants, finding qualified military participants was difficult in a period of high operational tempo. The current formative evaluation is based on the same hypothesis-driven design as the summative experiments, but used substantially fewer participants. Although it could not yield statistically significant quantitative results, it did yield significant findings about the potential value of TSOs early enough in their development to cost-effectively inform development decisions.

This paper is organized as follows. Section 2 discusses summative experiments and formative evaluations. Section 3 lays out the evaluation design and the reasoning that led to this design. Section 4 provides the average participant responses and a summary of qualitative results of the evaluation and Section 5 discusses our conclusions.

2. Summative Experiments and Formative Evaluations

The goal of the BTRA-BC project is to develop geospatial tools that empower commanders, soldiers, and systems with information that allows them to understand and incorporate the impacts of terrain and weather on their functional responsibilities. Providing useful tools to the warfighter means more than simply meeting product design requirements. In addition, it means evaluating the tool and assessing its performance relative to the purpose for which it was designed. Both summative experiments and formative evaluations can provide this kind of feedback on the ability of geospatial tools to provide useful geospatial information to military decision-makers. The difference between these two methods of evaluation is one of intent and, by extension, the type of information that needs to be gathered.

Summative experiments are conducted near or after project completion to provide information to some external decision-maker usually for strategic decisions such as continued funding (Scriven, 1991). The evaluation of a project near its completion should capture quantitative as well as qualitative data on the project to aid in strategic decisions about the project. Our first two summative experiments provided this kind of information on completed BTRA-BC TSOs. The single drawback to conducting summative experiments was the lack of availability, because of the current operational tempo, of participants with the necessary training and experience to represent future system users.

Because of this drawback and the need to provide feedback to the developers, the project team decided that quantitative results were not as important as qualitative results. Consequently, the program focus shifted to formative evaluations. Formative evaluations can be conducted at almost any point during development and are designed to provide information to decision-makers internal to the project for use in improving the product (Scriven, 1991). Without the necessity to provide statistically significant quantitative feedback, the number of participants in a formative evaluation could be reduced from the minimum of 15 participants required in a summative experiment and still provide good qualitative feedback. The formative evaluation described in this paper has the same hypothesis driven rigor as our summative experiments. Although fewer participants (4-6) were required, the goal was still to provide valuable information to the TSO developers.

3. Evaluation Design

The goal of the first formative evaluation was to determine which aspects of the various TSO engines were valuable to the decision-maker and which were not. The formative evaluation was designed to provide both positive and negative feedback. High Definition Documents (HDDs) drafted by Subject Matter Experts (SMEs) for each TSO engine provided doctrinal background on the tasks each TSO was designed to support and specified the inputs, outputs, and functions of each TSO engine. These HDDs provided the basis for evaluating each TSO. The evaluation also stressed the TSOs and evaluated the available inputs and generated outputs produced relative to supporting appropriate missions (tasks).

Hypotheses

Like the previous summative evaluations, the problem of evaluating the value of the TSOs requires defining what is valuable to the decision-maker. The evaluation team initially identified dozens of criteria of value for the various TSOs under evaluation. These criteria were vetted by SMEs to confirm their validity. Many of the criteria could be grouped to form concise hypotheses. These hypotheses are:

A TSO would be considered valuable if

1. *Using the TSO to complete the task would improve the quality of the participant's solution to the task.* Rationale: A tool that allows the participants incorporate to geospatial information that is not available with currently fielded systems will improve their plans.
2. *Using the TSO to make completing the task easier.* Rationale: A tool that presents additional information that is relevant to the task and is easily assimilated would make complex planning less difficult.
3. *Using the TSO to complete the task saved the participant time.* Rationale: The TSOs' analytical capabilities will permit planners to make decisions more quickly in a time-constrained planning environment.

4. *The functions and setting of the inputs are easy to understand*. Rationale: TSO functions and inputs that are not easily understood will preclude their use and deny the decision-maker information relevant to the task.
5. *Adjusting the inputs provides additional information about the effect of terrain on the task.* Rationale: Varying the user-defined inputs should generate outputs that display different information about the terrain which in turn should aid in the analysis of the impact of terrain on the task.
6. *The TSO highlights the terrain information most important to completing the task.* Rationale: The TSO information presented to the decision-maker should emphasize the information most relevant to the task. Information deemed to be less relevant should still be available for option analysis.
7. *The participants would elect to use the TSO to complete the task instead of, or in addition to, currently available methods/tools.* Rationale: Electing to use the TSO indicates that the information it generates is more valuable to the decision-maker than any negative aspects of the TSO.

These hypotheses were used to generate questionnaires designed to gather information on the value of the TSOs to participants. Not all of the criteria generated by the evaluation team fit into these hypotheses. Many were specific to the individual TSOs and did not generalize to the other TSOs. These criteria were included as part of the group discussions (hot-wash). The results from the questionnaire and hot-wash discussions are presented later in this paper.

Graphical User Interface (GUI) Concerns

Because BTRA-BC TSO engines are designed to be incorporated into any geospatial system that uses the Commercial Joint Mapping Toolkit (C/JMTK), the HDDs do not delineate the GUIs for specifying user-selectable inputs or the actual graphics associated with the outputs of the TSO engines. These GUIs will vary depending on the system in which the TSOs are used.

In order to evaluate the TSO engines, generic GUIs were built to support the selection of user inputs and the graphical display of the output of the TSO engines. Because the participants used these GUIs during their evaluation of each TSO, it is impossible to totally divorce their evaluation of the value of the TSO engines from their evaluation of the GUI. Efforts were made by the evaluation team to make the participants aware of the difference between the functioning of the TSO engine and the GUIs used. Participant comments concerning the GUIs were noted separately from comments on the TSO engines.

Basis for comparison

In our previous summative experiments with other TSOs, we were able to have the participants complete identical tasks on near identical terrain data using the TSOs and without using the TSOs. The tasks completed without the use of the TSOs provided for a direct comparison between the two cases. In this formative evaluation the participants conducted each task once, using the TSO. The participants had sufficient prior experience to provide a basis for comparing how completing the tasks using the TSOs compared to having to complete the tasks without the TSOs.

Computing Environment

The evaluation was conducted using the BTRA-BC C2 Test and Demonstration Application (TDA) as the geospatial environment which provided the geospatial data handling and user interfaces. TDA is an ESRI ArcEngine Application which also leveraged software from the Commander's Support Environment (CSE). TDA provides the task organization, map background, and graphic control measures necessary to create a scenario (mission). Specific scenarios were created to provide a context for evaluating TSOs. CSE was originally developed for the Defense Advanced Research Projects Agency (DARPA)/Army Multi-Cell and Dismount C2 Program (M&D C2) which was continued from the Future Combat System Command & Control (FCS C2) program (Viecore FSD, 2010). TDA was developed and maintained by Viecore, FSD and Northrop Grumman.

Participants

The participants were Army officers who have been previously trained in military planning and have tactical and operational experience at least up to the battalion (Bn) level. All the participants were familiar with the tasks that the TSOs were designed to support. One participant was active duty and three were retired US Army officers ranging from Captain to Lieutenant Colonel. Three of the participants' combat specialties were infantry, and one was an engineer. All four participants were comfortable using computers and had used computers to aid in decision-making previously.

Training

Because none of the participants were familiar with TDA or TSOs, the participants needed to be trained to the point that they were comfortable using the TDA interface and TSOs. The TDA training consisted of a hands-on tutorial which concentrated on how to manipulate the displays, run the TSOs, select the various TSO outputs for display, retrieve information relative to the TSO output, and save outputs. The training required to navigate the system was minimal because TDA is a Windows-based geospatial system and its controls are very similar to systems already in use in Army units. After training, all the participants asserted that they were comfortable manipulating the TDA environment.

The participants were briefed on the overall scenario prior to receiving training on individual BTRA-BC TSOs. The training on BTRA-BC TSOs consisted of three parts: (1) a briefing on the purpose of each TSO concentrating on what data the TSO will use, the doctrinal requirements that the TSO is designed to meet, how the TSO generates its output graphics and information, and how to interpret the output of the TSO; (2) a briefing on the specific task used to evaluate each TSO, terrain considerations, and enemy disposition; and (3) practice with the TSO consisting of a hands-on tutorial walking through the use of the TSO and its options on terrain similar to that to be used for the evaluation. The hands-on training was conducted in different terrain so that the evaluation did not turn into a demonstration. In demonstrations, participants can become just observers. Having the participants use the TSOs on different terrain from that used in training kept them mentally involved in the evaluation process. The participants were encouraged to explore the limits of the TSO and all the options available throughout the evaluation.

Tasks

The tasks the participants were asked to complete were carefully planned to support the specific TSOs under evaluation. The participants were provided a general scenario consisting of a mission, commander's intent, mission graphics, intelligence on enemy disposition and intent, and friendly order of battle for a movement to assault for one battalion of a Stryker Brigade Combat Team (SBCT). This scenario provided the overall framework within which the individual TSO evaluation tasks were run and gave the participants a proper background for completing the planning tasks for each TSO. This scenario provided a context and simulated as well as possible, in the informational realm at least, the tactical conditions under which each of the tasks would be performed in the field.

Factors considered in creating the tasks included the intended user for each TSO, the TSO inputs, the information the TSO was designed to generate, and the type of terrain in which the TSO would be most useful. The first factor required identifying for whom the TSO was intended (e.g. planner or operator), the level at which the TSO was targeted (e.g. company, battalion, brigade), and the specialty of the user. Understanding the perspective of the intended user was crucial to generating tasks that would provide useful information about the value of the TSO. The appropriate user will likely understand the task, understand the impact of the terrain on the task, and be able to evaluate the potential usefulness of the TSO in completing the task. Generating a task for someone other than the intended user will still generate feedback on the usefulness of the TSO to that user, but the feedback may not address issues that are most useful to the developers. In this evaluation the intended users were both planners and operators at the company and battalion level with expertise in infantry operations.

The last three factors – TSO inputs, output information, and terrain data – are closely tied together. The tasks must be designed such that the task would logically require using all the available inputs and that the information generated by the TSO would be applicable to the task. The background provided by the HDDs generally addressed the information the TSO would generate and the general tasks it was designed to support, but the actual information generated depended on the inputs selected and the terrain data. Fully evaluating each TSO required that the participants exercise all the inputs available. Using all the inputs would only be necessary if varying each input in turn generated additional information that aided the user in understanding the impact of the terrain on the task. In actual use, given the terrain, varying some inputs may not generate additional information. However, for evaluation purposes, the terrain must demonstrate the affect of varying all the inputs on the output.

Careful selection of the terrain data was necessary to ensure continuity during the evaluation. Since we desired that the participants feel that the evaluation simulated real world missions, one general mission with several TSO specific tasks would minimize the participants' need to shift focus and allow them to become familiar with the terrain. Ideally, in support of one general mission, a single terrain data set was selected that would require the participants to exercise the various inputs. For this evaluation, one of the TSOs, Assembly Areas (AAs), operated differently on open terrain (few obstacles) than on closed terrain (obstacles). All the TSOs used the same open terrain data set (National Training Center), but for the AA TSO an additional closed terrain data set (Korea) was used. The overall scenario and tasks were identical for both of the AA TSO tasks. Only the data set was changed to support this specific TSO.

The actual tasks the participants were asked to perform consisted of vignettes and instructions that defined the specific task and subtasks specific to the TSO. Each task also specified a frame of reference from which the task was to be evaluated. This was done by designating the staff position role the evaluator was to take (e.g. battalion operations officer (S3) or company commander). This role was chosen to be representative of the user for which the TSO was intended, and was important in getting the appropriate feedback from the participants. The tasks also provided detailed instructions indicating how to access the terrain data and operational graphics for the task as well as the recommended initial settings for user-defined inputs. The evaluation team emphasized that the recommended initial inputs were only a starting point and encouraged the participants to vary the inputs to stress the TSO and to determine how varying the inputs affected the output of the TSO. The general the tasks for this evaluation consisted of:

(1) identifying company-sized Assembly Areas (AAs) in both
  a) open terrain (desert),
  b) closed terrain (wooded hills),
(2) identifying Choke Points (CPs),
(3) identifying potential Engagement Areas (EAs),
(4) generating routes using:
  a) four route weighting options (shortest, fastest, on-road, off-road),
  b) four routing options (one-to-one, one-to-many, many-to-one, many-to-many),
  c) time-constrained routing,
  d) time- and objective-constrained routing,
  e) force-on-force routing,
  f) Named Area of Interest (NAI) routing,
  g) capacity flow routing,
  h) suitability-seeking routing, and
  i) multi-constrained routing.

An example of the actual task given to the participants for evaluating Named Area of Interest (NAI) routing was:

Open mission *I:\JGES_eval1\mission\NAI.mdx*
You are the S2 [intelligence staff] for 1/3 SBCT [1st battalion, 3rd Stryker Brigade combat Team]. Given the Red mechanized infantry company will transit the

provided Named Areas of Interest (NAIs), evaluate the routes and times of Red Force of arrival in and departure from NAIs 1, 2 & 4 in your Area Of Operations (AOO). ***Generate and Evaluate*** *an NAI analysis using the following setting:*

Red Force movement: Objective Slam to Phase Line (PL) PHILLIES
Cost Attribute: Med Mobility Wheel
Accumulate: Med Mobility Wheel, Shortest
Default Break: 80 minutes
Restriction: Obstacles, roads
Boundaries: PL PHILLIES, PL HOMERUN, in the 1/3 SBCT AOO

Save your mission

Below is typical output resulting from NAI TSO using the recommended initial inputs above:

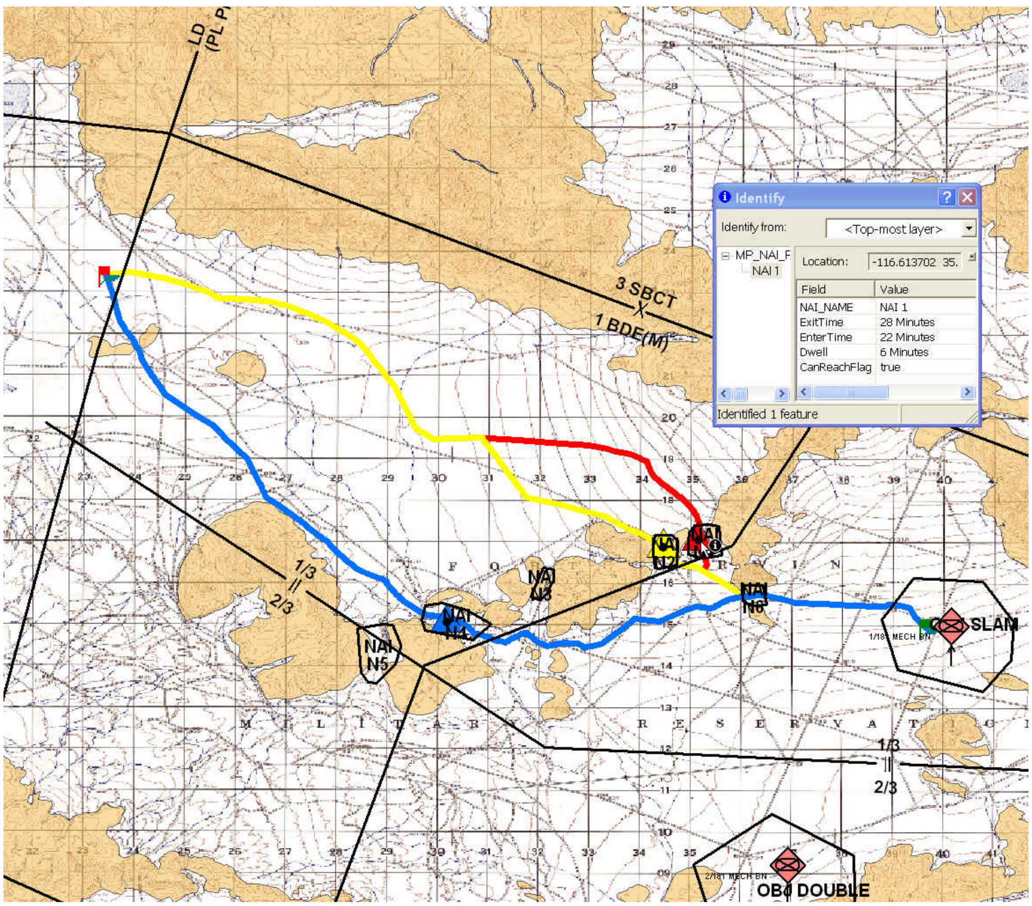

Figure 1: Typical NAI TSO Output

The TSO was designed to assist planners in estimating when reconnaissance assets would need to be in position to observe the enemy units pass through the NAIs. The TSO predicts potential enemy routes (in red, blue and yellow) from their position in Objective SLAM through the NAIs (outlined in black). The inset "identify" box shows the time the enemy unit will enter and leave NAI I and the amount of time it will be within the NAI.

Measures

All of the measures used in the evaluation were subjective; participants provided subjective assessments of the value of the TSO engines in supporting each task. Of the seven measures derived directly from the seven hypotheses, four of the measures directly concern the value of the TSOs. Three of the measures of the value of the TSOs were based on measures used in our previous summative experiments. The participants were all trained and experienced military officers and they were asked to compare using the TSO to complete the tasks to how they would have completed the tasks without using the TSO. The measures of value were:

(1) *Output quality*: Would completing the task using the TSO help generate a higher quality output?
(2) *Ease of task completion*: Would using the TSO make completing the task easier?
(3) *Time saved*: Would using the TSO save time in completing the task?

A fourth measure, a global measure, was used to summarize the positive aspects of the TSO relative to any negative aspects.

(4) *Usefulness*: "Would the participant use the TSO for the task specified if the TSO were available?"

A positive response to the fourth measure would indicate that the information generated by the TSO was more valuable to completing the task than any negative factors that detracted from completing the task with the TSO. A negative response would indicate that any negatives associated with the TSO would outweigh any value of information generated by the TSO. While the individual measures would provide feedback to the developers on specific functions of the TSO, this latter measure would summarize its overall value.

The other three measures concern the mechanics of the TSOs and would provide direct feedback to developers. They were:

(5) *Input settings and functions*: Provides feedback on the usefulness of user-selectable inputs.
(6) *Adjusting inputs provided additional information*: Provides feedback on the internal processes of the TSO.
(7) *Outputs*: Provides feedback on the presented TSO-generated information.

Participant responses to all the measures were elicited after they had completed each evaluation. Two methods of elicitation were used: (1) a questionnaire and (2) group hot-wash discussions.

The questionnaire served three purposes; (1) to give the participants a framework in which to think about the TSOs; (2) to attach numeric values to the participants' subjective evaluation of the TSO; and (3) to provide a starting point for the group hot-wash discussions. The participants were briefed on the questionnaire and knowing the measures on the questionnaire gave the participants criteria to consider as they were evaluating each TSO. In order to quantify the participants' responses, metrics and associated numeric scores were developed for each measure on the questionnaire. The metrics defined positive, negative, and neutral levels of response associated with that measure. The metrics allowed the participants to score the measures positively, neutrally, or negatively and corresponded to a 5-point Likert scale. The positive responses were assigned a score of 4 or 5, neutral responses were given a 3, and negative responses were given scores of 1 or 2. The measures follow directly from the seven hypotheses, but were tailored slightly to be specific to each TSO. Participants were encouraged to elaborate on their responses in the comment section associated with each question. Below is an example of a question supporting a measure and its associated metrics:

1. In comparison to traditional mission planning without the BTRA-BC TSO, would using the TSO improve the quality of your mission planning?	**The TSO would…………**				
	significantly reduce the quality of your mission plan	**somewhat reduce** the quality of your mission plan	make **no difference** in the quality of your mission plan	**somewhat improves** the quality of your mission plan	**significantly improves** the quality of your mission plan
Comments					

Figure 2: Example of questionnaire measure and metrics

Post-evaluation group hot-wash discussions provided a means for the evaluation team to elicit information which amplified the participants' responses, to obtain comments on the questionnaire, and to discuss measures which did not correspond to the seven hypotheses. The evaluation team led the discussion, which was recorded and later transcribed. The participants' responses during the group hot-wash discussions provided qualitative information on the value of the TSO, positive and negative feedback to the developers, and suggestions to improve both the TSOs and their associated GUIs. A final group hot-wash discussion was conducted to provide feedback to the evaluation team on the conduct of the evaluation including scenarios, tasks, questionnaires, and hot-wash discussions.

4. Results

The formative evaluation design and procedures described above provided useful feedback to the sponsoring organization. The results were presented to both management and development personnel. The level of agreement of the participants on each point was presented along with their qualitative comments. Information was presented on the value of each BTRA-BC TSO as well as general comments applicable to all TSOs.

The averages of the participants' scores from the subjective questionnaires are summarized in Table 1 below. Because there were only three or four participants in each trial, we are unable to provide standard deviations or confidence intervals for the data. Instead we coded the results into three categories to indicate the agreement of the responses: (1) all positive responses were coded with bold text and a green background, (2) positive responses with one or more neutral response were coded with non-bold text and a white background, and (3) indeterminate responses (contained both positive and negative responses) were coded with double-outlining and a yellow background. Asterisks are presented for the Engagement Area TSO because one of the participants was so negative about the TSO that he refused to complete the questionnaire. Since he convinced the other two participants of his concerns during the group hot-wash discussions, we thought it was inappropriate to present their numeric scores for the TSO in Table 1.

The scores and comments from the participants' questionnaires provided a basis for discussions in the hot-wash discussions. Scores other than those in the all positive response category indicated areas which would provide potential feedback to the developers. Guided by the participants' responses, the evaluation team encouraged the participants to expound on their own written comments as well as the responses of the other participants. When appropriate, the evaluation team introduced for discussion some of the previously developed criteria that were not included in the questionnaires. Even though the numerical data provided a snapshot of the participants' responses, their comments provided the most potentially useful feedback. Below is a summary of the major points generated during the post-evaluation hot-wash discussions:

1. Most TSOs were valuable in their current form and participants would use them if they were available.

	Assembly Areas		Choke points	Engagement Areas	Movement Projection														
					Weighted Costs				Route Options										
	Closed Terrain	Open Terrain			Shortest	Fastest	Off-Road	On-Road	One-to-One	One-to-Many	Many-to-One	Many-to-Many	Time-Constrained	Time- and Objective-Constrained	Force-on-Force	NAI	Capacity Flow	Suitability Seeking	Multi-Constrained
1. In comparison to not using the BTRA-BC TSO, would using the TSO **improve the quality** of your output?	4.20	4.00	3.75	**	4.25	4.50	4.25	4.75	4.25	4.33	4.33	4.33	5.00	4.33	3.67	4.00	5.00	4.00	4.00
2. In comparison to not using the BTRA-BC TSO, would using the TSO using the TSO make the task **easier**?	4.25	3.75	4.00	**	4.25	4.50	4.25	4.50	4.25	4.33	4.67	4.67	5.00	4.33	3.33	4.00	5.00	4.00	4.50
3. In comparison to not using the BTRA-BC TSO,, would using the TSO allow you to complete the task **more quickly**?	4.25	3.75	4.00	**	4.00	4.50	4.25	4.25	4.50	4.67	5.00	5.00	5.00	4.67	3.33	4.50	5.00	4.00	4.50
4. With respect to the task, are the functions and settings of available inputs **easily understandable**?	3.25	3.25	N/A	N/A	3.25	3.25	3.25	3.25	3.50	3.33	3.33	3.33	4.00	4.00	3.67	3.50	4.00	3.00	3.50
5. With respect to the task, does **adjusting the inputs** provide additional information about the terrain?	3.00	2.67	N/A	N/A	3.75	4.00	3.75	4.00	3.75	3.50	3.50	3.50	4.33	4.33	3.00	3.00	3.00	3.00	3.00
6. With respect to the task, do the graphics highlight the information you feel is the **most important**?	4.75	5.00	4.50	**	4.67	4.67	4.67	4.67	4.67	5.00	5.00	5.00	5.00	4.67	3.50	4.50	5.00	5.00	4.50
7. Overall , with respect to the task, would you **use** this TSO if it were available?	4.75	4.00	4.50	**	4.50	4.75	4.50	4.75	4.75	4.67	5.00	5.00	5.00	4.67	2.00	4.50	5.00	4.50	4.67

Table 1: Numerical Results (averages)

2. Their initial confidence in the TSOs would be enhanced if:
    a) It were easier to see the data on which the TSO outputs were based and
    b) The participants better understood the process for generating output graphics and tables.
3. Participants were enthusiastic about most of the Movement Projection TSO options, but found little utility for the Force-on-Force option.
4. Current version of the Engagement Area (EA) TSO was not ready to be fielded:
    a) It did not provide any information beyond than that available from the Choke Point TSO;
    b) Additional information was needed, such as size of units; weapons, observation, fire support, & battle positions; and the commander's desire about where to engage the enemy;
    c) The EA TSO should be developed at a more mission-specific level.
5. Addressing GUI issues would allow participants to better understand the potential value of the TSO engines.
6. It was difficult to access information other than that presented as graphics.
7. There was concern that in order for these TSOs to be used in the field, significant preparation by topographic experts would be needed to prepare the data.

In addition to the summary of results above, results specific to each TSO were obtained. An example of results specific to a TSO were those from the NAI TSO of Figure 1: (1) all the participants asserted that they would use it if it were available; and (2) the participants evaluated this TSO as potentially useful for both offensive and defensive missions including generating multiple friendly Courses of Action (COAs).

5. **Conclusions**

This evaluation demonstrated that informative feedback can be obtained from a well-constructed formative evaluation using a small number of participants. Ideally, five to ten participants would be preferred in order to ensure the stability of the finding. Moreover, to generate the most useful feedback possible, it is critical to find participants who have the proper training and experience to represent the intended users of the TSOs. During the current period of high operational tempo, the single most difficult task facing the evaluation team continues to be the availability of participants with the necessary expertise.

Developing a formative evaluation that will make the best use of qualified participants requires close cooperation among developers, SMEs, operational units, and the evaluation team. A well-designed formative evaluation of GDSPs early in development can test the viability of the underlying concepts prior to investing intellectual and monetary resources. This was exemplified by the evaluation of the Engagement Area (EA) TSO, which indicated that the TSO was not ready for fielding. As a result, the TSO was designated for further development. In contrast, the Choke Point and Movement Projection TSOs are scheduled for inclusion in the Common Ground Operational User Assessment scheduled for August 2010.

The development of BTRA-BC TSOs is an ongoing project. Since the completion of the formative evaluation described in this paper, the evaluation team conducted a second formative evaluation of five additional BTRA-TSOs for ERDC in December 2009. A third formative evaluation is scheduled for May of 2010. Future formative evaluations will be conducted at various points in the development cycle.

REFERENCES

Laskey, K. B., Powell, W. A., Adelman, L., Hieb, M., Kleiner, M., & Powers, M. (2007). Evaluation of Automated Geospatial Tools. In *12th International Command and Control Research and Technology Symposium*. Newport, RI.

Powell, W. A., Laskey, K. B., Adelman, L., Dorgan, S., Johnson, R., Klementowski, C., et al. (2009). Results of an Experimental Exploration of Geospatial Decision Support Products: Agility in complex planning. In *14th International Command and Control Research and Technology Symposium*. Washington, DC.

Powell, W. A., Laskey, K. B., Adelman, L., Dorgan, S., Johnson, R., Klementowski, C., et al. (2008). Evaluation of Advanced Geospatial tools: Agility in complex planning. In *13th International Command and Control Research and Technology Symposium*. Seattle, WA.

Scriven, M. (1991). *Evaluation Thesaurus* (4th ed.). Newbury Park: Sage Publications.

Viecore FSD. (2010). Products - Science and Technology - BCSE : Nuance - Viecore FSD. Retrieved January 20, 2010, from http://www.nuance.com/viecorefsd/products/science-technology/bcse.asp.