Casewise Visual Evaluation for High-Performance Collaborative Visioning of PGDP Nuclear Enrichment Plant End State

Keiron Bailey[1] Ph.D, Ted Grossardt[2] Ph.D, John Ripy[3] B.S. and Ben Blandford[4] M.A.

Abstract

This article describes the adaptation of the authors' Structured Public Involvement, or SPI, framework for community involvement in the end-state visioning process for the PGDP facility. The SPI framework is designed to elicit community values and provide access to these as a decision support system for the development of feasible, legitimate, and durable end-state management plans. This SPI protocol is designed around community evaluation of visualizations. Key properties of visual evaluation methods for large group visualization are discussed and the Casewise Visual Evaluation method is outlined. CAVE uses a fuzzy logic based neural network modeling approach to build a knowledge base for community preferences across all feasible end-state scenarios. The potential PDGP end-state land-use properties developed from focus group work are integrated into a sample range of dynamic visualizations and the sampling protocol is described. Preliminary results will be presented at the conference.

Introduction

The issues confronted by the Department of Energy in developing a stakeholder-driven plan for the decommissioned Paducah Gaseous Diffusion Nuclear Enrichment plant are detailed by Ormsbee and Hoover (2010). Many of these problems are representative of public processes dealing with environmental management and energy and public goods infrastructure under conditions of uncertainty and risk, with high potential costs as well as benefits (Department of Energy 2008). A further complication is that public processes of this type are often conducted under conditions of poor historical trust between stakeholders and project sponsor (Thomas 1998). The project therefore is of great interest to management organizations and public officials, as well as stakeholder groups and citizens at large.

The PGDP end-state visioning process represents an extension of the authors' Structured Public Involvement (SPI) protocol into the domain of environmental management and facility rehabilitation. SPI relies on John Rawls' Theory of Justice, using procedural justice and access to justice as principles around which the public involvement framework is designed. These principles of SPI are documented elsewhere (Bailey and Grossardt 2010). The intent of applying the Structured Public Involvement, or SPI, process to this challenging issue is to improve the quality of the decision making process by more fairly, and more accurately, eliciting and incorporating stakeholder valuations into the PGDP end-state management decisions. Decision process quality is defined as a function of multiple criteria including; the inclusion of both a

---

[1] Associate Professor, School of Geography and Development, University of Arizona. Harvill 409, 1103 East 2nd Street, Tucson, AZ 85641. Email: kbailey@email.arizona.edu

[2] Director, Academy for Community Transportation Innovation, University of Kentucky. Email: tgrossardt@uky.edu

[3] GIS Manager, Kentucky Transportation Center, University of Kentucky. Email: jripy@uky.edu

[4] Department of Geography, University of Kentucky. Email: bblandford@uky.edu

large number and a wide range of stakeholders; the quality of the data obtained; the efficiency of the overall protocol in time and money expended; and, ultimately, real-time anonymous stakeholder performance evaluation of the process. SPI protocols have been applied to numerous other public infrastructure processes over the previous ten years with notable success in terms of these criteria. In particular, high process efficiency and high process quality values for large-scale open stakeholder evaluations have been documented (e.g. Bailey and Grossardt 2001, Bailey, Grossardt and Pride-Wells 2007, Jewell et al. 2009).

To achieve these performance aims with respect to end-state visioning for the PGDP, the first step was to embed the SPI process within the larger framework for stakeholder value elicitation (Anyaegbanum 2010). The SPI framework was then adapted to incorporate key informational elements from the initial round of focus group meetings, to incorporate these valuations into land use and site properties. The team then converted these into land use plans and landscape scenarios that could be visualized and evaluated at large public forums using 3D visualization software (Grossardt et al. 2010).

Within the customized SPI framework detailed by Grossardt et al. (2010), the Casewise Visual Evaluation methodology was adapted for this new application. This article outlines the properties of CAVE in comparison with other visual assessment methods and details the adaptation. The methodology for the dynamic visual evaluation phase of the SPI protocol is outlined and the build and interpretation of the CAVE model are discussed. Key concepts including the design vocabulary are explained. Similarities and differences between the PGDP end-state visioning process and previous SPI applications using dynamic visualization, such as integrated transportation and land-use planning, are discussed. The open public meetings are scheduled for Spring 2010. Therefore, the stakeholder data required to populate the model has not yet been acquired. By the time the conference convenes, the authors expect to provide an overview of this data.

Dynamic Visual Evaluation Methods for PGDP End-State Visioning Process

The team desired to build a database of community preferences for alternative end states as a decision support tool for the project sponsors. The process of evaluation was formalized by working through the logic detailed here. Dynamic visual evaluation is defined as real-time evaluation of visualizations of end-states containing an interactive element, both in the presentation media and in the value elicitation framework, allowing the team to elicit, document and evaluate the interpretations provided by stakeholders. The purpose of the dynamic visual evaluation phase is to help large groups of stakeholders to evaluate visualizations of feasible PGDP end-states in real time and to explore the qualities perceived in these end-states by a meaningful cross-section of the attendees. This data is used to build the model of community preferences.

Although the word “visualization” is most often associated with 3D computer-generated renderings, or Virtual Reality environments, in its broad sense it means a visual representation of an environment. This could be computer animation, still image, diorama, charette, or virtualization i.e. interactive 3D environment.

Whichever representation mode is selected, visual assessment of these images or representations is a complex problem domain (Steinitz 1990). There are two primary philosophies of visual assessment; scenario evaluation, in which one complete visual representation is compared to another; and elemental decomposition, in which a visual representation can be disaggregated into components, each of which is assumed to influence viewer response both individually, and synergistically.

There are several methodological issues associated with visual scenario evaluation by large groups. Table 1 shows properties of visual evaluation methods.

Table 1. Properties of visual evaluation methods

Visual assessment method	Philosophy	Advantages	Disadvantages
Traditional visual assessment	Composite	Intuitive, rapid evaluation	Unstructured, very limited analytic capacity
Visual Preference Survey (VPS®)	Elemental	Rapid scoring	Marginal discrimination unreliable
		Intuitive	No analytic method
			Not open for public inspection
Exhaustive pairwise comparison	Composite	Explicit elemental scoring	Too data-hungry
		Reliable marginal preference discrimination	Far from intuitive
			Potential for inconsistency (i.e. preference intransitivity)

There are conflicting goals for visual evaluation. The need for a high volume and quality of input data from a large number of participants must be balanced against the cost and time involved in acquiring this data. The need for a large number of samples must be balanced against logistical and feasibility considerations for each meeting. The desirability of interval or ratio numerical quality inputs for statistical and numerical analysis must be balanced against the seamless functioning of human perception and cognition systems. These factors must all be taken into account during process design and selection of visual method.

Many of the problems encountered with large-scale group visual evaluation are associated with the reality of hosting large public meetings. A key constraint in real public processes is the useful time available. The authors have hosted over one hundred public meetings dealing with infrastructure issues and 90 minutes is an upper bound for this evaluation. Less than 60 minutes is preferable. Experience shows that between twelve and twenty-five visualizations can be evaluated effectively during this timeframe, depending on whether these are still images, or animations requiring run times prior to evaluation. The extent of post-scoring focused verbal evaluation of specific visualizations and their properties must also be considered in the time budget.

Traditional visual assessment consists of showing a small number of images to respondents and eliciting unstructured verbal feedback, or ordinal rankings of one scenario versus another. This method is cheap and easy to implement. It is often employed by consultants and designers for large group evaluations of design proposals. However, despite its convenience, it is a data-poor way to evaluate preference. It leaves unanswered the questions of whether, and how, specific design elements are influencing public valuations, and in which combinations, and it does not address the problem of preference intransitivity.

The Visual Preference Survey, or VPS®, (Nelessen 1994) is widely used by architects, designers and public involvement practitioners for visual evaluation of structures and built environments. It consists of rapid evaluation of images on an integer Likert scale and is quick and intuitive. However, the interpretation of the data, and the way in which elements interact with one another, is left to the minds of the survey designers. No database is generated and third-party analytic inspection of community values in relation to the design elements is not possible. The success of this system depends strongly on the participating group's trust of the individuals administering and interpreting the survey, and of the designers' understandings of how people react to composite scenarios.

Marginal discrimination is most effectively maintained by performing exhaustive pairwise image comparison (e.g. Zube et al. 1982, Whitmore, Cook and Steiner 1995). The function that describes the necessary number of comparisons is given by the *Combination Equation*:

$${}_nC_r = \frac{n!}{r!(n-r)!}$$

Despite the quality of the analysis, this evaluation method is not viable when a realistic number of design elements exists. This is because, even with few design properties, hundreds of potential combinations, C, exist. Environmental behavior, or environmental design, research of this type is often conducted with captive subjects such as students or advisory panel members or small numbers of paid attendees solicited for an experiment (e.g. Whitmore, Cook and Steiner 1995, Stamps 1998). However, the expectations of large numbers of citizens attending open public meetings cannot be met in this way.

If exhaustive evaluation is not possible, but feedback on elements and their interactions is desired, it follows that the visual evaluation decision support system needs to be able to convert the information from a smaller sample set into a function that will predict outputs for all possible input combinations i.e. it will estimate stakeholder preference for scenarios that may not yet have been modeled or tested, if such scenarios can be defined from feasible combinations of the input parameters. This is a very efficient process because it eliminates the need to score all possible combinations of inputs. It also provides more analytic information than traditional visual assessment or the Visual Preference Survey.

Therefore, the problem domain is challenging. Standard statistical methods cannot generate useful properties with such small sample sizes and limited coverage of the state space.

For several decades, fuzzy set approaches have been used effectively to model analogous complex nonlinear systems under conditions of sparse data and high uncertainty (e.g. Zadeh 1965, Ridgley and Ruitenbeek 1999). The authors designed a fuzzy-set based system modeling approach for visual evaluation called Casewise Visual Evaluation, or CAVE (Bailey et al. 2001). The aim of CAVE is to map the output, y, i.e. mean stakeholder preference for the scenario, to the known inputs x_1, x_2, $x_3 \ldots x_n$, which in this case are the planning, design and management parameters that define the properties of each visualized scenario. A relatively small set of sample evaluations can be used to generate a community knowledge base covering all potential configurations. The software *FuzzyKnowledgeBuilder* is used to build the community knowledge base. A series of neural network algorithms are employed to build outputs around the known points. The functions are compiled and saved as a multidimensional mapping function that relates the output to all of the inputs across the entire range of every input parameter. Once verified and built, the community knowledge base exists as a multi-dimensional inputs-output model that can be interrogated by the design team across this full range of all input parameters. The community knowledge base now functions as a decision support system. The research team and project managers can inspect this knowledge base, examining the sensitivity of stakeholder preferences with respect to various input parameters. It also allows for trade-off analyses, or constrained optimizations, to be performed in cases where the community knowledge base must interoperate with other factors e.g. cost, or areas of the design envelope that are not feasible for constructability reasons, etc.

Various tools exist to facilitate the inspection of the community knowledge base. A knowledge slicer allows a 3D graphical output to be presented. Two input variables (x_1, x_2) are presented across their entire ranges, and the output (z) is mapped to a surface. Figure 1 shows a sample output.

Exhaustive inspection of a range of these surfaces allows the team to interpret likely public response to changes in one input parameter, with all other inputs held constant. By working sequentially through each input, high spots or plateaux, i.e. combinations that the community values highly, can be identified. Also “sinkholes” or undesirable areas in the planning and design envelope, can be identified. The outputs are categorical, corresponding to numerical ranges for each parameter. The principle behind fuzzy logic application is the trading of false precision for greater accuracy between broader categories of output. For example, this means that, unlike a multivariate statistically-based analysis, discrimination based on numerical outputs within a given output class is not reliable. However, the discrimination between categories is robust. Normally, five categories or more are used to map output ranges. It must be borne in mind that this method is not directly comparable to standard statistical approaches to visual decomposition, because statistical methods cannot function at all with such limited data input.

The Design Vocabulary

CAVE requires that the inputs be parameterized and that each parameter be divided into classes that are meaningful for the design team. Each potential end-state can therefore be defined by a specific set of input properties and the corresponding visualizations can be engineered as composites of the input factors. To create a meaningful design vocabulary, the team hosted a series of internal meetings with various stakeholders. The team evaluated the correspondence between what citizens felt they were responding to in the visualizations, and what the design team needed to know to convert these perceptions into usable, actionable policies, plan and design guidelines. Following the process explained by Anyaegbanum (2010) key valuation clusters among stakeholders were identified to assist in this process.

The team then held a series of project meetings at which these values were brainstormed, examined and converted into properties that could be represented using dynamic visualization. This process is described by team members (Grossardt et al. 2010). Table 2 shows the land use matrix, i.e. the design vocabulary, for the PGDP End State visions.

Table 2. PGDP land use matrix for CAVE application

Variable	Categories					
PGDP Land Use	Nuclear plant	Heavy industry	Light industry	Recreational	Wildlife Management Area	Ins control
Wildlife Management Land Use	Recreational	WMA				
Waste disposal alternative	Removal	Part onsite	All onsite			
Legacy waste	Dig up	Leave as is				

Some properties, for example, the concern posed by the dissolved solvent plume, could not be incorporated directly into the visualizations. A GIS planform image of the site with plume extents for different time horizons, at small scale, was selected as a secondary presentation medium. The intent was to present the estimated spatial extent, depth and intensity of this plume for inspection in tandem with the proposed land use visualizations.

The visualizations are shown, and repeated if required. Audience members suggest navigation through the model to investigate the scenario at different scales and from different perspectives. Then the visualization is scored for suitability. Typically, a single criterion termed “suitability” is the metric used for evaluation. The scale used is a variation of a Likert system with a range from 1 (extremely unsuitable) to 9 (extremely suitable). The number of participants, mean score and standard deviation are shown in real time to the audience. The next visualization is then presented and scored, and so on, until all have been evaluated. The outcomes are ordered by mean score, and by standard deviation. Visualizations are inspected again. The team initiates an audience feedback session, asking them to brainstorm their reactions to the scenarios, either positive, or negative. The verbal commentary is appended to the output file. In some cases, the commentary is then scored for significance by the entire group. This data is valuable to the project management. It clarifies, for example, why some visualizations show high standard deviations and others do not, and whether certain features elicit bimodal response patterns from the participants.

Anonymity is preserved by means of the electronic polling system. Each keypad possesses a unique identifier. At open public meetings, the team does not record who takes possession of which keypad and therefore all valuations are recorded anonymously and simultaneously. Moreover, all participating stakeholders can see these features of the process during the meeting. These properties of transparency and integrity resist interest-group gaming and they are critical in terms of delivering high levels of process justice from the viewpoint of the stakeholders. These properties account for a portion of the high performance documented in previous SPI evaluations (Bailey and Grossardt 2010).

Similarities and differences with respect to previous CAVE applications

Dynamic visual evaluation using CAVE has previously demonstrated high performance for design support in large-scale planning and infrastructure applications such as noise wall design (Bailey et al. 2006); context-sensitive large bridge design (Bailey et al. 2007); community-driven visioning for transit-oriented development (Bailey et al. 2007); and integrated transportation and land use planning (Blandford et al. 2008).

Several dimensions of the PDGP end-state visioning problem are similar to these previous cases. For example, the complexity of presenting attributes of utility and disamenity in the same visualization is similar to the TOD case. These divergent valuations are captured implicitly in the “suitability” criterion and the reasons for a high standard deviation are disaggregated using verbal feedback, if prompted by participants during the discussion phase.

Another similarity is the time horizon over which the valuations are elicited. The values of the end-state are intended to be more than those pertaining to an instantaneous snapshot at the instant of plan approval. For the long-term viability of the final PGDP end-state, when a management plan is developed using the end-state visioning process as input, certain commitments are envisaged which could include land use controls, development planning, access controls, site management programs and so on. Likewise, investment in the built environment of TOD entailed sunk or unrecoverable costs and the fixed capital. When participants evaluate the visualizations using the “suitability” criterion, they are not only evaluating visual amenity, there are “bundling” other valuations into their score including perceived risk, environmental impacts, economic impacts and other factors. The team does not seek to make explicit which factors could, or should, be included. Respondents self-evaluate the meaning of “suitability.”

However, there are notable differences compared with previous applications. Risk levels of attributes such as the dissolved solvent plume, and the site uses, are not comparable (Freeman and Godsil 1994). In the case of PGDP, these are much more significant than other cases, and recent studies by team members have demonstrated that, to some extent, they are unknown in spatial extent, duration and intensity (Chandramouli, Ormsbee and Kopp 2007). Although the community benefits can be clearly apprehended in the form of the utility of different land uses for the decommissioned plant, the risks cannot be so easily defined and delineated.

The actual extent of the physical plant itself is not enormous, but the area impacted by the PGDP end-state is considerably larger than the plant. The impacted area extends over several counties, across the Ohio river to Illinois, and it includes tens of thousands of citizens and residents, as well as businesses, organizations, recreational land users and other groups from outside the immediate area. The geographic scale at which the visualizations are engineered and presented takes this factor into consideration.

Also, the historical development legacy of the plant and its impact on community valuations of end-state uses extend over a wider area and affect more stakeholders than in the land-use planning case. The plant has a significant history spanning over five decades and the complexity of stakeholder relationships with the plant runs a gamut from acceptance to intolerance. The process described here is designed to depolarize citizen valuations and decouple them from one another, as well as attach meaningful quantitative valuations to all feasible end-state scenarios and allow for reliable and defensible value comparisons. In these ways, the SPI process using CAVE is structured similarly to those used on the previous large infrastructure cases.

Application

The team will be conducting the evaluations at a series of public meetings in the area in spring 2010. Because the SPI protocol is designed to be scalable and modular, the team wishes for the maximum possible participation. The larger the audience, the greater the volume of data, and the more robust the conclusions derived from the community knowledge base. At previous SPI project meetings up to three hundred attendees have been accommodated at each session, although groups of thirty to eighty are more manageable. The meetings will be repeated in the same format at different times, in various locations in the study area, to facilitate the participation of as many in the broader community as possible. The data can be aggregated for final evaluation.

At these meetings, the visualizations will be shown, scored and then some of them will be verbally evaluated by the participants. The verbal evaluations will be of assistance in cases of high, or low, suitability, or where the standard deviations are high i.e. where there is a lack of agreement about the value of the scenario. In this case the reasons why will be elicited. The process will also elicit hidden concerns, and identify value polarities among stakeholders with respect to specific features or parameters of the scenarios. Comparison of data sets across geographic meeting sites, and among specific stakeholder groups, will allow identification of clusters of stakeholders who share similar beliefs.

It is expected that the CAVE database will assist the team in identifying and proposing a much smaller sub-group of scenarios as candidates for more intensive and detailed evaluation by the citizens and by experts including economic development and land use professionals. The SPI process does not involve the selection of a specific, predetermined scenario, or the advocacy of specific, predetermined outcomes on behalf of interest groups or project team and management. These process qualities are essential in terms of developing a measure of trust between stakeholders and the project team, and will ultimately impact the quality, robustness and legitimacy of the final end-state proposals.

The authors will discuss their preliminary findings from these meetings at the Conference.

References

Anyaegabanum, C., Hoover, A. and Schwartz, M. 2010. Use of community-based participatory communication to identify community values at a Superfund site. World Environmental and

Water Resources Congress 2010: Challenges of Change. *Proceedings of the World Environmental and Water Resources Congress* 2010. American Society of Civil Engineers.

Bailey, K. and Grossardt, T. 2010. Towards Structured Public Involvement: Justice, Geography and Collaborative Decision Support Systems. *Annals of the Association of American Geographers* 100(1): 57-76.

Blandford, B., Bailey, K., Grossardt T. and Ripy, J. 2008. Integrated Transportation and Land Use Scenario Modeling by Visual Evaluation of Examples: Case Study of Jeffersonville, Indiana. *Transportation Research Record* 2076:192-208.

Bailey, K., Grossardt, T., Williams, J., Toole, L. and Bryant, B. 2007. Context-sensitive large bridge design using Casewise Visual Evaluation: Case Study Section 2 Ohio River Bridges Project. *Transportation Research Record* 2028:85-92.

Bailey, K., Grossardt, T. and Pride-Wells, M. 2007. Community Design of a Light Rail Transit-Oriented Development using Casewise Visual Evaluation (CAVE). *Socio-Economic Planning Sciences* 41(3): 235-254.

Bailey, K., J. Brumm, and T. Grossardt. 2001. Towards structured public involvement in highway design: A comparative study of visualization methods and preference modeling using CAVE (casewise visual evaluation). *Journal of Geographic Information and Decision Analysis* 5 (1): 1–15.

Bailey, K., Grossardt, T., Ripy, J., Toole, L., Williams, J. and Dietrick, J. 2007. Structured public involvement in context-sensitive large bridge design using casewise visual evaluation: Case study of section 2 of Ohio River bridges project. *Transportation Research Record* 2028:19– 27.

Chandramouli, V., Ormsbee, L, and Kopp, J. 2007. Land Acquisition Study at Paducah Gaseous Diffusion Plant Site Using MODFLOWT Modeling. World Environmental and Water Resources Congress 2007: Restoring our Natural Habitat. *Proceedings of the World Environmental and Water Resources Congress 2007*. American Society of Civil Engineers.

Department of Energy. 2008. *Site Management Plan*. Paducah Gaseous Diffusion Plant, Paducah, Kentucky. Annual Revision, FY 2008. DOE/LX/07-0105&D2.

Freeman, J. and Godsil, R. 1994. The question of risk: Incorporating community perceptions into environmental risk assessments. *Fordham Urban Law Journal* 21(3): 547-604.

Grossardt, T., Ripy, J. and Bailey, K. 2010. Use Of Structured Public Involvement For Identifying Community Preferences For A Superfund Site End State Vision. World Environmental and Water Resources Congress 2010: Challenges of Change. *Proceedings of the World Environmental and Water Resources Congress 2010*. American Society of Civil Engineers.

Jewell, W., Gill, R., Bailey, K. and Grossardt, T. 2009. A New Method for Public Involvement in Electric Transmission Line Routing. *Transactions of the Institute of Electrical and Electronic Engineers.*

Nelessen, A. 1994. *Visions for a New American Dream: Process, Principles and an Ordinance to Plan and Design Small Communities*. Chicago and Washington D.C.: American Planning Association Press.

Ormsbee, L. and Hoover, A. 2010. End State Vision Process for the Paducah Gaseous Diffusion Plant. World Environmental and Water Resources Congress 2010: Challenges of Change. *Proceedings of the World Environmental and Water Resources Congress 2010*. American Society of Civil Engineers.

Ridgley, M. and Ruitenbeek, R. 1999. Optimization of Economic Policies and Investment Projects Using a Fuzzy-Logic Based Cost-Effectiveness Model of Coral Reef Quality: Empirical Results for Montego Bay, Jamaica. *Coral Reefs* 18(4): 381-392.

Steinitz, C. 1990. Toward a Sustainable Landscape with High Visual Preference and High Ecological Integrity. *Landscape and Urban Planning* 19:213-150.

Stamps, A. 1999. Physical determinants of preferences for residential facades. *Environment and Behavior* 31(6):723-751.

Thomas, C. 1998. Maintaining and Restoring Public Trust in Government Agencies and their Employees. *Administration and Society* 30(2): 166-193.

Whitmore, W. Cook, E. and Steiner, F. 1995. Public Involvement in Visual Assessment: Verde River Corridor Study. *Landscape Journal* 14(1):26-45.

Zadeh, L. 1965. Fuzzy Sets. *Information and Control* 8(3):338-353.

Zube, E., Sell, J. and Taylor, J. 1982. Landscape perception: research, application and theory. *Landscape Planning* 9:1–33.