ACHIEVING CONSENSUS IN ENVIRONMENTAL PROGRAMS

Harold A. Kurstedt, Jr., R. Martin Jones, J. Andrew Walker, and Louis I. Middleman
Management Systems Laboratories
Virginia Polytechnic Institute and State University
1900 Kraft Drive, Blacksburg, VA 24060

ABSTRACT

In this paper, we describe a new research effort on consensus tied to the Environmental Restoration Program (ERP) within the U.S. Department of Energy's Office of Defense Waste and Transportation Management (DWTM). We define consensus and explain why consensus decisions are not merely desirable but necessary in furthering ERP activities. As examples of our planned applied research, we first discuss Nominal Group Technique as a representative consensus-generating tool, and we conclude by describing the consensus-related mission of the Waste Management Review Group, established at Virginia Tech to conduct independent, third-party review of DWTM/ERP plans and activities.

INTRODUCTION

The key defining feature of the organizations responsible for ERP activities isn't autonomy but "polyonomy," a term we've invented to signify the diffusion of responsibility and power among many agencies and sub-agencies. The U.S. Department of Energy (DOE) is what we call a "Government Oversight Agency" (GOA); GOA's at all levels of government (federal, state, and local) must implement laws made by the Congress and by state legislatures--must coordinate their overlapping roles and responsibilities. Further, they must learn to cooperate in an arena that rewards adversariality, and they must learn to maintain this cooperation over long periods, to adapt to inevitable change.

RESEARCH CONTEXT

The popularity of books and articles on the "new" management and manager, emphasizing consensus rather than edict, is in part the result of a drastic increase in the number of organizations whose authority isn't complete and autonomous but partial and polyonomous --literally, "many-portioned" or "many-ruled." Even the few autonomous organizations left today are undergoing changes to be or remain competitive. Although many people are "only vaguely aware" of it, "modern society is a complex of interdependent groups or teams" (1). Autocratic management may work in small, centralized, independent, isolated organizations where one person's word is law and where formal and informal public scrutiny is low. Nothing could be further from describing the characteristics of GOA's: large, decentralized, regulated, dependent, and in a veritable fishbowl of public scrutiny. Getting things done within and among GOA's means implementing national policy derived from Congressional legislation. A GOA manager's first step toward such implementation normally meets with resistance from outside agencies and from various levels of parallel sub-organizations within the manager's own organization. That's no accident. Indeed, in principle it's a fine setup: a multitude of mandated checks and balances on everything actually or potentially affecting the public good.

Consensus within such a polyonomous system requires "acceptance-level decisions" (2) in addition to and as distinguished from "technical-level decisions." The latter refer to decisions about what will or won't "work"--for example, what kind of treatment will or won't reduce PCB contamination at a certain waste disposal site to an acceptable amount. The point is, whether there be only one or a number of equally effective and efficient technical solutions to such a problem, *none will be implemented unless all affected parties agree (or are compelled).*

The consensus necessary to achieve acceptance-level decisions within, between, and among organizations is often the result of what are called "informal" and "lateral" processes, as contrasted with those of the formal hierarchy. Galbraith (3) stresses the importance of not leaving informal structures to chance. "These informal processes," Galbraith writes, "are thought to arise spontaneously and are the processes through which most organizations accomplish their work despite the formally designed structure. A typical point of view is, 'If we had to go through channels, we would never get anything done.' The point of view being taken here is that these informal processes are necessary as well as inevitable, but *their use can be substantially improved by designing them into the formal organization* .... [A] more important reason for formalization is that these processes do not always arise spontaneously from the task requirements, especially in highly differentiated organizations. *When the relevant participants have different and sometimes antagonistic attitudes...and are separated geographically, the effective use of joint decision making requires formally designed processes* " (italics added).

DEFINING CONSENSUS

Consensus denotes "collective opinion or concord; general agreement or accord"(4). Further elucidation comes through the word's origin in the Indo-European root *sent* , meaning "to head for, go" following the prefix *con* - meaning "along with" or "together." This etymology suggests why consensus is necessary to the effective and efficient work of a system (like a GOA) with many parts. A *system* is a collection of entities, related by structure or communication, such that a perceptible or measurable change in one part causes a perceptible or measurable change in all the other parts. And *work* is the application of a force through a distance. For efficient work, the sum of the movement of the system's interdependent parts must point toward the objective and exert maximum energy in that direction. GOA's need consensus about these two things: the objective and the movement--the ends and the means. If you have the objective but not the ability to move as a unit, you won't get there (zero or low efficiency). If you have the ability to move but don't know where to go, your energy is pointless and will be wasted (zero or low effectiveness).

Defined in this way, consensus is pragmatically and morally neutral. Consensus on means and ends is a necessary condition of progress toward an objective, but it isn't sufficient. For, as a "collective opinion," consensus by itself implies neither accuracy nor morality nor even feasibility. Opinion is, in Plato's words, "something between ignorance and knowledge." The better opinions are those closer to knowledge, and the way you show this closeness is through adequate supporting information. An opinion, though collective among a certain group, can still be bad (evil); wrong (incorrect); and difficult, perhaps even impossible to implement if it has to fight against a strong or stronger opposition. Consider the ancient consensus achieved by Pontius Pilate and his lieutenants, the ongoing consensus among the membership of the Flat Earth Society, and the 1982 Congressional consensus that high-level radioactive waste be permanently disposed of in a deep geologic repository.

Consensus as a State Variable

Consensus, then, is a *parameter* , a variable that must always be taken into account in describing the state of a system. Our research focuses on interactions within and among systems of people (GOA managers)--and between these systems and the population they serve; we're using the term as a gross or macro-level "state variable" characterizing the degree to which a group of people behaves as one person. It's the measure of a group's tendency, as it moves through time, to behave as a unit--irrespective of the force behind this tendency (reward/punishment), the source of the movement (internal/external), or the level of the source (horizontal/vertical/mixed).

Consensus Distinguished from and Favored Over Compulsion

When consensus results from a directive that doesn't fit what the group would have come up with had it been asked, we call it compulsion, meaning that the environment (i.e., one or more people coupled with one or more conditions) forces the group to behave in a certain way and prevents the group from behaving in any other way.

No group is ever completely free of compulsion. The environment always more or less constrains this freedom, making certain behaviors undesirable or impossible and requiring certain behaviors nobody likes. Galbraith (3) calls compulsion "forcing," which he defines as "power or position or knowledge being used to force a preferred alternative on the rest of the group." Forcing, says Galbraith, will lead to ineffective decisions if it is the dominant mode. If one function or dominant department always forces, then there is no need for a group effort, since information from other departments is ignored. *Suboptimal decisions and poor implementation result when a forced solution is based on local information in the presence of interdependence The preferred approach to conflict resolution therefore is to use confrontation and problem solving backed up by occasional forcing when lack of agreement stymies the group* " (italics added).

ACTUAL VERSUS PERCEIVED CONSENSUS

When we look at consensus as common or unified behavior, we see consensus if the behavior shows it. The observable behavior, however, may or may not be congruent with group members' unobservable internal states. We can define the degree of actual as opposed to perceived consensus as the difference, if any, between the assumptions an external observer or a group member would make about a group's attitudes based on the group's objective (visible) behavior, and the group's subjective (invisible) thoughts and feelings about this behavior. Consensus can be more perceived than actual, or vice versa, both from the perspective of a group member (including the leader or the person who called the group together) and from the perspective of someone outside. The importance of the distinction between actual and perceived consensus may be seen in the following example.

Avoiding the Trip to Abilene

Unhealthy agreement--the false pretense of consensus born of fear of self-disclosure--can move a group to unwanted behavior as much as, perhaps even more than, an excess of explicit conflict can stymie them. The "Abilene Paradox" (5) is the cautionary tale of a Texas family that reaches a false but powerful consensus and ends up doing something none of them actually wants to do. They somehow find themselves driving a dusty 106 miles in an un-air-conditioned car on a 104-degree summer afternoon from Coleman to Abilene to eat unpalatable food in a fourth-rate cafeteria, instead of doing what they all really want--to stay out on the electric fan-equipped screened-in porch, play dominos, listen to the radio, drink lemonade, and chat. Having returned from Abilene thoroughly disgruntled, they reveal their true thoughts and feelings--and their bewilderment at how they ever decided on that stupid trip.

Consensus like that which motivated this family is false to the group members' thoughts and desires. In its effects, however, their consensus is actual, and therein lies the danger. We humans are skilled at hiding thoughts and feelings from others, but we forget others are just as good at hiding theirs from us. To avoid and prevent trips to Abilene, group members (meaning, one time or another, everyone) must learn the skills of self-disclosure and learn to overcome the fear of risking this disclosure.

HOW MUCH CONSENSUS IS ENOUGH?

Our behavior-based definition of consensus provides an answer. A group member's behavior, either toward or away from the group's ostensible objective (what an observer would predict to be the effect of the group's decision) is analogous to a vector. If the vector sum of all members' behavior is toward the ostensible objective, then there's enough consensus. And "enough" isn't strictly a matter of numbers, of how many participants' vectors point the same way, but of their relative magnitudes, and moreover, the changing magnitude of their vector sum over time. Over time, the objective will be reached if and only if the sum of all consensus vectors is greater than the sum of all antagonistic compulsion vectors.

There's no necessary one-to-one correlation between a group's achievement of consensus on an objective and the reaching of that objective. Consensus by itself, though necessary to reaching the objective, is not sufficient. The

consensus reached in a football huddle may be total, yet the play may fail because, literally and figuratively, it bumps into another and stronger consensus. In terms of the effects of GOA group decision-making, we must measure consensus by discovering how to measure the extent to which the actions and conditions (including costs and schedules) that should follow from the group decision actually do follow, and by how long they remain in effect in comparison to how long the decision-makers expected them to remain in effect.

WHY CONSENSUS IS NECESSARY FOR GOA'S

Nowhere is interdependency more apparent than in and among GOA's and in interdependent programs like the ERP. These organizations differ from product and service organizations in the private sector and from other government field offices and public works units. Whereas the latter are close to the provision of services to the public, GOA's are close to the legislative bodies from which they take their mandates. As such they are the interface points between public policy-making and implementation and must participatively interact with legislatures, other GOA's, and other elements within their own organizations. However, when called upon to solve problems GOA's don't have tools, methods, or processes uniquely tailored to fit their participative environment. Rather they have tools, methods, and processes designed for hierarchically structured environments.

Given so much to do with such a (relative) paucity of resources, those people responsible for managing and implementing the ERP must invite consensus on task and funding *prioritization* . Numerous organizations at all levels of government, as well as private interest groups and individual citizens, are bound to be dissatisfied no matter how the funds are spent, *unless there's consensus on the priorities* But what--this is the key question-- *nationally* constitutes the greatest overall benefit to be realized from the available dollars? It's one thing to prioritize activities vertically at each DWTM field site (and that's no small matter in itself, for the potential hazards of many waste sites may take years to estimate accurately). It's quite another to prioritize activities horizontally across all field sites. There are simply too many squeaky wheels and too little grease. The first formal activity of the Waste Management Review Group, discussed at the end of this paper, is to review a prioritization model, the Program Optimization System (POS), and recommend ways to assess the degree of consensus likely to result among stakeholders nationally.

If short-term consensus about ERP priorities is hard to achieve at all, it's surely harder to maintain in the moderate run, and it's exponentially more difficult in the long run. This is partly because, like any organization, this DWTM and ERP will see changes in leadership and personnel (nobody working in DWTM will be working there sixty years from now). It's also because GOA managers are, as Kotter (6) says, "rewarded almost entirely for short-run performance." Though DWTM's programs are all long-term efforts, funding comes just once a year, and it's impossible with any confidence to allocate dollars you don't yet have. As a result of this short-run focus, "they spend far too much... of their discretionary resources trying to keep current processes effective and efficient, far too little of their time and other resources trying to create or maintain a coalignment, and far too little effort trying to create adaptive element states."

HOW WE CAN APPLY THEORY TO UNDERSTAND ERP PROBLEMS AND FIND SOLUTIONS

We're at the very beginning of our study, with more questions than answers--and probably with not enough questions. How can we move forward? In this section we give two examples of our research approach. In the first example, we start with a popular method for getting consensus, analyze this method, and consider its usefulness both as a vehicle for learning about consensus and as an aid to ERP success. In the second example, we start not with a method but with a group--with the assembling of an expert review group--so we can observe and understand how that group's recommendations can affect consensus beyond the group, in the wider population of which the members constitute we hope, a representative sample. In short, our research involves intra-group consensus, inter-group consensus, and (most importantly), extra-group consensus.

Examining the Nominal Group Technique (NGT)

We want to find out what works and why it works. At first we'll find out more of what works (and doesn't). We'll apply tools and methods people have tried to use for generating consensus, and see what happens. We assume no one method will work all the time. What works will depend on the group's characteristics and the problem to be solved or the decision to be made. When we find something that works, what we'll be finding is something that works in a particular, constrained situation. If we understand the constraints, we ought to be able to make it work in a similarly constrained situation. And if we get things to work enough times in enough differently constrained situations, we can begin to generalize about why things work.

The Nominal Group Technique (NGT) was first developed by Van de Ven and Delbecq (7). NGT is a method--a series of tools applied in sequence--to provide free and equal expression and ranking of opinions in groups which otherwise might be dominated by certain individuals or certain paths of thought. Both the tools for idea proposal and the tools for ranking limit the potential for conflict between group members. The four steps of the NGT proposed by Van de Ven and Delbecq are as follows: a) Individual members first silently and independently generate their ideas on a problem or task in writing; b) This period of silent writing is followed by a recorded round-robin procedure in which each group member (one at a time, in turn, around the table) presents one of his or her ideas to the group without discussion. The ideas are summarized in a terse phrase and written on a blackboard or sheet of paper on the wall; c) After all individuals have presented their ideas, there is a discussion of the recorded ideas for the purposes of clarification and evaluation; d) The meeting concludes with a silent independent voting on priorities by individuals through a rank ordering or rating procedure, depending upon the group's decision vote. Other practitioners expand NGT to six steps (8, 9),

affording more interaction among group members: e) After the voting and ranking, participants consider the aggregated results to measure the extent to which they've supported their true positions or taken a trip to Abilene. At this point they may revise their rank ordering of ideas; f) Finally they consider the resources available to implement their high-priority action items, choose a set of these items to scope for implementation, and divide into smaller groups to begin the scoping process, which involves measuring the feasible application of available human, funding, and material resources within known or estimated time constraints.

"Nominal" is the key word describing this method; the group that meets for an NGT session isn't (or isn't necessarily) a regular work group within one organization. It's not a *real* group (a group sharing common aims and values and therefore predisposed to cooperate) but a group in name only (a *nominal* group). The question arises whether consensus is possible within a nominal group; if not, then you have to take a nominal group and somehow transform it into a real group. Can NGT help in this transformation? We need to ask such questions because the set of ERP stakeholders would have to be called a nominal group.

At least one NGT study, Mahler's (10), suggests group dissatisfaction with the resulting decision is high; however, the results come from 45-minute applications with student groups working on unsolvable problems in which they had no stake. Studies made on day-long applications outside classroom situations and with real, pressing, solvable problems have found more favorable results. Mahler is correct in linking the use of pure NGT primarily to the generation rather than to the evaluation of ideas. Indeed, in its original formulation, since ranking is numerical and done by straw vote, a "pure" NGT application gives little chance for conflict between members and therefore little opportunity for persuasion, bargaining, compromise, or what can loosely be called the politics of consensus. In the generative and in portions of the evaluative stage of consensus reaching, this is an advantage; later, however, a group would probably wish to tailor the technique to provide for surfacing enough essential conflicts (ad hoc rather than ad hominem) to increase the probability of results participants won't torpedo once they leave the meeting.

NGT is a popular technique. Though until just recently we never felt the compulsion to research it systematically and longitudinally, at Management Systems Laboratories we've collectively had more than 40 person-years experience conducting, observing, and participating in a wide variety of NGT-type sessions, several recently within the ERP itself. We know what works in one circumstance will fail in another. We've seen an apparent deadlock become a consensus as if by magic, not in the meeting itself but in the hallway out to the parking lot. We've seen an apparent consensus fall to pieces when participants realized some were using a term to mean one thing and others something different. We see NGT as a group of tools and we believe there must be a fit between the users and the problem to be solved. Just as with computers, you don't simply accept NGT because you heard it worked somewhere else.

We can hypothesize an ideal group on which NGT should work. Such a group, of about eight to 12 people, would likely share interest in common issues and be all at about the same level in their organizations. They'd probably be more divergent than convergent in their information gathering styles, preferring open-endedness to closure. NGT would both satisfy their need to brainstorm--it's a great idea-generating technique--and constrain their divergence through the requirement to vote and rank.

Okay. Say (no simple matter) we find the tools that work for our ideal nominal group. We're likely to encounter an actual ERP nominal group that's "all wrong"--convergent if we want divergent; at levels all over the hierarchies of their respective organizations, larger than we think any group should be to function well together but needing to be that size or else we lose legitimate stakeholders who if not included will vitiate the efforts of the rest. Well, we don't have a perfect tool, so we try to combine or modify tools we know to get as close as possible. In our research we must ask, What are the important characteristics of groups? and What are the important characteristics of tools? Then we can hope to synthesize an answer to the question of the relationships between these sets of characteristics and why they hold.

We must identify all tools, methods, techniques that might apply to our real-world nominal groups and keep a running list, a living list. We'll do test runs, we'll discover that a certain tool works nine times out of 10 and use that fact as a reference point. We'll ask, if you change the group in such-and-such a way, what must you do to the NGT to make it work? We'll start to understand what parts of NGT work in different situations and begin to generalize and to understand why. In our attempt to understand why and confirm that understanding, we'll first generalize tool by tool; then group of tools by group of tools. Finally, we'll hope to generalize in terms of all consensus tools. When we can do this, we can predict outcomes and prescribe the sets of tools likeliest to facilitate consensus within various given situations.

The Waste Management Review Group (WMRG): A Forum for Consensus

Whereas with NGT we start with a tool or a group of tools (a method) and try to figure out how that method yields or doesn't yield consensus, with WMRG we start with a group, a group specifically selected for the necessity of its members' knowledge and experience to the furthering of the ERP. The idea for such a group came from the realization that another group, the Technical Review Group (TRG) convened by Martin Marietta Energy Systems, Inc. was increasingly spending its time on policy issues rather than on purely technical issues. *Since* both kinds of issues are vital to the ERP, it was decided to form a second group devoted entirely to policy. The WMRG's chartered scope is to provide 1) objective reviews, evaluations, and assessments of current plans, projects, and activities related to DWTM policies and actions mandated by Congress, the Executive Branch, and the DOE; and 2) research, analysis, and communication necessary to substantiate the reviews, evaluations, and assessments and provide useful

information, conclusions, and recommendations to a new organization, the Waste Policy Institute (WPI).

The WMRG was chartered and convened under the auspices of WPI, a not-for-profit university-related corporation at Virginia Tech. The principles for assembling the WMRG were instinctive, intuitive. It had to be done that way because we don't know enough about GOA team building processes to have put it together scientifically. What we tried to do was exhaust the categories representing those constituencies in the general population which must form part of the consensus necessary ensure ERP success. The five categories from which we chose members are universities, industry, States and Indian tribes, the Federal government, and public interest groups and professional societies.

Though not all members know each other, the WMRG starts out probably closer to a real group than a nominal group. But though for this reason it should find it easier to achieve consensus on prioritization in its review of the POS, what DWTM needs is far beyond the consensus of that group. What DWTM needs us to do in our research is figure out how to configure WMRG so it's a faithful microcosm of the population it needs to represent. To put it another way, our job wasn't to set up the group. Setting up the group was a necessary precondition, like setting up experimental apparatus. Our research mission is to discover the extent to which our partitioning of the ERP-related world has yielded the right constituencies--all the right ones--so this wider world will look at the WMRG's conclusions and adopt them because the WMRG's consensus is representative; because the WMRG sample represents the relevant population.

The convener of the TRG attends but is not a member of the WMRG. The convener of the Ad Hoc Waste Contractors Group attends but is not a member of the WMRG. And I, the convener of the WMRG attend but am not a member of the TRG and the Ad Hoc Waste Contractors Group. We call it "interdigitation." We watch the extent to which consensus on technical issues and consensus on policy issues combines to yield a larger, synergistic consensus. The WMRG and related activities thus provide an ideal real-world laboratory for our research, a perfect opportunity to study intra-group consensus, inter-group consensus, and extra-group consensus.

ACKNOWLEDGEMENTS

The preparation of this paper is funded by U.S. Department of Energy (DOE) Grant No. DE-FG02-88DP48058. Management Systems Laboratories thanks the DOE's Office of Defense Waste and Transportation Management for providing us a real-world laboratory for the research, development, and testing of state-of-the-art management tools and the frameworks for understanding how to make them successful.

REFERENCES

1. Dyer, William G. *Team Building: Issues and Alternatives* , 2nd edition. Reading, Massachusetts: Addison-Wesley Publishing Company, 1987.
2. Vroom, Victor H. and Phillip W. Yetton. *Leadership and Decision-Making* . University of Pittsburgh Press, 1973.
3. Galbraith, Jay. *Designing Complex Organizations* . Reading, Massachusetts: Addison-Wesley Publishing Company, 1973.
4. William Morris, ed. *The American Heritage Dictionary of the English Language* . Boston: Houghton Mifflin Company, 1976.
5. Harvey, J. "Managing Agreement in Organizations: The Abilene Paradox." *Organizational Dynamics* , Summer 1974, pp. 63-80.
6. Kotter, John P. *Organizational Dynamics: Diagnosis and Intervention* . Reading, Massachusetts: Addison-Wesley Publishing Company, 1978.
7. Van de Ven, Andrew H. and Andre Delbecq. "The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes." *Academy of Management Journal* , vol. 17, no. 4, 1977, pp. 605-621.
8. Rohrbaugh, John. "Improving the Quality of Group Judgment: Social Judgment Analysis and the Nominal Group Technique." *Organizational Behavior and Human Performance* , vol. 28, 1981, pp. 272-288.
9. Sink, D. Scott. "Using the Nominal Group Technique Effectively." *National Productivity Review* , Spring 1983, pp. 173-184.
10. Mahler, Julianne G. "Structured Decision Making in Public Organizations." *Public Administration Review* , July/August 1987.