

Draft Deliberative Document - For Discussion Purposes Only
Not Releasable Under FOIA

Date: 18 June 2004

To: Alan R. Shaffer,
Director, Plans and Programs, ODDR

From: Don J. DeYoung

Subj: Military Value and Capacity Analysis of DoD Laboratories and Centers

Encl. (1) Analysis: *Shadows on the Wall: The Problem with Military Value Metrics*

1. I have taken the opportunity to update my ongoing analysis of the TJCSG Military Value scoring approach, sent on previous occasions to both you and Dr. John Hopps, the late Deputy, Director Defense Research and Engineering (earlier versions were dated 17 February, 1 April, and 24 May).

The changes to our approach, mandated on 4 June by the Infrastructure Steering Group (ISG), had a nominal effect on the outcomes for the test case that I alerted you of in the 24 May version. Minimal as the changes were, I nevertheless wanted to make you aware of them before Data Call #2 is issued to the field next week. The ISG's changes served to ameliorate the situation, albeit very slightly. The results of the trial run on a world-class research group are now merely irrational, instead of patently absurd. Finally, you will see that I offer some comments within this memo regarding our proposed approach to capacity analysis.

2. Military Value. In the 24 May version, I referred to the findings of the National Defense University (NDU) Lab Relevance Study. The NDU study team, and in particular Admiral H. Gehman (later appointed as the chair of the Columbia Accident Investigation Board), had singled out a DoD biosensors research team for their work and its potential criticality to public safety and national defense. ADM Gehman specifically stated that this world-class research group,

“demonstrated bio-warfare detection systems that are light years better than anything I have seen in DOD, CDC, DTRA, the Raid teams or anywhere else. This Center may be a life-saver sooner than we think.”¹

Using the TJCSG's official analytical approach, I scored this bona fide world-class research group (with real data) against two hypothetical groups to see how it fared. The two hypothetical groups, Projects X and Y, have large budgets but were otherwise deliberately designed to be unexceptional and unproductive. *This was done to determine whether our analytical process could in fact recognize world-class talent and evaluate each site accurately for its military value.*

In short, if we cannot identify the exceptional talent within the DoD labs and centers, or the field impact they make, we will risk doing extensive damage to long-term national security. A sound analytical process would obviously rank the world-class group highest (by a significant amount) given its own elite credentials, and given the unexceptional personnel and non-existent operational impact of Projects X and Y.

¹ Section 913 Report #1: *Sensors Science and Technology and the Department of Defense Laboratories*, (National Defense University: March 2002), p.31.

Draft Deliberative Document - For Discussion Purposes Only
Not Releasable Under FOIA

The Results. Despite the ISG's changes, we still have a big problem. The 15-person (*13 PhDs, 2 MSs*), world-class research team — *with one technology transition to the U.S. Marine Corps, a successful rapid response project for the U.S. Fleet in Bahrain, a Homeland Security Award, a Presidential award, a Technical Society fellow, CRADA income, 3 patents and a license* — ranks second, with a total Military Value (MV) of **3.50** (1.93 for Intellectual Capital and 1.57 for Operational Impact). *This is little more than half of the top-ranked Project Y's score of **6.06** (3.00 for Intellectual Capital and 3.06 for Operational Impact). Project Y is a 35-person project with only 2 PhDs, no awards or recognition, no product, and no impact.* It does have a \$15 million dollar budget.

Even more amazing, the world-class research team ranks *only 18% higher than Project X's score of **2.96**. Project X is an unexceptional, 2-person contract shop, with no recognition, no product, and no impact. But like Project Y, it has a fat wallet.*

Another disturbing sign is how insensitive these results are to artificial increases of exceptional talent. A quick experiment reveals the problem. If we add 10 Nobel Laureates, all with PhDs with more than 20 years experience, to the world-class group, it still finishes second to Project Y in Intellectual Capital (2.61 versus 2.85).

The reasons for these irrational scores are not surprising. Results like this were predicted in the earlier versions of the attached paper, as well as in numerous emails to the Capabilities Integration Team (CIT) and in meetings of both the CIT and TJCSG. The current version, which is provided as enclosure (1), contains discussion of the ISG changes and a detailed description of the above test. I should note that one arithmetic error in the last version was found and corrected.

The attached analysis shows that the People Metrics are “dumbed-down” to the point where we will be blind to exceptional talent. This jeopardizes the success of our entire endeavor. Absurd point value compression and use of an inappropriate DAWIA certification metric for world-class S&T performers are the major reasons for the problem. The other root flaw is that *dollars* (a lame surrogate for what some TJCSG members generously call “product”) have achieved relative rock-star status as a preferred metric for military value.

Any community-hired BRAC consultant, worth his or her salt, will zero in on these problems a means of discrediting the study. For example, our compressed point values (e.g., 3 points for a PhD, 2 for a MS, and 1 for a BS) will be vulnerable to ridicule by anyone with the statistical aptitude of an avid baseball fan and an understanding of the non-linear difference between a PhD and a BS. It is unfortunate that my early proposals to eliminate the point compression failed to gain any traction.

The evidence strongly suggests that we have a fatally flawed study on our hands. If our approach to assessing military value is, as the attached analysis shows, blind to exceptional talent and to real field impact, then there is an unacceptably high probability of making serious mistakes. *BRAC actions built upon defective military value scores will almost certainly damage the DoD's ability to develop new defense technologies. And our capacity analysis will dictate the extent of that damage.*

A large calculated excess capacity within the DoD RDT&E infrastructure will cause large movements of workload. After the military value scores (which are almost certain to be inaccurate) determine the “winners” and “losers,” these large movements will almost certainly damage some, if not many, technical centers of excellence. Large excess capacity will maximize

Draft Deliberative Document - For Discussion Purposes Only
Not Releasable Under FOIA

the damage; small excess capacity will minimize it. That brings me to some concerns that I have about our approach to calculating excess capacity.

3. Capacity Analysis. Selecting an effective unit for measuring RDT&E capacity is critical. Unfortunately, it is not easy to find one without disadvantages.

"Excess capacity is a simple concept when applied to most installations, such as naval stations, air bases, hospitals, and test centers. Fewer ships need less berthing, fewer aircraft need less hangar space, fewer personnel need fewer hospital beds, and reduced weapons procurement equals less test range use. But unlike conventional bases, there is no direct relationship between size of the force and that of Laboratory infrastructure (for example, buildings, roads, and utilities). If there were such a relationship, the Navy would not have resorted to a surrogate metric (that is, work-years) for capacity during the 1995 Base Realignment and Closure (BRAC) round...

...the surrogate metric counted only in-house work-years, which means contractor work-years were excluded. This was not an oversight. Contractor numbers are notoriously hard to verify. With the high stakes of a BRAC, this raises the risk of fraud or, almost as bad, of rumors of it. Nevertheless, contractors perform about half of Navy RDT&E, and a great many of them work at the Laboratories and use their infrastructure. Therefore, the metric provided an incomplete picture, yielding inaccurate conclusions...this is like counting only [hotel] guests who occupy even-numbered rooms.”²

For BRAC-05, the TJCSG is taking a more complicated approach than the Navy did in BRAC-95. I believe the TJCSG’s approach has one advantage, but it also has several major problems that are potential “showstoppers.” On the positive side of the ledger, by deciding to count on-site contractor work-years, the TJCSG fixed the problem identified in the excerpt above. There is of course the downside of verifying the numbers of on-site contractors, but this metric does stand the best chance of producing an accurate estimate of a site’s true capacity.

Unfortunately, the TJCSG has developed four additional units of measure that will cause major problems for us down the road. They are (a) Acquisition Category (ACAT) programs (both numbers of and funding levels) as a capacity unit for Development & Acquisition (D&A), (b) extramural funding as a capacity unit for Science and Technology (S&T), (c) square footage, and (d) a Force Structure Adjustment (FSA) to be based on collective expert military judgment.

The first problem with the additional units is the complexity. We are making the job harder than it needs to be. The following advice is based on Service-specific experience, but it could help us sort things out.

As a former member of the Navy Base Structure Analysis Team (BSAT), I can say that the capacity unit for all RDT&E — including the acquisition function — was the work-year. The Navy’s report to the BRAC Commission stated that,

“Budgeted work-years were used as a measuring tool for capacity because of its commonality within the functionally diverse Technical Centers whose products range from published scientific papers to the installation of a new piece of shipboard equipment to the live testing of a new warhead or airframe.”³

² D.J. DeYoung, “The Silence of the Labs,” *Defense Horizons*, No. 21 (January 2003) [http://www.ndu.edu/inss/DefHor/DH21/DH_21.htm]

³ Report to the Commission: Department of the Navy Analyses and Recommendations, Vol. IV (March 1995), p. X-5, [<http://www.defenselink.mil/brac/navy.htm>].

Draft Deliberative Document - For Discussion Purposes Only
Not Releasable Under FOIA

This approach was successful. In BRAC-95, the GAO examined the closure process and decisions of each Service, including their capacity and military value analyses. It found that “the Navy’s process and recommendations were sound.”⁴ In short, the Navy’s use of *work-years as a capacity metric was effective and supportable*. In fact, the same GAO report stated about the Navy process that, “The configuration analysis for this subcategory (Technical Centers) involved complicated assessments of the existing capabilities and requirements for 29 functional categories, such as undersea and surface ship platforms, *across four phases of work: RDT&E, acquisition, lifetime support, and general* (p.96-7).” This comment shows that the work-year even satisfied requirements of functions beyond RDT&E and acquisition. In the end, the Navy recommended 21 lab/center closure or realignment actions, and was successful with all but a few. The process for analyzing capacity stood up to the inevitable challenges by being both defensible and equitable. *In short, work-years did the job — for S&T and D&A.*

(a) ACATs. On the other hand, the use of ACATs (count and funding) is analytically unsound and will be hard to defend. ACAT programs exhibit large ranges in cost and have great variances in complexity. This leads to considerable differences in personnel, funding, and infrastructure requirements between programs — even at the same ACAT level. ACATs have some use in measuring military value, but as a capacity unit they are much too imprecise. Finally, no proponent of using ACATs as a capacity unit has yet been able to assure me that we will not miss *non-ACAT* development programs when evaluating D&A (e.g., See “Major Navy Non-ACAT Programs”⁵). My concern here is that we will have compromised the whole process should we miss counting substantial workload at some sites.

(b) Extramural Funding. To be blunt, this unit is absurd. Dollars provided to external organizations in the private sector, and to other government (DoD and non-DoD) agencies, is not a measure of on-site capacity. How does spending the most money on private sector performance show that one performs the most work? This unit introduces private sector infrastructure into an analysis of the public sector. BRAC is about closing, reducing, and/or realigning government, *not private sector*, infrastructure. Also, by using dollars sent to other DoD organizations, we are ensuring double-counting (or worse) of the same dollar as it passes from sponsor, to program manager, to performer, and to sub-contractor.

Lastly, the unit is based the faulty assumption that the level of dollars is directly related to the workload level of a contract manager. Proponents for this unit should prove there is a one-to-one correspondence between number of dollars and number of required contract managers before this unit is approved. This unit does, however, continue a theme found in our military value approach where *more dollars equals greater operational impact*.

(c) Square Footage. If ever there were a seductive measure of physical infrastructure, it is square footage. It promises simplicity, clarity, and accuracy, but delivers none. Take the case of the estimates of excess capacity contained in the DoD’s March 2004 report to Congress.⁶ In it, the DoD estimates, by using square footage, that the FY09 excess capacity for Army and Air Force labs/T&E sites will be 62 and 18 percent, respectively.⁷ Looking

⁴ GAO, “Military Bases: Analysis of DoD’s 1995 Process and Recommendations for Closure and Realignment” (GAO/NSIAD-95-133) April 1995, p.87.

⁵ <http://www.abm.rda.hq.navy.mil/navyaos/content/view/full/2876>

⁶ Department of Defense, “Report Required by Section 2912 of the Defense Base Closure and Realignment Act of 1990, as amended through the National Defense Authorization Act for Fiscal Year 2003,” (March 2004), p.47 and 52.

⁷ Unlike these estimates using square footage, Navy estimates were based on in-house work-years.

Draft Deliberative Document - For Discussion Purposes Only
Not Releasable Under FOIA

more closely, these estimates are ratios where the “acquisition workforce” divides total square footage. But what is that workforce? Is it both contractor and in-house personnel, or is it a partial picture that uses just in-house government employees? The following evidence suggests the latter.

In a 1997 report to Congress, the Department’s total (all Services, plus Defense Agencies) acquisition workforce was stated to be 617,000 *employees* in FY89.⁸ It happens that the March 2004 report identifies 158,000 in the Army acquisition workforce for that same year — FY89. At the risk of being simplistic, assume an equal share of the acquisition workforce among the Army, Navy, Air Force, and Defense Agencies. An equal share of 158,000 among the four would yield about 632,000, which is very close to the number of employees cited in the 1997 report. It appears then that the 158,000-person Army workforce is made up of government employees, and therefore the estimate does not include the on-site contractors who also use the base infrastructure. This matters because the estimates of a cavernous 62% (825 square feet per person) excess capacity in the Army and 18% (750 square feet per person) in the Air Force, may be overstated at best, and way off at worst.

It should be recalled that since 1996 (a year after the last BRAC round) the Services have been complying with ambitious outsourcing goals levied by the DoD. Many of the positions formerly filled by government workers are now performed on-base by private sector employees. Assuming that 50 percent (which in many places is a significant under-estimate) of the on-site population is comprised of contractors, then both Services have instead about 400 square feet of available space per person. But what does that really mean? Is that a lot of space? Is it too much?

In 1876, Thomas Edison opened what has been called the first research and development laboratory, as well as one of the most productive, at Menlo Park, New Jersey. The lab building was a 100-foot by 25-foot structure with two floors (5,000 square feet).⁹ Edison’s staff numbered 25, which amounted to 200 available square feet per person. This was roughly half the space available to the average member of the Army and Air Force “acquisition workforce.” When one factors in the facility requirements set by more powerful technical equipment that is much more dependent on carefully controlled environments than Edison’s 19th century equipment, maybe 400-sq ft per acquisition worker is to be expected.

There is one last problem with square footage, one that is best revealed by using the example of the Air Force’s McKinley Climatic Chamber. The 6-chamber facility is huge, with its main chamber being 65,520 square feet.¹⁰ Assume the site downsized its acquisition workforce by 18 percent. I doubt anyone would argue that this unique, state-of-the-art facility would then have an excess capacity of nearly 12,000 sq. ft. All 65,000-plus sq. ft. would be necessary whether 100 persons, or 1 person, worked there. Again, the key metric for capacity appears to be work-years, not the amount of space available per person.

(d) The Force Structure Adjustment (FSA). This metric is supposed to somehow identify any of today’s capacity that may not be necessary in 2025 given the military force structure in place then. The plan is to use the expert military judgment resident in the TJCSG sub-groups for such determinations, and the idea is to adjust the estimated required capacity, up

⁸ Office of the Under Secretary of Defense (Acquisition & Technology), “Right-Sizing the Department of Defense Acquisition Workforce”, (28 January 1997).

⁹ <http://www.edisonnj.org/menlopark/taemenlo.asp>

¹⁰ <http://www.eglin.af.mil/TS/climlab/main.html>

Draft Deliberative Document - For Discussion Purposes Only
Not Releasable Under FOIA

or down, by what they think will happen. It is unclear how we will be able to defend a quantitative value based on such speculative judgments. Moreover, these judgments will be subject to the following significant limitations.

- First, over time, “the threat” shapes the force structure. Sometimes the threat is predictable, and sometimes it is not. For example, the DoD’s concepts for future force structure after September 11 are different than they were before that date.
- Second, S&T’s impact on the force structure 20 years hence is unknowable, especially given that basic research is unpredictable and often produces unexpected benefits. Moreover, many of the most revolutionary technologies born in S&T, like radar and GPS, will take as many as 20 years to reach operational use.
- Third, the impact of current D&A is less speculative than for S&T, but it is guess work nonetheless. For example, if the FSA metric was used during the first BRAC round in 1988, the Navy’s experts would have said that the DoN’s 1998 force structure (i.e., only 10 years later, not 20) would have had more than 850 A-12 Avengers streaming from the Fleet’s carriers.¹¹ Things happen.

4. As before, the ideas and suggestions offered in both this memo and in the attached analysis are my own, and therefore do not necessarily represent the official views or positions of the U.S. Navy, RADM Jay Cohen, or Mr. George Ryan. I offer them again because I want to see our process succeed, and believe they provide effective ways to avoid the pitfalls of BRAC-95.

Very Respectfully,



Don J. DeYoung

Capabilities Integration Team, Alternate
Technical Joint Cross Service Group
U.S. Navy

¹¹ <http://www.fas.org/man/dod-101/sys/ac/a-12.htm>

SHADOWS ON THE WALL

THE PROBLEM WITH MILITARY VALUE METRICS

Behold! human beings living in a underground den...here they have been from their childhood, and have their legs and necks chained so that they cannot move, and can only see before them, being prevented by the chains from turning round their heads...

Like ourselves, I replied; and they see only their own shadows, or the shadows of one another, which the fire throws on the opposite wall of the cave?

True, he said; *how could they see anything but the shadows if they were never allowed to move their heads* (emphasis added)?

— Plato, “The Allegory of the Cave”, from Book VII, *The Republic*

Background

The use of Plato’s famous allegory is not meant to suggest that we are held prisoner by the base closure process, though it may feel that way for those of us who have done this before. Instead, the allegory is intended to help articulate this paper’s thesis — which is that, *without changes to our approach, we will fail to accurately and fairly measure the military value of the Department of Defense (DoD) laboratories and warfare / product centers.*¹

Like the prisoners in Plato’s cave who could not see objects in their real form, we are making judgments about the labs and centers indirectly, by way of “shadows” cast by very problematic metrics. The basic difficulties fall under three general areas. The first is that our metrics are unable to measure the actual *impact*, or effect, of the technical work itself on national security. A second is the study design itself, which does not recognize and protect the innovative connectivity within, and between, the three functional areas: science and technology (S&T), development and acquisition (D&A), and test and evaluation (T&E). Both problems were experienced during BRAC-95. Neither was solved.²

The third flaw is all but certain to be a fatal one. The following analysis show how we have built a “Dollars and No Sense” approach where (1) dollars have achieved relative “rock-star status” as our preferred metric for military value, and (2) the Intellectual Capital metrics have been “dumbed-down” to the point where they will generate irrational results like ranking “world-class” talent lower than mediocre talent. This latter issue remains a problem even though the BRAC Infrastructure Steering Group fortunately overturned our highly flawed “Percentage Approach” to scoring the data. If this situation is not fixed, BRAC-05 will damage the DoD’s ability to develop new warfighting technologies. In other words, actions based on our flawed process will weaken national security.

The bad news is that we will never attain a perfect way to determine the military value of research and development (R&D) institutions. The very nature of R&D does not lend itself to easy quantification. But, the good news is that with some changes we can (using Plato’s allegory) widen our constrained field

¹ The military value of test ranges is more easily measured because of its very strong correlation to physical parameters, such as airspace, seospace, and isolation from population centers. Constraints on the parameters (e.g., explosive test limits, inadequate airspace, urban encroachment, etc) are the critical metrics and easily quantifiable.

² The author was a member of the BRAC-95 Navy Base Structure Analysis Team and the DoD T&E Joint Cross-Service Working Group.

of view and sharpen distorted images, without sacrificing the objectivity and certifiability demanded by the BRAC process.

Two Challenges

To achieve the transformation and capacity reduction goals of BRAC-05, while preserving an infrastructure that will continue to create new warfighting capabilities, the Technical Joint Cross Service Group (TJCSG) must meet two challenges.

- (1) Accurately and fairly assess the military value of the Services' corporate laboratories, warfare/product centers, and test ranges
- (2) Provide a credible way to judge their potential to create new warfighting capabilities that will help defeat currently unforeseen threats over the next 20 years

The First Step

To determine military value, we need to first know what it is. The BRAC Selection Criteria are only minimally helpful because they do not explicitly address RDT&E (although they speak directly to other facets of national defense, like joint warfighting, training, and readiness). During the public comment period on these criteria, at least one letter by a Congressional delegation was written to the DoD about this obvious omission,³ but the criteria unfortunately remained as originally written.

Instead of using the deficient BRAC criteria, a clear working definition for the military value of labs and centers is suggested below:

Military Value: the level of effectiveness in meeting national security interests, both in terms of warfighting impact and contributions to science and technology.

The next, and much more difficult step, is to develop metrics that accurately measure each sites' effectiveness in meeting national security interests. Given the inadequacy of the BRAC criteria, these metrics will be the determinant of military value.

Necessary — But Not Sufficient

One fact of life drives us to collect a broad range of disparate *input* metrics (e.g., experience, education level, dollars, etc.) and *output* metrics (e.g., technology transitions, patents, etc.), and then to segment the analysis into three functional areas (i.e., S&T, D&A, T&E). *This driver, experienced in all previous BRAC rounds, is the fact that corporate laboratories, warfare/product centers, and test ranges have very different missions that require diverse, and often unique, workforce skill mixes and types of facilities.*

Our menu of metrics for labs and centers contain the same as those used by the BRAC-95 Laboratories Joint Cross-Service Group (LJCSG),⁴ plus a number of new ones to address acquisition functions, synergy, and jointness. But, like BRAC-95, none measure the *impact*, or real effect, of the work on national security. Such metrics have been the BRAC equivalent of the quest for the Holy Grail.

³ U.S. House of Representatives, letter dated 30 January 2004 (signed by 16 U.S. Representatives).

⁴ A Test & Evaluation Joint Cross-Service Group developed T&E metrics separately.

The TJCSG's input and output metrics are necessary, but not sufficient. If they showed where the best and most relevant R&D was being conducted today, and where the best and most relevant work were likeliest to occur tomorrow, then our task would be simplified. Nearly all of our metrics lack the power to distinguish critical differences, and a number of them are dubious. For example, the "Workload Focus Metric" scores highest in military value those sites with the largest workloads in a technical capability area (e.g., C4ISR). While this metric does show the degree of focus and scale of the work, it does not address quality or relevance. A more apt name might be the "Jiffy Lube Metric." After all, Jiffy Lube focuses on one area, but you may or may not want to bring your car there.

All but one (i.e., Elite & Prestigious Awards) of the metrics fails to directly gauge impact and quality. But unlike the prisoners of Plato's cave who were forced to surmise the real world by its shadowed images, we can widen our narrow field of view with 13 changes.

Proposal #1: Go Back to the Future

Dr. Robert Frosch (former NASA Administrator, Deputy Director of ARPA, Assistant Secretary of the Navy for Research and Development, Associate Director for Applied Oceanography at the Woods Hole Oceanographic Institution; and vice president of GM Research Laboratories) once observed that in R&D,

“... you cannot measure the future; the only thing you can measure is past performance. You have to measure R&D by what you have done”⁵

In other words, the only viable metric for evaluating a site's effectiveness in meeting national security requirements is its *track record*. After all, we routinely judge sports teams that way. What might the reaction be if we were tasked to determine the NFL's top teams, and we responded by collecting data on stadium square footage, revenue expended, number of luxury box seats, proximity to other sports complexes, number of first round draft picks, tackles made/missed, or whether the stadium had a dome? Why treat national security less seriously than athletic competition?

We can assess past performance by collecting information that directly ties the sites' work to national security interests. The following is an example of how this can be done.

- Identify and describe up to 15 technical contributions that demonstrate your site's impact on developing major new warfighting capabilities that were introduced into routine operations *since the end of the Cold War*, new technical developments during that period that may or may not be yet deployed, and new scientific breakthroughs which occurred during that period. (2,000 words max)
 - Specify exact role played by *in-house* personnel in the development of that capability
 - Cite any other DoD, Federal, or private sector organizations that also played a role and attempt to quantify those roles
 - Provide copies of primary references to substantiate all claims. References must be from authoritative sources external to the laboratory/center.

This is one way to capture critical information relating directly to military value. For example, this approach would gather information on: state-of-the-art night-vision technologies developed by the Army's Night Vision Laboratory and tank ammunition developed by the Army Research Laboratory, both used in Desert Storm with devastating results; the thermobaric bomb developed by the Navy's Indian Head laboratory and used in Afghanistan to avoid bloody tunnel warfare; and the F/A-18 SHARP

⁵ R. Frosch, "The Customer for R&D is Always Wrong!," *Research•Technology Management*, (November-December 1996), p. 27.

reconnaissance system developed by the Naval Research Laboratory, which provided real-time digital imagery (vice the 3-9 day norm) and was credited with saving lives in Operation Iraqi Freedom.

Some may claim this approach is not objective or certifiable, or that it somehow violates the integrity of the BRAC process, because the information would come from the sites under study. This argument is not convincing. Most of the data we are gathering comes, in some form, from the sites themselves. The real issue is whether the requested information can be verified with official documentation and corroborated with sources outside the reporting site.

A preliminary IG review, performed after our 2-3 December off-site at the Bolger Conference Center, determined that the use of authoritative documentation would make this approach auditible, and therefore defensible.

Proposal #2: Use Expert Judgment

Again, using Plato's allegory, if we cannot look directly at the work of the labs/centers, then it makes sense to listen to those defense experts who have seen and evaluated it themselves. Since World War II, the DoD labs and centers have been the subject of more than 100 studies.

One of the more recent was a Congressionally mandated study to evaluate the projects of the labs/centers for relevance to future warfighting requirements. The defense experts on the "Sec. 913 Lab Relevance Study", completed in 2003, included retired "4-stars" General A. Zinni, Admiral H. Gehman (chair of the *Columbia* Accident Investigation Board), General W. Crouch, and General M. Carns, as well as a former Air Force Secretary and DDR&E, and various other technical experts. The findings of this expert team, along with others who have looked directly at the work, can help us measure the impact of labs and centers in meeting national security interests.

We can use DoD's extensive library of authoritative information to help assess actual performance. The following proposed data call question to the sites is an example of how this can be done.

- Quote up to 15 findings or comments made by DoD / Service chartered review panels, or made independently by national leaders in defense R&D, about your site since the last BRAC round in 1995. Provide references to substantiate all claims. (1,000 words max)

The sites' responses might take the form of the following sample excerpts from the Sec. 913 Relevance study, sent to Congress in early 2004:

- Regarding the Army Research Laboratory (ARL) at Aberdeen, the Armaments Research, Development, and Engineering Center (RDEC), and the Aviation and Missile RDEC, the report stated:

"The Team was impressed with the relevance of the ARL and RDEC program, the high potential value it has to the warfighter, the strong links to the requirements of the warfighter, and the positive approach to jointness." One of the experts stated, "The Army ballistics program is one of the areas wherein the Army holds a dominant lead technically and the private sector carries out the Army's concepts. There is no concern over privatizing this work."⁶
- Regarding the Naval Air Warfare Center (NAWC) at China Lake, one of the Team's experts made the following statement:

⁶ Section 913 Report #3: *Weapons Science and Technology and the Department of Defense Laboratories*, (National Defense University: December 2002), p.18.

“I believe their location is really important. Their proximity to other ranges is very valuable. They are close to the aerospace industry in this country as well as some quality academic organizations. Their distance from population centers allows for the testing of dangerous systems in a safe way. All in all it is an extremely valuable resource for the DoD and should be nourished and protected.”⁷

- Regarding the Communications—Electronics Command (CECOM) RDEC at Fort Monmouth, the report stated:

“The panel was impressed with the relevance of the CECOM RDEC’s program...the work being done is competent and relevant to the missions of the Army...it was noteworthy that some important work has already been done in support of homeland security, and a good number of projects have potential applications in that area. There may be value in supporting and encouraging more contacts with non-traditional customers such as first responders, the National Guard, and border agencies, to make these technologies relevant to their missions.”⁸

- Regarding the Air Force Research Laboratory (AFRL) at Rome, New York, the report stated:

“The work being done at AFRL Rome is very relevant. The Laboratory is appropriately focused on the information and knowledge dominance components of information technology, and the work presented to the Study Team supports full spectrum dominance as defined in *Joint Vision 2020*...AFRL Rome is clearly tied in with various Air Forces exercises...One example was a team that was updating the chemical response monitoring capabilities at Osan air base. This was not transformational work, but it did vastly improve the CINC’s capabilities in this area.”⁹

- Regarding the Naval Research Laboratory (NRL), the report stated:

“With significant funding from the other Services and DoD agencies, the program clearly addresses interoperability issues and joint warfighting requirements...Many of the projects represent new and innovative approaches to serious national and service problems...NRL moves new technology into field use and has, according to one Study Team expert, ‘probably the best record of any DoD organization for transitioning products’...one of the Study Team ‘four-stars’ made the following assessment, ‘What we saw was a Category A+ laboratory.’”¹⁰

- Regarding the CECOM Night Vision & Electronic Sensors Directorate (NVESD), the report stated:

“The Laboratory’s ‘Own the Night’ mission statement guides it in the conduct of a highly relevant S&T program. NVESD has done an outstanding job developing sensors that will enhance the ability of ground UAV and rotary-wing assets to detect and engage time-critical targets. In this area they are undoubtedly the world’s best...its efforts in uncooled detectors must be viewed as a high-risk undertaking that, if successful, will have monumental implications for all U.S. IR sensors and IR-sensor-guided weapons.”¹¹

Some may claim this approach compromises the BRAC’s objectivity because such studies can be biased (a legitimate concern), or that it is not certifiable because it draws from information outside the closure process. These arguments are not convincing. First, we would ask the sites to self-report the comments,

⁷ Ibid., p.12.

⁸ Section 913 Report #2: *Information Science and Technology and the Department of Defense Laboratories*, (National Defense University: July 2002), p.14 and 19.

⁹ Ibid., p. 20, 22.

¹⁰ Section 913 Report #1: *Sensors Science and Technology and the Department of Defense Laboratories*, (National Defense University: March 2002), p.26, 30-31.

¹¹ Ibid., p. 14.

which will presumably be positive ones, much like those in the above sample excerpts. Self-reporting eliminates the concern about studies that might convey undeserved criticisms.

Second, this data will come from official reports, authorized and approved by the DoD / Services, and from documented statements by national leaders in defense R&D. *If this information cannot be considered authoritative and certifiable, then why does the Defense Department continue to charter these studies — at considerable public expense — and provide them to Congress?*

Third, BRAC-05 will — for the first time in five rounds — entertain “transformational options” *proposed by private groups outside the Government*, such as the Business Executives for National Security. These options will be used during the phase where the DoD generates and evaluates proposed closures. *Surely, if private sector opinions can be used for such a sensitive phase of the process, then the official findings of DoD chartered and approved studies, and the independent conclusions of national leaders in defense R&D, must be acceptable to use when determining each site’s military value.*

Finally, as noted previously, the DoD IG determined that the use of such authoritative DoD studies would be auditable, and therefore defensible. It would be unfortunate not to capitalize on the wealth of information and expert judgments available in such studies.

Proposal #3: To Transform and Enhance Jointness — *Stay Connected*

New technology is one of the two primary engines of military transformation, with the other being tactics. So it follows that if we can ensure, through our BRAC process, an innovative and agile R&D infrastructure for the future, we will also ensure that the DoD will gain important new technologies. *The point here is that we don’t need to pick the technologies ahead of time* (which we are trying to do with the dubious Future Warfighting Capability metric), *but we do need to pick the most innovative and agile sites.* The two previous proposals can help us do that, but more can be done.

To quote Dr. Frosch again, “Great R&D must preserve the *connections* (emphasis added) between various kinds of knowledge.” He continues,

“...The problem of R&D management is, in a sense, the problem of the management of a variety of forms of knowledge that are deeply interconnected, and whose interconnectedness one learns as one tries to solve the problem. It turns out to be a problem of *maximizing collision cross-sections among kinds of knowledge* (emphasis added): making sure people who need knowledge they don’t have—and may not even know they need—have a good chance of learning about it.”¹²

A 1973 Battelle study on the elements of the innovative process makes a similar point.

“Confluence of technology, unplanned in most instances, was important in all case histories, and to a substantial number of decisive events.”¹³

These comments are relevant to our task because they reveal the importance of *synergy*: i.e., the degree of lab / center “interconnectedness” within its own organization, with each other, the private sector, and with the national and international scientific and technological community. This synergy finds expression in three different ways: *multidisciplinary programs, vertical integration, and jointness*.

¹² Ibid. p. 23-24.

¹³ Battelle, “Interaction of Science and Technology in the Innovative Process,” (1973).

A multidisciplinary program of collocated scientific disciplines and technology areas enhances the horizontal interconnectedness that Battelle found important to innovative success. And a multidisciplinary program, in Frosch's words, "maximizes collision cross-sections among kinds of knowledge." A vertically integrated program, on the other hand, optimizes connectivity across technical functions (S&T, D&A, and T&E). Finally, jointness happens when these horizontal and vertical connections jump Service boundaries.

The logic is as follows: (1) the key to transformation is identifying innovative and agile sites; (2) a key to recognizing those sites (in addition to the two previous proposals) is to measure connectivity, or synergy; and (3) we can measure synergy by gauging a site's level of multidisciplinary projects, vertical integration, and jointness. The following military value questions can help measure that synergy.

- How many patent citations of published refereed journal articles have there been over the last 5 years? (This shows the connectivity of real applications to research published in journal articles.)
- How many on-site projects (include project name and dollar level for each) have been funded by other Services since 1995? (This cross-service connectivity shows an ability to meet joint warfighting needs and indicates that the work is of superior quality for another Service to fund it.)
- How many refereed journal articles were written in collaboration with personnel in other scientific disciplines and technology areas, both inside and outside the lab / center over the last 5 years? List separately with name of collaborating organizations. (This shows the level of horizontal and vertical connectivity. This question can also include collaborative patents granted)
- What percentage of the lab / center budget is discretionary and allows the Director to pursue high-risk, high-payoff projects? (This measures the agility of a lab / center to pursue connections and "maximize collision cross-sections", without rigid conformance to higher level Service management.)

The above questions pertain largely to S&T, which makes sense given its seminal role in creating tomorrow's disruptive technologies. Data on shared patents can help gauge the connectivity of D&A.

Proposal #4: To Optimize Test Range Operations — *Stay Connected*

Connectivity, or synergy, is also important to T&E. It can be quantified by asking the test ranges to report total square miles of restricted airspace, seospace, and land area of other contiguous sites that it uses to conduct tests. It is often said that the military fights the way it trains, which is why training is conducted in the most realistic ways possible. *Similarly, our analysis of the ranges should capture they way they actually test.*

For example, if Edwards AFB uses the airspace at NAWC China Lake, as well as the warning areas off NAWC Point Mugu for its tests, then it should report the total square miles that the *composite* test area represents. If it also uses the Utah Test and Training Range, the airspace above Nellis AFB, and elsewhere, then those totals should be included in the composite.

Other examples would include NAWC China Lake's similar composite use of available airspace and warning areas. Eglin AFB, White Sands Missile Range, and NAWC Patuxent River would likewise report whatever contiguous air, land, and sea space they actually use to test. *This information would capture the connectivity that exists between the DoD's test ranges during actual testing operations.*

Air, land, and sea space that is owned or controlled by each installation would be reported and scored separately. Only restricted airspace and warning areas shown on U.S. Government civil aeronautical

charts should be scored in order to eliminate the ambiguities that plagued BRAC-95's T&E analytical process and compromised the credibility of its results.¹⁴

Proposal #5: Avoid the Misguided Notion that Intellectual Capital is Fungible

Scientists and engineers were treated by BRAC-95 as interchangeable, conveyable, replicable items — much like military housing, piers, and hangar space — regardless of their competence and professional accomplishment.¹⁵ Such simplistic treatment can be harmful to national security because top technical talent is critical to the success of defense RDT&E programs. While it is true that, with sufficient time and money, some personnel may not be difficult to replace, we all know the blunt truth is that the best will not move with the work. If we repeat BRAC-95's approach, we risk “decapitating” the intellectual drivers from the realigned workload.

Rather than repeat that flawed approach, our central focus should be on identifying and preserving high quality intellectual capital. Toward that end, the Intellectual Capital or “People Metrics” should receive weightings on the upper end of the ranges discussed thus far. However, we need to ensure that the quality of the personnel data we collect warrants the high weighting. *It does no good to give the People Metrics high weightings if they cannot identify the top talent.* The next proposal will address that concern.

Proposal #6: Fix the “Dumbed-Down” People Metrics

The accurate evaluation of intellectual capital — its quality, and in some cases, actual impact — is essential because, in the words of a former DDR&E,

“The presence of *a few* (emphasis added) individuals of exceptional talent has been responsible for the success (and even the existence) of outstanding research and technology development organizations.”¹⁶

Unfortunately, as currently written, our People Metrics (i.e., Elite & Prestigious Awards, Experience, Education) will, in all likelihood, fail to discriminate the essential differences among the labs / centers due to: (1) the failure to assign value to some of the DoD’s best and brightest, (2) the highly compressed range of point values, and until recently, (3) the “Percentage Approach” to scoring the data, which was fortunately rejected by guidance issued by the Infrastructure Steering Group (ISG) on 4 June 2004 (see Attachment A). The result of these flaws will be the loss of meaningful information in the noise of large aggregate populations.

We can avoid this problem by *focusing on a very limited set of critical data and scoring it appropriately*. In the excerpt below, Peter Drucker describes how effective decisions are derived from studying the critical items that drive a given process.¹⁷ The excerpt is lengthy, but best read in its entirety.

“That the procurement and inventory policies of the U.S. armed services were in bad shape had been known ever since the Korean War. There had been countless studies – but things got worse, rather than better. When Robert McNamara was appointed Secretary of Defense, however, he challenged the

¹⁴ The suggestions of proposal #4 are not the “no-brainers” they appear to be. During BRAC-95, the T&E JCSG awarded one site military value points for airspace 550 miles away that it did not control, while another was denied points for airspace 150 miles away that it did control. The scoring rules must be clear and equitable.

¹⁵ D.J. DeYoung, “The Silence of the Labs,” *Defense Horizons*, No. 21 (January 2003).

¹⁶ Hans Mark and Arnold Levine, *The Management of Research Institutions* (Washington, DC: Scientific and Technical Information Branch, National Aeronautics and Space Administration, 1984).

¹⁷ The OSD BRAC office director, Mr. P. Potochny, made the same point in our 23 January TJCSG meeting by suggesting that we capture only the most important data.

traditional measurements of military inventory — measurements in total dollars and in total number of items in procurement and inventory. Instead, Mr. McNamara identified and separated the very few items — maybe 4 percent of the items by number — which altogether account for 90 percent or more of the total procurement dollars. He similarly identified the very few items — perhaps again 4 per cent — which account for 90 per cent of combat readiness. Since some items belong in both categories, the list of crucial items came to 5 or 6 per cent of the total, whether measured by number or by dollars. Each of these, McNamara insisted, had to be managed separately and with attention to minute detail. The rest, the 95 per cent or so of all items which account neither for the bulk of the dollars nor for essential combat readiness, he changed to management by exception, that is, to management by probability and averages.”¹⁸

The above statements by both Drucker and the former DDR&E show the value of identifying *the few*. Therefore, *a way for us to recognize the top performing sites is to identify only the exceptional talent*. In our first step toward this goal, we originally developed an excellent framework to identify this talent with a three-tiered filter that included:

Elite Awards (e.g., the Nobel Prize, Robert J. Collier Trophy, National Medal of Science, National Medal of Technology, Draper Prize, Bower Award and Prize for Achievement in Science, and members of the National Academy of Sciences and the National Academy of Engineering)

Prestigious Awards (e.g., A.T. Waterman Award, Stellar Award, Goddard Astronautics Award, William Streifer Award, Lord Rank Award, National Inventors Hall of Fame, Space Technology Hall of Fame, and Technical Society Fellows, such as IEEE Fellows, AAAS Fellows, etc.)

Patents, Citations in Refereed Journals, and Software Licenses

Unfortunately, some subsequent decisions jeopardize our success. They are as follows:

(1) Refusal to include Science and Technology (ST) positions in the analysis. This decision will be impossible to explain to the Commission. The TJCSG’s subgroups had the patience of Job when addressing the physical features of the Department’s facilities, but steadfastly resisted including ST positions in the analysis. Deciding to ignore ST’s means that zero value will be assigned to this competitively selected talent, or to the sites where they do their innovative work.

The reason given for ignoring the ST was that “it is just a grade.” On a minimal level, that is a true statement — much like saying the McKinley Climatic Chamber is “just a building,” and China Lake’s bombing range is “just a high desert plain.” ST’s are some of the best and brightest technical personnel in the DoD. This is a well-supported fact. For example, the Army Science and Technology Master Plan (1997) refers to the ST as a “world-class” scientist. And a former Deputy Director for Defense Research and Engineering stated,

“Science and Technology (ST) positions recognize a *world-class* expert level for researchers with pay comparable to equivalent managers, thus giving laboratories more tools to retain their *best researchers* without losing them to the private sector...The occupant of a ST position serves as the nucleus for ‘growing’ the capability of the organization, attracting skilled personnel, and *setting the technical standard* for the laboratory.”¹⁹

STs should be captured and scored under the Experience metric. As currently structured, this metric contributes little to our analysis. We are assigning 1-point to all those with 10 or less

¹⁸ Peter F. Drucker, *The Effective Executive* (New York: Harper & Row, 1966), p.145.

¹⁹ Office of the Director of Defense Research and Engineering, *Memorandum For Director, Defense Performance Review Task Force*, (30 July 1993).

years of experience, 2-points for all those between 10 and 20 years, and 3-points for all those with more than 20 years. *At best*, this metric grants higher military value to those workforces that have spent the most years doing work with significant military value. *At worst*, this metric grants the highest scores to workforces with the greatest seniority, and punishes those that have refreshed themselves with younger talent. Moreover, since many people begin new areas of endeavor at various points in their career, there will be no assurance that the reported level of experience is in fact relevant to the technical work at issue. Most likely, the metric will be doing all of the above — both the positive and the negative — to some unknown degree.

Collecting the STs and scoring them at 70 points (see T&E ratio in Section (4) below) would infuse this moribund metric with an actual confirmed measure of quality. Pegging the STs' value at 70 times that of an unknown person with less than 10 years of experience working at an unknown level of technical competence should be defensible because an ST is a “world-class” scientist who is of the quality that this metric purports to be awarding. STs also attract skilled personnel, so the 70-point value would also help counteract some of the unintentional penalty meted out to laboratories and centers that have refreshed their workforces with younger talent.

(2) *The Downgrade of Society Fellows.* Our original assignment of Society Fellows to the Prestigious tier for the awards / honors metric was changed by removing them and cutting their value by half. It is unclear why, or exactly when, that occurred. This decision further dampens our ability to differentiate high quality talent and separate it from the aggregate population. Why would we want to take this group of technically gifted government scientists and engineers, who have been recognized as such by the national and international technical community, and pull their value down toward that of the aggregate population? How do we defend the fact that our plan will now equate a Society Fellow to 5 unidentified patents of no known value?

It should be noted that each new class of IEEE Fellows, for example, is a highly select bunch that cannot number more than 0.1 percent of the total membership. It should also be noted that a non-scientific survey that I performed on the internet shows some IEEE fellows with between 100 and 200 patents to their name. In short, these are extraordinarily talented individuals, with an extensive record of accomplishment that has a confirmed, known, and substantial value.

In short, the decision to downgrade their value is inappropriate and compounds the already significant problem of point compression in our analytical plan.

(3) *Compressed Point Values.* The TJCSG sub-groups balked at assigning high point values to the *elite* and *prestigious* awards and honors, saying that it will unfairly “skew results.” Part of this resistance stems from a perceived lack of quantifiable parameters as compared to other metrics, such as those under the Physical Environment attribute. However, the Physical Environment metrics show that large differences, if quantifiable, are acceptable.

For example, the Army’s White Sands Missile Range (WSMR), the largest land test range in the Western Hemisphere, has nearly 4 million acres available for tests.²⁰ On the other end, the Aberdeen Proving Ground’s ATC test range has 56,707 acres,²¹ less than 1.5 percent of WSMR’s. None would think point values based on this 70:1 ratio would unfairly skew results. Therefore, some statistics might help justify an equivalent range of point values for the Awards Metric.

²⁰ <http://www.globalsecurity.org/space/facility/wsmr.htm>

²¹ Jeanne Ditter, Conference Briefing: “Test & Evaluation for Scalable Effects Capabilities: Aberdeen Test Center as the SEC Test Center of Choice,” (4 June 2002).

For the Elite Awards, consider that there have been only 143 Nobel Laureates in Chemistry over the last 100 years, and the DoD has one of them.²² Consider also that there are a total of 12,530,700 S&Es in the U.S.²³ but less than 1,900 members of the National Academy of Sciences (i.e., less than .01% of all American S&Es), and the DoD has a number of them.²⁴

Regarding the Prestigious Awards, consider that ARL and one of its scientists, were inducted into the Space Technology Hall of Fame for work on Quantum Well Infrared Photo-detectors, an area of huge importance to military sensor systems. There have been 44 technologies inducted into the Hall of Fame over the last 15 years, about *3 per year nation-wide*. The Rank Award is given for exceptional achievement in electro-optics. One recipient, no longer with the DoD, was recognized for discoveries leading to development of the rare gas halide excimer laser. Aside from its DoD applications, this is the laser that made Lasik surgery possible. The Rank award has been granted to 123 persons over the last 28 years, about *4 per year world-wide*.

These statistics provide a solid quantitative rationale for assigning 70 points to the Elite Awards, 10 to the Prestigious Awards, and 1 point to each patent and software license. This range also corresponds to the 70:1 WSMR / ATC ratio. This scoring scheme is reasonable and would be defensible to the Commission.

Unfortunately, the Subgroups instead chose highly compressed point values that will decrease our ability to differentiate the exceptional talent from the rest. As currently written, our plan is to assign 30 points to the Elite, 10 points to the Prestigious, 5 points to Society Fellowships (which were for some reason broken out separately from the Prestigious awards), and 1 point to each patent and software license. What this means is that 5 unidentified patents of unknown military value will equate to one IEEE Fellow. The effects of this decision grow more absurd as it affects the Elite Awards.

For example, DoD's Nobel Laureate did work that now results in the molecular structure determination of more than 10,000 new substances a year, and he continues his work with applications ranging from the characterization of potent toxins all the way to making explosives and propellants that are safer and more powerful. But, by our study plan, *the military value of this Nobel Laureate is equal to 30 unidentified software licenses of unknown military value* — despite the profound impact of his work on warfighting capabilities, his ability to attract a staff of exceptional talent, and the pervasive and continuing value of his accomplishment to the DoD.

The point values for the other People Metrics (i.e., Education and Experience) are even more compressed. For example, the Education metric assigns 3 points to a PhD, 2 points to a MS, and 1 point to a BS. A wider spread, such as 10:3:1 would increase our ability to distinguish critical differences among the many sites under study. The next section shows how the Percentage Approach to scoring for military value exacerbates the problem with point compression.

(4) The Flawed “Percentage Approach” to Scoring.

[Note: This highly flawed numerical approach was the TJCSG's official methodology for nearly 5 months, but it was fortunately overturned by the ISG on 4 June. The following discussion is

²² http://www.nobel.se/help/faq/nobel_laureates.html#5

²³ National Science Foundation, *Science and Engineering Indicators*, (2000).

²⁴ <http://www4.nationalacademies.org/nas/nashome.nsf/>

Draft Deliberative Document – For Discussion Purposes Only
Do Not Release Under FOIA

retained in this paper's fourth version to demonstrate the level of energy devoted to dumbing-down the People Metrics. The following discussion appears as it did in the earlier versions.]

The Navy's objection to the percentage approach has been voiced in a number of different forums, including two occasions to the DDR&E (see Attachments B, and C). Each time it did so, data was presented to substantiate its concern that this approach will likely yield some number of irrational results. Proponents of the percentage approach were invited repeatedly to show instances where the Navy's recommended "absolute numbers" approach would yield irrational results, but each time the response was silence.

The Navy's concerns go beyond the irrational results. The Percentage Approach depresses the value for large masses of high quality talent (whether measured by the awards, education, or experience metrics) and exaggerates the value of small masses — across the board. Take the patently absurd case where the Percentage Approach gives a one-person site with one PhD the same military value score as a 1,000-person site with one thousand PhDs. Here it is obvious that the military value of the small mass is exaggerated and the military value of the large mass is minimized. This absurd case can be seen as the outer boundaries of the problem "box", but every case within the box will be affected to varying degrees. *And some number of them will yield irrational results.*

To illustrate the Navy's concerns, the following scenarios compare a small 100-person site versus a large 500-person site.

Absolute Approach:

<i>Site A:</i>	100 S&Es (50 PhDs, 50 MS)	
	$3(50) + 2(50) = 250$	MV = 0.12

<i>Site B:</i>	500 S&Es (300 PhDs, 100 MS, 50 BS, 20 AD, 30 No degree)	
	$3(300) + 2(100) + 1(50) + .5(20) + 0(30) = 1160$	MV = 1.0

Percentage Approach:

<i>Site A:</i>	100 S&Es (50 PhDs, 50 MS)	
	$3(50) + 2(50) / 100 = 2.5$	MV = 1.0
<i>Site B:</i>	500 S&Es (300 PhDs, 100 MS, 50 BS, 20 AD, 30 No degree)	
	$3(300) + 2(100) + 1(50) + .5(20) + 0(30) / 500 = 2.32$	MV = 0.92

Therefore, *despite having 6 times the number of PhDs and 2 times the number of MS, Site B's military value is almost 10% less than Site A.* This is irrational. Site A lacks the greater intellectual horsepower and sheer idea generation that comes with 250 more PhDs and 50 more MS, as well as the talent that 100 less degreed, and generally younger, individuals have to offer.

Draft Deliberative Document – For Discussion Purposes Only
Do Not Release Under FOIA

Proponents of the Percentage Approach counter that the Absolute Approach gives Site B a MV that is almost ten times Site A, which they found excessive. However, this is more the function of the compressed point range. For example, if a PhD was worth 10 points, an MS worth 3 points, and a BS worth 1 point, the MV for Site A increases to 0.19. But when such a change (specifically a 30:10:1 ratio) to the point range was suggested as a way to address their concerns and increase the MV of the hypothetical small site, the proponents for the Percentage Approach were not interested (see Attachment D).

The problem grows more acute when scoring for the Elite and Prestigious awards:

Absolute Approach:

Site A: **100** S&Es
No Elite awards
No Prestigious awards
5 IEEE Fellows
30 unidentified patents of unknown value in last 3 years

$$5(5) + 1(30) = \mathbf{55} \qquad \qquad \qquad \mathbf{MV = 0.21}$$

Site B: **500** S&Es
1 National Medal of Technology, **1** Nobel Prize, **1** Collier Trophy
1 Space Technology Hall of Famer
10 IEEE Fellows, **10** AAAS Fellows
60 unidentified patents of unknown value in last 3 years

$$30(1) + 30(1) + 30(1) + 10(1) + 5(10) + 5(10) + 1(60) = \mathbf{260} \qquad \qquad \qquad \mathbf{MV = 1.0}$$

Percentage Approach:

Site A: **100** S&Es
No Elite awards
No Prestigious awards
5 IEEE Fellows
30 unidentified patents of unknown value in last 3 years

$$5(5) + 1(30) / 100 = \mathbf{0.55} \qquad \qquad \qquad \mathbf{MV = 1.0}$$

Site B: **500** S&Es
1 National Medal of Technology, **1** Nobel Prize, **1** Collier Trophy
1 Space Technology Hall of Famer
10 IEEE Fellows, **10** AAAS Fellows
60 unidentified patents of unknown value in last 3 years

$$30(1) + 30(1) + 30(1) + 10(1) + 5(10) + 5(10) + 1(60) / 500 = \mathbf{0.52} \qquad \qquad \qquad \mathbf{MV = 0.94}$$

These results are irrational. *A site with only 5 fellows and 30 unidentified patents of unknown value scores 6% higher in military value than a site with a Nobel Prize, a Collier Trophy, a National Medal of Technology, 20 fellows, a Space Technology Hall of Famer, and twice as many unidentified patents of unknown value?* Bear in mind that Bill Gates, Steven Jobs, Admiral Grace Hopper, Norm Augustine, and David Packard all won the National Medal of Technology. The Nobel's value is self-explanatory. And the Navy and Air Force shared the Collier Trophy for

inventing and developing the Global Positioning System. This would not pass the smell test with either the Commission or the communities.

Also, note how Site A, even under the Absolute Approach, scores more than 1/5 of the MV compared to Site B, which is proportionally five times larger, but *far* more technically accomplished. *This is due to the point compression problem.* A 70:1 Awards ratio, which parallels the test range ratio of WSMR / APG, and the restoration of the Society Fellows to the Prestigious tier, whose demotion in value only compounded the compression problem, gives us a more credible comparison below.

Absolute Approach + 70:1 Point Ratio + Restoration of Fellows to Prestigious Tier:

<i>Site A:</i>	100 S&Es No Elite awards 5 IEEE Fellows 30 unidentified patents of unknown value in last 3 years	
	$10(5) + 1(30) = \mathbf{80}$	MV = 0.16
<i>Site B:</i>	500 S&Es 1 National Medal of Technology, 1 Nobel Prize, 1 Collier Trophy 1 Space Technology Hall of Famer 10 IEEE Fellows, 10 AAAS Fellows 60 unidentified patents of unknown value in last 3 years	
	$70(1) + 70(1) + 70(1) + 10(1) + 10(10) + 10(10) + 1(60) = \mathbf{480}$	MV = 1.0

In reality, given that one point is awarded for any unidentified patent of absolutely no known value, a more accurate value for the Elite and Prestigious Awards would be on the order of 500 and 50 points respectively. Who would argue that an Elite award like the Collier Trophy — awarded annually for the greatest achievement in aeronautics or astronautics in America, to past winners like Glenn Curtiss, Orville Wright, Chuck Yeager, the “Mercury 7”, the Apollo 11 crew, NASA’s Voyager Team, and the Navy and Air Force GPS team — is not worth 500 unknown patents of no known value? *This 500:1 ratio eliminates the point compression problem and gives Site A the more realistic MV of 0.1 when compared to the internationally recognized heavyweight achievements of Site B.* The only reason this 500:1 ratio is not recommended is the convenient 70:1 fit with the test range ratio, which should make it acceptable to all TJCSG members.

In the end, this paper’s recommended approach (i.e., the “Absolute Numbers” scoring for all people metrics, a 70:10:1 point spread for the Awards Metric, a 10:3:1 point spread for the Education Metric, scoring STs under the Experience metric at a point spread of 70:3:2:1, and restoring the Technical Society Fellows to the Prestigious tier of the Awards/Honors metric) would yield closure and realignment proposals that we can better defend to the Commission. These steps must be taken to gain the necessary differentiation to properly value the Department’s intellectual capital. This is a vital key to identifying the most innovative sites, and it is critical to the credibility of the BRAC process.

On the other hand, if we fail to take this approach, we will have “dumbed-down” the People Metrics to the point where we jeopardize the success of our entire endeavor. Any community-hired BRAC consultant worth his or her salt would zero in on these problems a means of discrediting the study. For one, *the use of the Percentage Approach would put the TJCSG in the position of having to defend irrational results, as well as explain why it insisted on using it when the problems were apparent well*

beforehand and conveyed to the DDR&E. Furthermore, the omission of “world-class” scientists is leading with our chin, and devaluing Society Fellows loses them among the aggregated noise level of patents, software licenses, and paper citations. Lastly, our compressed point values will be vulnerable to dissection and ridicule by anyone with the statistical aptitude of an avid baseball fan and an understanding of the non-linear difference between a PhD and a BS, an ST and an unknown person with less than 10 years of experience, and a Nobel Prize and an unidentified software license.

Proposal #7: Do Not Include Contractors in Scoring for Military Value

[Note: On 4 June, guidance issued by the ISG prohibited the TJCSG’s use of contractor data for military value. However, the following discussion is retained in this fourth version because there is still an issue regarding the inclusion of personnel from Federally Funded Research and Development Centers. The following discussion appears largely as it did in the earlier versions.]

First, BRAC is not about private sector infrastructure; it is about the infrastructure of the Federal Government. If a lab or center is closed, but the DoD remains in need of a given contractor, then those services can continue to be bought by the sites receiving the workload of the closed site. The strength of the private sector is that its services will persist, if there is public money to pay for it. Contractors will follow the money.

Second, BRAC data *must* be auditable. However, it is doubtful that personnel data on *individual* contractors (i.e., education level, years of experience, awards, etc.) would meet the stringent audit standards of BRAC. If individual contractor data were not specified within the contract obtaining the contractor’s services (which is acknowledged to be unlikely), then how would it be collected? Unlike data on government personnel, which is retained in auditable central federal databases, *contractor data would need to come from the companies themselves.*

In Attachment E, the DoD IG finds that contractor data “does not provide the TJCSG with consistently useable data” and that “BRAC 2005 is a DoD process, not a contractor process.” Another interesting fact is that for the Capacity Data Call (which appropriately included on-site contractors to gauge base capacity) the DoD IG reports that, “in some cases, contractors *are requesting to be paid for their data.*” Money raises an issue that needs discussion — i.e., the financial dependence of private contractors on the continued operation of the host laboratory / center increases the risk of fraud. Given the recent corporate ethics meltdown exhibited by scandals at Enron, Worldcom, Arthur Anderson, Tyco, and ImClone, most would find the DoD IG’s judgment to be prudent.

Abiding by the DoD IG’s guidance does three things for us. It avoids inevitable challenges of data that are not sufficiently auditable, prevents the occurrence of an ethics scandal, and avoids scoring private sector infrastructure for a process that will close or realign public sector infrastructure.

Lastly, personnel from Federally Funded Research and Development Centers (FFRDCs) should not be included in the analysis and scored like government personnel, as one Service argues strenuously. Arguments that they are not contractors, or that they are more like federal employees than contractors, are disingenuous. Indeed, if FFRDC staffs were in fact government personnel, they would be called civil servants, have the authority to contract public dollars, make financial decisions in the name of the U.S. government, and be subject to the same DoD outsourcing policies levied on the government workforce. On the other hand, it might be interesting to ask CEOs of the FFRDCs if they are willing to, like government infrastructure, (a) close shop should their host base be closed by BRAC action, and (b) include their workforce and federally-funded budgets under the same DoD outsourcing quotas and

policies currently levied on the federal workforce. I bet we would find out in a New York minute that they consider themselves to be neither civil servants nor public sector infrastructure.

While FFRDC personnel may possess some special status that distinguishes them slightly from a typical contractor, but they are still private sector contractors. For example, The Aerospace Corporation, a private sector entity, manages the Aerospace FFRDC that is mentioned often by FFRDC proponents as a model example. While the FFRDC personnel may work side-by-side with civil servants of the Air Force Space and Missile Systems Center (SMC), there are strict contractual rules that pertain to their management. The following is an excerpt from the SMC FFRDC Users Guide (1 August 2002), p.1-4.

1.3 *The Aerospace Corporation.* The Aerospace Corporation, through an annual incrementally funded research and development Air Force contract, which coincides with the Government's Fiscal Year, operates an FFRDC, which provides scientific and engineering support...

The contract entered into between the sponsoring agency (AF) and The Aerospace Corporation for the operation of the Aerospace FFRDC places requirements on the Aerospace FFRDC for performance of technical work and specifies various terms and conditions under which that work shall be performed. It specifies the direct Staff-year of Technical Effort (STE) labor hours to be delivered, sets forth specific categories of effort (GSE&I, TR, and TS), and lists programs to be supported in each category. The categories of effort and related tasks in each program are specifically defined in the Technical Objectives and Plans (TO&Ps). Approved STE years and/or dollar constraints are contractually authorized by SMC/AXC...

b. *Government Direction to Aerospace Personnel.* Any direction to the Aerospace FFRDC to perform work other than that required by the contract is prohibited and may constitute a violation of the Anti Deficiency Act, 31 USC 1341. Any question concerning the Aerospace FFRDC's responsibility to perform a given task must be immediately referred to the Contracting Officer for resolution. In the event the Aerospace FFRDC performs work other than that required by the SMC contract or any other FFRDC contract, The Aerospace Corporation does so at its own risk.

In short, it is hard to defend the notions that such personnel are not contractors or that they are more like government employees than contractors. The above conditions and constraints under which the Aerospace FFRDC operates are very much contractual in nature.

Proponents for including the FFRDCs have also argued that (a) they are authorized by 10 USC Sec. 2367, (b) the government is authorized to make sole source awards to them, and (c) that they are covered under a different portion of the Federal Acquisition Regulations from other A&AS contractors. Presumably these items are cited to support the fact that FFRDCs are a different breed of contractor because they have a "non-profit" status. I am not an expert on such matters, so I'll instead cite some findings made by a political scientist, H.L. Nieburg, who was an adviser to both President John F. Kennedy and Robert Kennedy.

In Nieburg's book titled, *In the Name of Science*, which focused on the post-World War II growth of the phenomenon known as the Contract State, he examines the role of FFRDCs, and Aerospace Corp in particular. He states,

"The contracts under which they operate are largely comparable to those with profit-making firms, including a fee on top of costs and overhead, indistinguishable from profit taking..."

“with a fee that is at present 5 percent and in some years has been higher, Aerospace is receiving a higher percentage of fee in terms of retained income than a profit-making firm [quoted from a report by the House Committee on Armed Services, Subcommittee for Special Investigation]...²⁵

The point here is that FFRDC personnel are quite clearly contractors, whether they have “non-profit” status or not.

Should we nevertheless decide grant special dispensation for FFRDC personnel in our military value analysis, are we prepared to then open the door for other types of special contractor personnel? One such type is the Post-doctoral researcher. They work side-by-side and co-publish with their civil service colleagues. They have special security badges that distinguish them from other contractors, and as a result they have out-of-hours privileges that are not granted to typical contractors. They form a sizeable pool of talented future recruits; so giving them military value in our analysis would be a way of accounting for future Intellectual Capital at the host lab or center. They are high quality talent and are managed by *non-profit* professional associations (e.g., the National Research Council).

Does this mean that Post-docs should be considered “government workers”? No. Post-docs are contractors. Does this mean they should nevertheless be included in our scoring for military value? No. BRAC is not about the infrastructure or workforce of the private sector. This is merely to say that if we grant special dispensation to FFRDC personnel, then we had better be prepared to give fair and equitable treatment to other types of contractors that reside in a special status. But, that step will then open up another debate — what constitutes “special?”

Proposal #8: Eliminate the Effects of Service Policy and Organizational Differences

Two of our metrics will be significantly affected by differences in policy and organizational structures within the Armed Services. Each would be easy to fix.

(1) *Organizational Differences for Extramural Funding*. The corporate laboratories of the Army and Air Force, (i.e., ARL and AFRL), have organizationally embedded extramural funding organizations — the Army Research Office (ARO) and the Air Force Office of Scientific Research (AFOSR). Both ARO and AFOSR are organizational analogs to the Office of Naval Research (ONR), which performs the same function for the Navy. However, ONR is NRL’s parent command, and is not embedded within NRL. As such, ONR reports to the TJCSG data calls separately. Unless ONR’s funding is included, there will be an artificial penalty given to NRL for metrics that score extramural dollars, such as the Workload Focus metric.

There are two ways to fix the problem: include ONR’s funding in the NRL totals, or exclude ARO and AFOSR funding. Either approach can be addressed during the scoring phase and do not necessitate changes to the military value questions. This would level the playing field by eliminating an artificial organizational structure-based difference among the Services.

(2) *Scoring DAWIA Certification for the S&T Function*. The Navy regards its S&T workforce as technical performers, not acquisition managers, and therefore does not mandate universal Defense Acquisition Workforce Improvement Act (DAWIA) certification. The Army and AF implemented universal DAWIA certification throughout their S&T workforces. Because of this difference in Service policy, there will be a major *artificial* MilVal difference between NRL and the other Services’ corporate labs for this metric. This difference will have no real military value

²⁵ Harold L. Nieburg, *In the Name of Science* (Chicago: Quadrangle Books, 1966), p. 254.

significance, but would unfairly penalize one corporate laboratory by at least 4 percent for a different, but valid, Service policy.

This artificial difference is also easy to fix. Given that two of the five subgroups have zeroed out the weighting value of DAWIA certification for S&T in their scoring plans, it should be deleted in full for all S&T scoring. This would provide uniformity among the subgroups. It would also ensure defensibility of the process by leveling the playing field through the elimination of an artificial policy-based difference among the Services.

It should be noted that the 4 June ISG guidance changes the metric so that only those personnel at GS-14 and above are scored. Unfortunately this does not solve the problem, a fact that will be demonstrated later in this paper.

Proposal #9: De-Emphasize Dollars as an Indicator of Military Value

Fourteen of the 16 Operational Impact metrics use funding to measure output — despite the fact that dollars are a rather standard resource *input*. Oddly, despite talk by some about measuring “product” by these metrics, a few of them (i.e., “ATDs Currently In Work”, “ACAT Systems Currently in Work”, and “Workload Focus”) do not even meet the standard of having made an operational impact because they measure the total funding of work *in progress*. In other words, these metrics do not measure “product.”

All of the current “dollar metrics” suffer from the following problems:

- Dollars show level of investment, but reveal nothing about the work’s impact on warfighting capabilities.
- They introduce the use of an *input* resource as a surrogate output measure for Operational Impact.
- They are founded on an unsupportable assertion that *more dollars equals greater Military Value*, an assertion that will be hard to defend to the Commission (in one sense the opposite is likely true — large dollar contracts represent today’s, and in some cases yesterday’s, technology, not the disruptive, transformational technologies of tomorrow).
- The lack of specificity in the ground rules for reporting “intramural” versus “extramural” funding will lead to double-counting, or worse.

Two of the metrics — the *Workload Focus* and *Future Warfighting Capability* metrics — are especially problematic because they include contract dollars executed in the private sector. This is a dubious approach to determining military value for the reasons above, plus those cited below.

- They convey the scale of private sector infrastructure into an analysis of the public sector. BRAC is about closing, reducing, and/or realigning government, *not private sector*, infrastructure.
- They grant military value to a resource that is highly fungible, i.e., dollars can be awarded by most any DoD R&D contract shop.
- They introduce the use of a resource that will not be closed, reduced, and/or realigned due to any decision by the Commission.

Furthermore, what do dollars spent show us? For example, if Site A does the same type of work at the same in-house level of investment as Site B, but Site B contracts out more money to the private sector,

can we really support the conclusion that Site B has more military value? And, can we defend the assumption that “more dollars equals greater military value” (i.e., he who spends the most wins) to either the Commission or the communities?

Finally, the “dollar metrics” are rife with triple-counting where many of the same dollars will be scored multiple times under three different metrics: “Advanced Technology Demos Currently in Work”, “Workload Focus”, and “Future Warfighting Capability.” So, dollars with no discernable value to our analysis get counted three times, but *income* with confirmed value gets ignored (see next section).

Proposal #10: Include Income from Industry

The S&T metric for “Tech Transitions” counts transitions from S&T projects into DoD acquisition programs, but it misses *direct* transitions to industry via legal arrangements known as Cooperative Research and Development Agreements (CRADAs). Income to the labs and centers from CRADAs should be captured and scored because the private sector dollars received in fact confirm the quality and potential promise of the work. Industry would not provide resources for public sector research that is poor in quality or show little promise of application.

CRADAs have been used successfully for such things as the development of novel techniques for airport luggage screening, detectors for drugs of abuse, location and mapping devices for unexploded ordnance, fiberoptic dosimeters for radiation exposure, and other technological advances that have impacted the military and the civilian sectors. Nevertheless, the metric has not been adopted because of the claim that CRADAs occur too early in the process to make an operational impact. That is a curious argument given the as yet null operational impact made by “ACATs in Work,” “ATDs in Work,” “Workload Focus”, and “Future Warfighting Capability.”

In short, the Operational Impact metrics count dollars contracted to the private sector (in some cases three times for the same dollar), so why not dollars *from* the private sector? This omission is especially odd when the latter is a much better indicator of quality than the former.

Proposal #11: Kill the Future Warfighting Capability (FWC) Metric

Two factors make the idea of prognosticating a site’s *future* military value hard, if not impossible, to realize:

- Different missions (i.e., S&T, D&A, and T&E) yield different products, some of which will not have discernable military value for as much as 20 years or more (e.g., GPS).
- The dynamic and unpredictable nature of R&D, especially S&T, does not lend itself to predictions of future performance (e.g., witness the problems applying the Government Performance and Results Act to R&D agencies²⁶).

The analytical foundations for this metric, as currently fashioned, are much too speculative. How do we defend the dubious correlation between a site’s *future* military value and its *current* funding for technologies *pre-selected* by the TJCSG and prognosticated to be critical to the force structure in 2025?

²⁶ Testimony Before the Committee on Science, U.S. House of Representatives, *Managing for Results: Key Steps and Challenges in Implementing GPRA in Science Agencies*, (GAO/T-GGD/RCED-96-214), July 10, 1996.

A far more legitimate, effective, and defensible approach is to just identify today’s most innovative sites. With responsible stewardship of the in-house RDT&E community and infrastructure, they will most likely be the ones creating tomorrow’s new warfighting capabilities. As it stands now, this metric merely provides yet another opportunity to score the same dollars included and scored under the other “dollar metrics.”

Proposal #12: Break the Stovepiped Study Design

The BRAC-95 LJCSG performed its analysis by sorting 12 common support functions by 3 lifecycle areas: S&T, Engineering Development, and In-Service Engineering (T&E was addressed by a separate JCSG). That approach generated 36 possible “bins” of workload, as opposed to our 39. While past closure rounds are not the focus here, there is an important feature that our process shares with BRAC-95 — i.e., pushing highly interconnected work through technical and functional stovepipes.

It might be helpful to think of the 39 bins as a very complex Rubik’s Cube when the 3rd dimension of organization / installation is added. To be successful, we need to align the bins in a way that: *maximizes military value, reduces capacity, enhances jointness, and ensures an innovative end-state that meets transformation goals*. And, all this must be done in a way that does not sub-optimize the program of the organization / installation by severing innovative connectivity when workloads are realigned.

If that does not seem daunting enough, unless changes are made, those bins will be populated with data providing no clue as to the actual impact or value of the work. In that case, our process will be an arithmetic exercise where packets of workload are moved around in an almost arbitrary fashion, which is the inevitable result of having no data on work value. *This will sever the connectivity of critical multidisciplinary projects and vertically integrated programs, as well as decapitate top talent from any realigned work.* Realignment by arithmetic is not good enough.

A real example clarifies the problem. Last December, the DDR&E presented a prestigious award to a Navy researcher whose work on inter-operability and embedding simulations within C4I systems “represent major scientific advances satisfying critical military information technology requirements.” However, the award recipient was not from a “C4I lab” (thus demonstrating the dubious nature of the Workload Focus Metric, i.e., the “Jiffy Lube Metric”). Further, the researcher is a specialist in tactical electronic warfare (EW). In what bin should this EW specialist’s work be assigned? Information Systems, Sensors & Electronics, Sea Vehicles, Air Platforms, or Ground Vehicles? The relevant point here is that the 39 bins do not have clean, mutually exclusive borders.²⁷

After our analysis pulls the work of the labs and centers through the stovepipes, how many critical innovative connections will be severed by realignment? Given Dr. Frosch’s observation that “great R&D must preserve the connections between various kinds of knowledge,” we will potentially do a lot of damage. But one way to minimize this risk is to reengineer the analysis for conformance to real-world differences in missions, outputs, types of personnel, and connectivity.

²⁷ One benefit to this approach is that we may be able to tell if one, or more, of the Services attempts to “game” the answers to the data calls. If a Service issues “top-down” guidance to its field sites on how to fill the bins, then we may see a very clean *Service-wide* workload focus (from S&T through T&E) on a strategically limited number of *end-product* bins where inter-service competition is greatest (e.g., air platforms, weapons), vice a smearing of workload across a larger number of bins, which would likely be characteristic of data reported by field sites operating without a directed “game-plan.”

A solution is to compare the whole R&D program at a corporate lab to that of another corporate lab, and the whole RDT&E program at a warfare/product center to another warfare/product center. This way the horizontal connectivity at multi-disciplinary corporate labs would be evaluated intact, and the vertically integrated connectivity at warfare/product centers would be treated likewise. In addition, the military value of sites that maximize “collision cross-sections among kinds of knowledge” by performing significant levels of joint work would also be recognized.

This proposed solution means assigning Military Value at a higher level, such as at the activity / installation level, and not to the Rubik’s Cube “facilities.”

The Emerging Pattern: A “Dollars and No Sense” Approach

At this point, it should be beneficial to step back to assess the pattern that has emerged in the TJCSG’s Military Value data call. Some will disagree, but our interminable 6-month-long debate has exposed a noticeably high level of energy directed toward “dumbing-down” the metrics for Intellectual Capital. Clearly, there was no credible analytical defense for a numerical approach based on percentages, especially when it was shown to carry a significant risk of yielding irrational results. Fortunately, the ISG struck down that misguided approach, however we are still ignoring known world-class scientists (i.e., the competitively selected STs), allowing the unexplained downgrade of technical fellows from the Prestigious tier of awards/honors, subjecting innovative bench scientists to a DAWIA metric that is appropriate only for “technology shoppers”, and assigning ridiculously compressed point values for education, experience, and awards/honors.

Dollars, on the other hand, have achieved relative rock-star status as a preferred metric for military value. Fourteen of the 16 Operational Impact metrics use funding to measure output — despite the fact that dollars are a rather standard resource *input*. Eleven of these metrics (i.e., “ATDs Currently In Work” for S&T; “Workload Focus” for S&T, D&A, and T&E; “Future Warfighting Capability” for S&T, D&A, and T&E; Cost Effectiveness for S&T, D&A, and T&E; and ACAT Systems Currently in Work for D&A²⁸) measure funding for work *in progress*, so they fail to even meet the standard of having produced anything, much less something with an operational impact.

Most dubious of all, the “Future Warfighting Capability” metric uses *today’s dollars* to gauge a lab’s aptitude for inventing the unforeseen future technologies necessary to combat the unforeseeable threats of the year 2025.

If this paper’s thesis proves correct, we will not have the ability to identify the DoD’s highest-quality talent, which means we will fail to accurately appraise sites with that kind of talent. With its acute dependence on intellectual horsepower, this will greatly affect our assessments of S&T, which is the source of powerful, disruptive, new technologies. Given this concern, is there a way to test our analytical approach before it goes “prime time?” Can we bump a real-life, “world-class” S&T test group up against our analytical process to see how well it fares? Ideally, we would have a test group that was “benchmarked” recently as “world-class” by technical and military experts who examined the work directly and confirmed it to be of high potential pay-off and relevant to future national security needs.

Luckily, such a test group exists.

²⁸ While not measuring actual dollars expended, as done by the other cited metrics, this metric scores highest those ACAT programs with the biggest budgets.

Testing Our Analytical Process...

The Congressionally mandated “Section 913 Study”, completed in 2002 by the National Defense University (NDU), evaluated S&T projects at 10 DoD labs / centers for *their potential impact on future warfighting requirements* — an objective that is at the core of our BRAC endeavor. The customer for NDU’s study was the DDR&E.

(1) *The Defense Experts*. NDU’s defense experts included retired “4-stars:” General A. Zinni (replaced by General C. Wilhelm after being called away for State Department duties in the Middle East), Admiral H. Gehman (later to become the Chair, Columbia Accident Investigation Board), General W. Crouch, and General M. Carns. Technical experts included a former DDR&E (and Secretary of the Air Force), others with experience in DoD and non-DoD R&D management, and academic experts, of which one was also on the Columbia Accident Investigation Board. The project director was Dr. H. Binnendijk, Director of NDU’s Center for Technology and National Security Policy and a former Special Assistant to the President on the National Security Council.

(2) *The Test Group and its Work*: NRL’s Center for Bio/Molecular Science and Engineering (CBSE) was among a number of impressive research groups evaluated at the 10 DoD labs and centers. The CBSE performs a wide range of work in the area of biosensors and has been successful in transferring its technologies to industry. For example, a patented CBSE technology licensed to Lifepoint, Inc. was recognized by *Popular Science* magazine as one of the Top 100 Tech Innovations of 2002. The device, called the Impact System, is a drug and alcohol testing method based on the CBSE’s Flow Immunosensor. Another marker of success is the fact that the CBSE has attracted over \$1,000,000 in CRADA income from industry to support promising R&D. Also, the CBSE demonstrated the first fully automated optical biosensor for remote identification of biological warfare (BW) agents using a UAV.

Among the CBSE projects evaluated by the NDU team were two biosensors. The projects, RAPTOR and the Array Biosensor (AB), focus on developing portable, fully automated BW agent detectors for use by expeditionary forces, first responders, and UAV surveillance. In addition to detecting BW agents, these detectors are being developed to assure food/water safety, monitor treaties, and diagnose infectious disease. RAPTOR was successfully field-tested in Bahrain for monitoring Navy drinking water supplies and the technology was transferred to the U.S. Marine Corps. Funding sources for the projects have included the Navy, Marine Corps, Special Forces, Army, Defense Threat Reduction Agency (DTRA), and NASA (for monitoring recycled water and human arrivals to the Space Station). Because these two projects were evaluated by NDU, these will be the only CBSE projects scored in the following test.

(3) *Test Group’s Credentials*: Dr. F. Ligler, an ST and an elected fellow of SPIE (Society for Photooptical Instrumentation Engineering), leads the two projects. On 5 May 2004, she was personally congratulated at the White House after receiving the 2003 Presidential Rank Award for Distinguished Senior Professionals for her exceptional scientific accomplishments.²⁹ This was the ST cadre’s first year of eligibility for an award that had previously been given exclusively to managers. Dr. Ligler holds, and has held, many professional memberships, including the DoD Counter-terrorism Technology Advisory Group, the Department of State Technical Task Group L

²⁹ Dr. W. Bryzik, an Army ST, and member of a TJCSG sub-group, also received this prestigious 2003 award for his technical accomplishments. Dr. Bryzik strongly advocated the inclusion of STs in the TJCSG military value analysis, but his arguments were met with the same resistance that I experienced (i.e., an ST is “just a grade”).

(Biotechnology), Head of the U.S. delegation to NATO Panel 33 (Automated Sensors), U.S. representative to International Task Force 24 (New Detection Technologies), and the International Biosensors Congress.

Dr. Ligler recently received the *Homeland Security Award*, which is awarded by the Congressionally-established Christopher Columbus Fellowship Foundation. The award credits both Dr. Ligler and her staff with a,

“a profound impact on the development of bio-sensor based detection of biological warfare agents over the past 17 years, leading to significant improvements in the nation’s capability for early and accurate detection.”³⁰

As far as the staff’s intellectual capacity is concerned, there are 12 PhDs other than Dr. Ligler. Eight of the staff are postdoctoral researchers (i.e., postdocs). Another confirmation of the CBSE’s quality is a recent “Best Places to Work” survey of postdocs by *The Scientist*. A total of 91 institutions were included in the rankings, with NRL ranked 13th. Other top 15 institutions were the National Cancer Institute, Harvard School of Public Health, Lawrence Livermore National Laboratory, Cal Tech, and Vanderbilt University Medical Center.³¹

It was the work of this world-class research group that impressed the NDU study team when it visited on 7 September 2001. In particular, Admiral Gehman, stated that the,

“Center for Bio/Molecular Science and Engineering demonstrated bio-warfare detection systems that are light years better than anything I have seen in DOD, CDC, DTRA, the Raid teams or anywhere else. This Center may be a life-saver sooner than we think.”³²

Admiral Gehman’s comments carry a prescient quality given that his critique was made a year or so before the White House and the Columbus Foundation conferred the above awards, and only four days before the terrorist attacks on America.

...By Targeting Exceptional Talent

Lest my choice of test group be regarded as parochial, it should be emphasized that the CBSE was not the only group to impress the NDU study team. For example, there is the Quantum Cascade Laser work of ARL’s Dr. K.K. Choi, who was recently inducted into the Space Technology Hall of Fame. His work supports Army free space communications and chem-bio sensing programs.³³ And, there is the work of the SPAWAR Systems Center (SSC) in the area of nonlinear antenna technology. This work was found by the NDU team to be,

“Impressive and contained some potentially very transformational thinking related to the shipboard antenna problem...This work demonstrated that the Center is capable of attracting top-notch basic research people and doing first-rate science.”³⁴

³⁰ <http://www.ccolumbusfoundationawards.org/homeland/past.cfm>

³¹ Anderson, Grimwade, Hunter, and Park, “Best Places to Work Survey: Postdocs Speak Up”, *The Scientist*, (Feb 16, 2004), 17-19.

³² Section 913 Report #1: *Sensors Science and Technology and the Department of Defense Laboratories*, (National Defense University: March 2002), p.31.

³³ Ibid., p. 9

³⁴ Section 913 Report #2: *Information Science and Technology and the Department of Defense Laboratories*, (National Defense University: July 2002), p.10.

The NDU team was also impressed with the intellectual climate at ARL-Aberdeen, the Armaments Research, Development, and Engineering Center, and the Aviation and Missile Research, Development, and Engineering Center. Upon evaluating the weapons work of S&T groups within these three sites, the study concluded that:

“The Army ballistics program is one of the areas wherein the Army holds a dominant lead technically and the private sector carries out the Army’s concepts. There is no concern over privatizing this work.”³⁵

The reason none of these groups were chosen to be the test group is because I have no personnel, funding, and programmatic data by which to score them. On the other hand, as an employee of NRL, I know enough about the CBSE to score it accurately. Regardless of my choice, all of these groups have exceptional S&T personnel. Therefore, *the exceptional talent working on the Quantum Cascade Laser at ARL-Adelphi, nonlinear antenna technology at SSC-San Diego, and ballistics S&T at ARL-Aberdeen, ARDEC, and AMRDEC will no doubt share the NRL CBSE’s fate in our military value analysis.*

The Test Design

This following test compares the CBSE against two hypothetical S&T groups: Project X and Project Y. Project X is a two-person shop that manages a contracted biosensors Advanced Technology Demonstration (ATD) program of 15 million dollars. Project Y is also an ATD, performed mostly by on-site personnel at the same \$15 million dollar funding level. Both have unexceptional staff and no product, or impact, thus far. *The data is purposely constructed this way in order to isolate the effect that both dollars and “dumbed-down” people metrics will have on the overall military value scoring of a benchmarked world-class research group.*

The data required for scoring are shown below in Table 1. The column for the CBSE contains real data.

Table. 1

	<u>CBSE</u>	<u>Program X</u>	<u>Program Y</u>
PhDs	13	2	2
MSs	2	0	23
BSs	0	0	10
STs	1	0	0
<10 yrs	11	0	5
10-20 yrs	4	0	15
>20 yrs	0	2	15
DAWIA (3)	n/a	2	5
DAWIA (2)	n/a	0	5
DAWIA (1)	n/a	0	5
DAWIA (Multi-3)	n/a	2	5
Patents	3 (1 is licensed)	0	0
Society Fellow	1	0	0
Funding/Year	\$750,000	\$15,000,000	\$15,000,000
IH / OH Ratio	\$500 K / 250 K	\$0.5 M / 14.5 M	\$10 M / 5 M
Tech. Transitions	1	0	0
ATD	No	Yes	Yes

³⁵ Section 913 Report #3: *Weapons Science and Technology and the Department of Defense Laboratories*, (National Defense University: December 2002), p.18.

Draft Deliberative Document – For Discussion Purposes Only
Do Not Release Under FOIA

Rapid Response Awards	Yes (but no UNS) Homeland Security Award Presidential Award	No None	No None
CRADA Income	\$90,000	None	None

The calculations and the assumptions used are provided in Appendix A. The data is scored using three methods, of which two were developed by the TJCSG. I call the first one the “Mass Approach” (which, after the ISG’s 4 June directive, is the TJCSG’s official process), and the second is called the “Mass Plus Percentage Approach” (which was our official process for 5 months). The third scoring method is called the “Defensible Approach,” which is a collection of alternatives and proposals, most of which have been advocated in the three earlier versions of this paper over the last four months.

The point of this test is to determine whether our analytical process can evaluate each site accurately. If our process is sound, then the CBSE will rank highest in military value given: (1) its world-class credentials, and (2) the unexceptional personnel and non-existent operational impact of Project X and Project Y.

On the other hand, if the CBSE does not rank highest, then we have a big problem on our hands — *and it is better to know that before we do irrevocable harm to national security.*

The Results — *Blind to Exceptional Talent and Field Impact*

We have a big problem on our hands. The CBSE ranks *second with just more than half the value* of the top-ranked Project Y under the official TJCSG scoring approach. It ranks *last* under the former TJCSG approach. This is a strong sign that we will be unable to identify exceptional talent and work. Therefore, *our analysis will almost certainly fail to accurately evaluate each site’s military value, especially those with exceptional talent.*

- The *Mass Approach* ranks the CBSE second, with a total Military Value (MV) of **3.50** (1.93 for Intellectual Capital and 1.57 for Operational Impact). This is just more than half of the top-ranked Project Y’s score of **6.06** (3.00 for Intellectual Capital and 3.06 for Operational Impact). Even more amazing, it ranks only 18% higher than Project X’s score of **2.96** — which is an unexceptional, 2-person, no-impact program with a fat wallet.
- The *Mass Plus Percentage Approach*, which was our official approach until the ISG fortunately rejected it, yields absurd results that are not hard to fathom given the previously described vagaries of the Percentage Approach. The CBSE ranks last, with a total MV of **3.97**. *This is less than ¾ that of the top-ranked 2-person Project X (5.51).*

First, let us look closely at the Intellectual Capital metrics. The CBSE, with a 13-PhD staff of widely recognized *world-class* talent, ranks second to Project Y’s personnel — an unexceptional, undistinguished staff with just 2 PhDs. This kind of outcome was predicted months ago by earlier versions of this paper and in emails³⁶ to the TJCSG CIT. Now just how insensitive are these results to

³⁶ In Attachment F, the Air Force claims on page 4 that its sites would be more fairly scored against Army and Navy sites by use of the Percentage Approach. This claim raises a question. How does the Air Force know that the Percentage Approach has this effect given (a) that approach’s grossly unpredictable dynamics, (b) the large number of sites and “bins” to be compared, and (c) that official data have yet to be collected? By contrast, the best this paper can do is offer the much more limited claim that our study cannot identify intellectual excellence and

exceptional talent? If we add 10 Nobel Laureates, all with PhDs and more than 20 years experience, the world-class group still finishes second to Project Y in Intellectual Capital (2.61 versus 2.85).

Such irrational outcomes for Intellectual Capital are almost certain to occur in our BRAC scoring, and will probably do so on a widespread basis. Foreseeing this problem was not hard. They are due to the absurdly compressed point values, as well as the fact that the DAWIA metric creates a penalty for an artificial policy difference with no military value significance. Navy policy does not require “hands-on” R&D performers to attend classes for DAWIA certification. Such certification is appropriate for *acquisition* personnel, or “buyers” of technology, but not the innovators themselves.

But it is the Operational Impact metrics (i.e., the “dollar metrics”) that bury the world-class research group. *This is odd given that the CBSE is the only group of the three to deliver a product, or an impact, with a technology transition and a successful rapid response to a Navy need in Bahrain.* Most surprising is that the metric designed for gauging a lab’s potential to create transformational technology (i.e., “Future Warfighting Capability”), scores the CBSE last — roughly 80% less than Project Y. This is incompatible with Admiral Gehman’s expert assessment of the CBSE’s potential importance to national security, not to mention the special recognition conferred upon it by President Bush and by the Columbus Foundation, which has a Congressionally-established mission to “encourage and support research, study and labor designed to produce new discoveries in all fields of endeavor for the benefit of mankind.” *The relevant point here is that it took only a fat wallet to negate the confirmed military value of the CBSE.*

Finally, the CBSE scores a zero in “Rapid Response” and a minimal value in “Cost Effectiveness” due to the disqualification of its successful rapid response project. It is disqualified because it was not initiated by official paperwork known as an “Urgent Needs Statement” (UNS). The TJCSG’s analytical process does not allow for the fact that urgent field requirements are not all met in response to formal bureaucratic requests.

So, what does this test show us?

To again use Plato’s allegory, it shows that exceptional technical talent and real field impact *does not even cast a shadow*. In fact, our approach appears to be blind to locating “the best and brightest” and properly appraising operational impact. The test shows that the “dumbed-down” people metrics artificially suppress the CBSE’s high talent quotient, and then the dollar metrics bury it from view. Bear in mind that when we perform the real analysis, this data will be mixed in a large “bin” with a variety of similar, and not-so-similar, data in a manner that eliminates any means of identifying discrete research groups. *Unless its data rise to the top, the CBSE will disappear without us knowing it was ever there.* Then, using our stove-piped analysis, its staff and work would likely be realigned elsewhere. Should that occur, the research group that impressed Admiral Gehman for its potential importance to our country would be lost.

This test also shows that *both TJCSG approaches are certain to generate irrational results*. In fact, it shows us that the bias of the Percentage approach is as obvious as a tarantula sitting astride a slice of angel food cake. While a tad subtler, the point compression issue (where the value of one PhD is offset by just three BSs) is no less problematic when used alone (as in the Mass Approach), or in combination with the Percentage Approach. In short, point compression and percentages result in “dumbed-down”

meaningful operational impact, whether in the Army, Navy, or Air Force. How a given Service, or even site, will fare is completely unknowable without extensive data from all the Services.

people metrics, and *neither scoring approach will be defensible before the BRAC Commission or the communities.*

Proposal #13: Use a Defensible Approach

On the other hand, the third approach used in this test case, the Defensible Method, yields results one would expect given the data. This method is a collection of alternatives and proposals that have been advocated over the last four months, many of which appeared in earlier versions of this paper.

- The *Defensible Approach* ranks the CBSE first, with a total MV of **5.69** (3.00 in Intellectual Capital and 2.69 in Operational Impact). This is 20% more than Project Y's total of **4.71** (1.65 in Intellectual Capital and 3.06 in Operational Impact), and just more than 100% more than Project X's total of **2.73** (0.22 in Intellectual Capital and 2.51 in Operational Impact). Most importantly, the big difference is made by the fact that the CBSE scores well above the other two sites in Intellectual Capacity, as one would expect it should.

The “Defensible Approach” is comprised of the following scoring rules: (a) “Absolute Number” scoring approach for all people metrics, (b) a 70:10:1 point spread for the Awards Metric, (c) a 10:3:1 point spread for the Education Metric, (d) collecting STs and scoring them under the Experience metric at a point spread of 70:3:2:1, (e) restoring the Technical Society Fellows to the Prestigious tier of the Awards/Honors metric, (f) eliminating the superfluous DAWIA Certification as a scoring metric for S&T, (g) relaxing the Rapid Response requirement for a formal Urgent Needs Statement, (h) including CRADA income, and (i) allowing the flexibility to expand the very limited prescribed lists of “Elite” and “Prestigious Awards” to include such awards as the *Homeland Security Award* and the *Presidential Rank Award for Distinguished Senior Professionals*, both of which, amazingly enough, are not allowable for scoring in the two TJCSG approaches.

Conclusion

Left unchanged, our current course will produce a “Dollars and No Sense” approach to assessing military value. People metrics that fail to discriminate extraordinary talent from the mediocre, coupled with Operational Impact metrics that score by the dollar, will favor sites that spend a lot of money irrespective of their technical competence. To put it bluntly, that is a threat to national security.

Intellectual capital is the most critical resource to defense R&D (the physical environment is probably the most important for T&E). All else is secondary. By identifying the best talent in the DoD RDT&E system we also make the rest of our BRAC job easier because exceptional talent is an indicator of other important parameters. For example, the best talent does not choose to work with lousy facilities. It does not choose to work for an organization with no record of success and no chance to make a difference. It does not choose to work with mediocre colleagues and poor leadership. And, it does not choose to work on yesterday’s problems. If we can find exceptional talent, we will find state-of-the-art facilities, capable leadership, top colleagues, a record of impact on the nation’s security, a powerful desire for success, and a staff working on tomorrow’s challenges. *Find the best talent, and the rest falls into place.*

These 13 proposals do not ensure success, but they offer a good chance for it. Together, they present some advantages and solve a number of critical problems:

- For the first time in any BRAC round, military value would be based, at least in part, on a site’s actual level of effectiveness in meeting national security interests, both in terms of warfighting impact and contributions to science and technology

Draft Deliberative Document – For Discussion Purposes Only
Do Not Release Under FOIA

- Creative synergy (i.e., connectivity) would be protected in its three forms: multidisciplinary projects, vertical integration, and jointness
- Intellectual capital would be accorded the importance it deserves, and the metrics that measure it would not be “dumbed-down” to the point where critical differences in workforce expertise are lost in the noise of large aggregate populations
- The analytical process would conform to real-world differences in missions, culture, outputs, and types of personnel
- The analytical process would eliminate: the unsupportable assertion that *more dollars equals greater military value*, the risk inherent in using personnel data from private companies that are financially dependent on the continued operation of the host laboratory / center, and the effects of artificial differences in Service policy and organizational structure
- The analytical process would avoid the sub-optimal realignments that result from a process that segregates workload by technical and functional stovepipes
- Test range operations would be evaluated for their real-world connectivity
- By identifying today’s most innovative and agile sites, it will not be necessary to link unknowable future directions of technology to tomorrow’s unforeseeable warfighting requirements

We must not damage the competence of the in-house system as a performer of long-term, high-risk work. We must not cripple the ability of the laboratories, centers, and test ranges, to respond to crises. We must not break the “yardstick” that serves as the Pentagon’s strongest voice for independent, authoritative technical advice that is insulated from commercial pressures to make a profit.³⁷ And we must avoid the pitfalls of the previous BRAC.

Much rides on our decisions, even more so than ten years ago. Our country is engaged in a prolonged struggle with an opportunistic, fanatical enemy who has unlimited apocalyptic goals and is not deterred by traditional means. We need all of the technical options we can get. Moreover, the fast global pace of technological change will continue, and disruptive new technologies are unlikely to all be invented here. A creative and agile in-house system of defense labs, centers, and test ranges will be of great importance to our nation’s security. We cannot afford to make the big mistakes that most assuredly will happen if our eyes remain fixed on the shadows.

Don J. DeYoung
Capabilities Integration Team, Alternate
Technical Joint Cross Service Group
U.S. Navy

18 June 2004

³⁷ DeYoung, “The Silence of the Labs,” *Defense Horizons*.

APPENDIX A

Assumptions and Military Value Calculations for Case Study

A. ASSUMPTIONS

(1) The Case Study tests the TJCSG's analytical approach for the S&T function only, using only the metrics for "Intellectual Capital" and "Operational Impact." While partial in scope, these metrics should be sufficient to identify where the DoD's best S&T talent resides.

- Metrics for "Physical Structure & Equipment" were not scored due to the unavailability of data. This is mitigated by the fact that the "Intellectual Capital" metrics should identify the best talent, and the "Operational Impact" metrics should identify where the work with the greatest impact is done. All other things being equal, the best minds will gravitate to the labs with the best equipment and the best work will tend to be performed with the best equipment.
- Metrics for "Synergy" were not scored due to the large number of required assumptions, many of which would be arbitrary. This is mitigated by the fact that there is a strong consensus among the TJCSG working groups that the Synergy metrics will not differentiate among the sites.

(2) All metrics are given equal weighting.

(3) All "dollar metrics" assume constant level of funding from FY01-04. To keep it simple, dollar amounts are project-sponsored funds and do not include salary.

(4) The "Defensible Approach" is comprised of the following rules: (a) "Absolute Numbers" scoring for all people metrics, (b) a 70:10:1 point spread for the Awards Metric, (c) a 10:3:1 point spread for the Education Metric, (d) scoring STs under the Experience metric at a point spread of 70:3:2:1, (e) restoring the Technical Society Fellows to the Prestigious tier of the Awards/Honors metric, (f) eliminating superfluous DAWIA Certifications as a scoring metric for S&T, (g) relaxing the requirement for a formal Urgent Needs Statement, (h) including CRADA income, and (i) allowing expansion of the existing list of "Prestigious Awards" to include the *Homeland Security Award* (awarded by the Congressionally-established Christopher Columbus Fellowship Foundation) and the Presidential Rank Award for Distinguished Senior Professionals.

B. MILITARY VALUE CALCULATIONS

INTELLECTUAL CAPITAL METRICS

<u>Metric (1)</u>	<u>Education</u>	<u>Mass Approach</u>	<u>Mass + % Approach</u>	<u>Defensible Approach</u>	
Raptor/AB	$3(13)+2(2)=43$	MV=0.69	$43/15=2.9$ MV=0.97	$10(13)+3(2)=136$	MV= 1.00
Program X	$3(2)=6$	MV=0.10	$6/2=3.0$ MV= 1.00	$10(2)=20$	MV=0.15
Program Y	$3(2)+2(23)+1(10)=62$	MV= 1.00	$62/35=1.8$ MV=0.60	$10(2)+3(23)+1(10)=99$	MV=0.73
<u>Metric (2)</u>	<u>Experience</u>	<u>Mass Approach</u>	<u>Mass + % Approach</u>	<u>Defensible Approach</u>	
Raptor/AB	$2(4)+1(11)=19$	MV=0.24	$19/15=1.3$ MV=0.43	$70(1)+2(3)+1(11)=87$	MV= 1.00
Program X	$3(2)=6$	MV=0.08	$6/2=3.0$ MV= 1.00	$3(2)=6$	MV=0.07
Program Y	$3(15)+2(15)+1(5)=80$	MV= 1.00	$80/35=2.3$ MV=0.77	$3(15)+2(15)+1(5)=80$	MV=0.92

Draft Deliberative Document – For Discussion Purposes Only
Do Not Release Under FOIA

<u>Metric (3)</u>	<u>DAWIA Certification</u>		
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	0	MV=0.00	0 MV=0.00
Program X	3(2)+3(2)=12	MV=0.27	12/2=6 MV=1.00
Program Y	3(5)+3(5)+ 2(5)+1(5)=45	MV=1.00	45/35=1.3 MV=0.22
			n/a

<u>Metric (4)</u>	<u>Honors / Awards</u>		
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	5(1)+ 1(2)+ 2(1)=9	MV=1.00	9/15=0.6 MV=1.00
Program X	0	MV=0.00	0 MV=0.00
Program Y	0	MV=0.00	0 MV=0.00
			MV=1.00
			MV=0.00
			MV=0.00

<i>INTELLECTUAL CAPITAL</i>			
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	1.93	2.40	3.00
Project X	0.45	3.00	0.22
Project Y	3.00	1.59	1.65

THE “DOLLAR METRICS” (OPERATIONAL IMPACT)

<u>Metric (5)</u>	<u>Technology Transition</u>		
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	1 MV=1.00	1 MV=1.00	1+ [90K / 90K]=2
Project X	0 MV=0.00	0 MV=0.00	0 MV=0.00
Project Y	0 MV=0.00	0 MV=0.00	0 MV=0.00

<u>Metric (6)</u>	<u>ATDs Currently in Work</u>		
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	0 MV=0.00	0 MV=0.00	0 MV=0.00
Project X	15 x 3= 45.0 MV=1.00	1.0 MV=1.00	1.0 MV=1.00
	45.0/15 x 3=1.0		
Project Y	15 x 3=45.0 MV=1.00	1.0 MV=1.00	1.0 MV=1.00
	45/45=1.0		

<u>Metric (7)</u>	<u>Rapid Response</u>		
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	0 MV=0.00	0 MV=0.00	1 MV = 1.00
Project X	0 MV=0.00	0 MV=0.00	0 MV = 0.00
Project Y	0 MV=0.00	0 MV=0.00	0 MV = 0.00

<u>Metric (8)</u>	<u>Workload Focus</u>		
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	[.9(.5x3/10x3) + .1(.25x3/10x3) + 15/35] / 2 = 0.24 MV = 0.25	0.25 MV = 0.25	0.25 MV= 0.25
Project X	[.9(.5x3/10x3) + .1(14.5x3/14.5x3) + 2/35] / 2 = 0.10 MV = 0.10	0.10 MV=0.10	0.10 MV=0.10
Project Y	[.9(10x3/10x3) + .1(5x3/14.5x3) + 35/35] / 2 = 0.97 MV = 1.00	0.97 MV = 1.00	0.97 MV = 1.00

Draft Deliberative Document – For Discussion Purposes Only
Do Not Release Under FOIA

<u>Metric (9)</u>	<u>Future Warfighting Capability</u>	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	[.045 + .0025 + .43 + .75/15] / 3 = 0.18	MV = 0.18	0.18	MV = 0.18
Project X	[.045 + .1 + .06 + 15/15] / 3 = 0.40	MV = 0.41	0.40	MV = 0.41
Project Y	[.9 + .03 + 1 + 15/15] / 3 = 0.98	MV = 1.00	0.98	MV = 1.00
<u>Metric (9)</u>	<u>Cost Effectiveness</u>	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	[1 + 0 + 0] / 15=0.07	MV = 0.14	0.07	MV=0.14
Project X	[0 + 1 + 0] / 2=0.50	MV = 1.00	0.50	MV= 1.00
Project Y	[0 + 1 + 0] / 35=0.03	MV = 0.06	0.03	MV=0.06

<i>"DOLLAR METRICS"</i>			
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	1.57	1.57	2.69
Project X	2.51	2.51	2.51
Project Y	3.06	3.06	3.06

<i>TOTAL MILITARY VALUE: INTELLECTUAL CAPITAL METRICS + "DOLLAR METRICS"</i>			
	<i>Mass Approach</i>	<i>Mass + % Approach</i>	<i>Defensible Approach</i>
Raptor/AB	3.50	3.97	5.69
Project X	2.96	5.51	2.73
Project Y	6.06	4.65	4.71



ACQUISITION,
TECHNOLOGY
AND LOGISTICS

Deliberative Document – For Discussion Purposes Only – Do Not Release Under FOIA
THE UNDER SECRETARY OF DEFENSE

3010 DEFENSE PENTAGON
WASHINGTON, DC 20301-3010

JUN 4 2004

MEMORANDUM FOR CHAIRMAN, TECHNICAL JOINT CROSS-SERVICE
GROUP

SUBJECT: Technical Joint Cross-Service Group Report on Military Value Analysis

The Infrastructure Steering Group (ISG) has reviewed the Technical JCSG Military Value report submitted to it on March 25, 2004. The ISG appreciates the dedicated effort and military judgment that your members, as the experts in the field, put into revising this report in response to our initial comments. As part of the process to review issues raised by the Military Departments, the ISG tasked the BRAC Deputy Assistant Secretaries (DASs) to develop recommendations for resolving issues with the military value reports. The BRAC DASs met to review the issues pertaining to your JCSG. Representatives from the TJCSG participated in the discussion with the DASs to ensure the issues were well understood and potential solutions were implementable and reasonable. Subject to incorporation of the comments expressed herein, the ISG approves your report as the basis for your military value analysis.

In its report, the Technical JCSG proposes the use of out-year spending data as an indicator of the “Future Warfighting Capability.” The use of out-year dollars is not appropriate for assessing the military value of facilities where your functions are currently being performed. The metric “Future Warfighting Capability” should be limited as follows:

$$S(fwc) = [.9X(FTFEi/MTFEi_i) + .1X(FTFEe/MTFEe) + (FTFEf/MTFEf) + (FFTEs/MFFTEs)]/3$$

FTFEi = funding executed internally by the technical facility (includes personnel salaries) over the last three years (FY01-03)

MTFEi= maximum funding executed internally by any like-facility (includes personnel salaries) over the last three years (FY01-03)

FTFEe = funding executed externally by the technical facility over the last three years (FY01-03)

MTFEe = maximum funding executed externally by any like technical facility over the last three years (FY01-03)

Deliberative Document – For Discussion Purposes Only – Do Not Release Under FOIA

Attachment A

FTFEf = funding appropriated identified over the FYDP in FY 04 (FY04-FY10) by the technical facility

MTFEf = maximum funding appropriated identified over the FYDP in FY 04 (FY04-FY10) by the technical facility

FFTEs = In house FTEs at the technical facility over the last three years (FY01-FY03)

MFFTEs = maximum # of FTEs at any like facility over the last three years (FY01-FY03)

In its report, the Technical JCSG asks for specific personnel names in addition to qualifications. The TJCSG said that the names were requested to assist in auditing the data. The facilities submitting the personnel qualification information must have source data indicating the specific individuals having the qualifications and this source data is sufficient to allow auditing of the input. The Military Departments and Defense Agencies have data certification procedures to prevent respondents from providing false information. The audit process is the method to ensure the integrity of this process. Therefore, revise the relevant questions to eliminate the request for specific personnel names.

The rationale for the weighting and scoring process proposed by the TJCSG must be included somewhere in the group's formal record. This can be in the Military Value report or in the group's deliberative meeting record. Please review the group's deliberative meeting minutes and revise the Military Value report as necessary to provide a complete record of the rationale for the weighting and scoring process in the TJCSG formal record.

The TJCSG's current Military Value scoring plan uses percentages instead of absolute numbers as a measure of value in certain people-related areas. The use of percentages would skew Military Value by equating large and small installations. Please revise your report to change the method of measurement to use absolute numbers.

Because of differences in the way the military departments apply the Defense Acquisition Workforce Improvement Act (DAWIA) certification requirements, please revise your report to only request DAWIA certification data for those employees who are at the grade level of GS-14/NH IV and above.

The TJCSG proposes to use the qualifications of on-site contractors in their military value calculation. Because of the Office of the Inspector General concern about the auditability of this contractor data as well as the concern that this data is not consistently available at all locations, the metrics and related questions that measure the qualifications of on-site contractor data need be stricken from the TJCSG military report.

After incorporation of the directed revisions above, please provide a final copy of your Military Value Analysis report to the OSD BRAC Office no later than two weeks after the date of this memorandum. If you have any questions regarding these comments, please contact Peter Potochney, Director, Base Realignment and Closure, at 614-5356.



Michael W. Wynne

Acting USD (Acquisition, Technology & Logistics)
Chairman, Infrastructure Steering Group

cc: Infrastructure Steering Group Members
MilDep BRAC DASs

**Close**

From: DeYoung, Don
To: 'Short, James'
Cc: De Young, Don; 'Robert.Buckstad@osd.mil'; 'Harshad.Shah@osd.mil'; 'Andrew.Porth@osd.mil'; 'rflorence@dodig.osd.mil'; 'desalvapn@mcsc.usmc.mil'
Subject: TJCSG Minutes
Sent: 3/20/2004 6:56 PM

Importance: Normal**DELIBERATIVE DOCUMENT - FOR DISCUSSION PURPOSES ONLY****DO NOT RELEASE UNDER FOIA**

Jim,

Please ensure that the minutes of the 19 March 2004 TJCSG meeting contain my objection to the percentage approach for the People Metrics. Also, please make sure the minutes show that I advised the DDR&E that in all likelihood we will be in a position where "we have to defend an approach that yields irrational results." You may want to add the specific case I described to him.

The plausible scenario for the Awards Metric that I described to Dr. Sega is included in the paper I sent to the subgroup leads in a 27 Feb email appended below. I haven't re-done the arithmetic yet, but irrational results may be even likelier now that the award point scales have been compressed from a ratio of 30:10:1 down to 10:1.

In emails subsequent to the one below, I challenged the subgroup leads several times to show me where I'm wrong, and to cite examples where the alternative Absolute Numbers approach yields irrational results. Other than the woodland botany advice I received about my tendencies to examine the colors of individual leaves in the forest, I've received no responses.

Without any evidence provided to show where I'm in error, I remain convinced that we have a very flawed study on our hands. I can think of no rational defense to our approach if it is challenged by the Commission or the communities.

v/r,

Don DeYoung
U.S. Navy, CIT Alternate

-----Original Message-----

From: DeYoung, Don
To: 'Mathes, Thomas'; 'Rohde, Robert S Dr SAALT'; 'Blake Christopher L SES HQ AFMC/XP'; Mleziva Matt Civ ESC/NI; Shah, Harshad C Mr OSD-ATL; Shaffer, Alan Mr OSD-ATL; Ryan, George R CIV; Goldstain Alan B Civ AEDC/CD; Berry, William Dr OSD-ATL; "karen.higgins@navy.mil"; "schuette@nrl.navy.mil"; "brian.simmons@dtc.army.mil"

Cc: Buckstad, Robert COL OSD- ATL; Short, James Dr OSD-ATL; Strack, Gary Mr SAF/IEBB; DeYoung, Don
Sent: 2/27/2004 6:27 PM

Attachment B



Close

From: DeYoung, Don
To: 'Shah, Harshad C, Mr, OSD-ATL'; 'Buckstad, Robert, COL, OSD-ATL'
Cc: 'Ryan, George R CIV'; 'Cohen, Jay'
Subject: RE: Decision Time//BRAC FOUO -- MIL Val Rpt for Dr Sega Approval
Sent: 3/24/2004 1:46 PM **Importance:** Normal

Sirs,

I non-concur with the "Military Value Analysis Report" of the Technical Joint Cross Service Group (TJCSG), dated 24 March 2004.

I cannot with integrity approve a base closure study plan that has major, known, and yet uncorrected, flaws that will substantially affect the assigned military value for each of the Department's laboratories, warfare/product centers, and test ranges.

I have gone on record with my concerns, particularly to the “percentage approach” to determining the value of the Department’s intellectual capital. My objections have been made in a number of different forums, and each time I presented quantitative data to both substantiate my concerns and invite informed debate. The last was the 19 March TJCSCG meeting, where I showed how an irrational score is yielded for a likely scenario. In that meeting, the TJCSCG principals agreed to use the percentage approach – but only if the OSD BRAC office decides that individual contractor data is not sufficiently auditable.

On several occasions I invited the proponents of the percentage approach to show instances where the alternative “absolute numbers approach” yields irrational results. Without any responses to that request, I remain convinced that we are heading down a path that will lead to the discrediting of the study if it is challenged by the BRAC Commission or the communities. That the problems have been known for at least one month before final approval of the report will be especially difficult to explain. Use of the percentage approach risks placing the DDR&E in the untenable position of having to defend an analytical process that yields irrational results.

I am also concerned by the fact that the Scientific/Technical (ST) Corps will not be separately scored for military value. According to high level OSD officials and to the Army's Science and Technology Master Plan (1997), STs are "world class" scientists, some of the best and brightest in the Department's technical workforce.

In short, the report requires changes to ensure that we are not "dumbing down" the very metrics used to assess the value of the Department's most essential resource – its scientific and technical talent.

Very respectfully,

Don J. DeYoung
CIT Alternate
U.S. Navy

-----Original Message-----

From: Shah, Harshad C, Mr, OSD-ATL
To: 'Goldstayn Alan B Civ AEDC/CD'; 'Desalva Col Peter N'; Castle Fred F Brig Gen AF/XP; Shaffer, Alan, Mr, OSD-ATL; 'Blake Christopher L SES HQ AFMC/XP'; Ryan, George R CIV; 'Rohde, Robert S Dr SAALT'; 'Mleziva Matt Civ ESC/NI'; 'karen.higgins@navy.mil'; Berry, William, Dr, OSD-ATL; 'schuette@nrl.navy.mil'; 'MathesT@tacom.army.mil'; 'Blake Christopher L SES HQ AFMC/XP'; DeYoung, Don

Attachment C



Close

From: Mleziva Matt Civ ESC/NI [SMTP:Matt.Mleziva@hanscom.af.mil]
To: DeYoung, Don; Mleziva Matt Civ ESC/NI; 'Lawrence C. Schuette'; 'Shaffer, Alan, Mr, OSD-ATL'; Goldstain Alan B Civ AEDC/CD; 'Higgins, Karen L SES'; Blake Christopher L SES HQ AFMC/XP; 'Berry, William, Dr, OSD-ATL'; 'Simmons, Brian'; 'Rohde, Robert S Dr SAALT'; 'Shah, Harshad C, Mr, OSD-ATL'; 'Ryan, George R CIV'; "Mathes, Thomas"; 'schuette@nrl.navy.mil'
Cc: 'Buckstad, Robert, COL, OSD- ATL'; 'Short, James, Dr, OSD-ATL'; Strack Gary Mr SAF/IEBB
Subject: RE: People Metrics -- Part II
Sent: 3/2/2004 7:55 AM

Importance: Normal

Don, et.al. - my overall observations are:

. that there are a variety of views (percentage, absolute, combinations, etc.) that one can take to each metric
. not only must each metric be "fair and equitable" but the overall set must be "fair and equitable" too - this leads me to part company with the argument that all metrics (or an attribute set) must use the same view - views should be tailored to provide differentiation between reporting facilities, and must be "fair and equitable" for the topic being measured - hence I can rationalize quantity for one metric, quality for another and some combination for yet another

So for education, I envision "normal" organizations (not "corner cases" we can all construct to prove a point) and I think the objective to measure is the quality of the organization as the quantity is measured in other metrics - I think having all the metrics in a set being either quality or quantity may not be "fair and equitable" as it may inadvertently introduce a "cultural" bias in the result - hence diversity of "views" helps insure a lack of inadvertent bias

On the suggestion of changing the multipliers from (e.g.) 1,2,&3 to 1,20 & 30 that does not pass my "giggle" test - that is, is an MS worth 10x a BS - I have both and I can't rationalize that - and the same for grade/experience - and I have experience with both of those (no pun intended)

In summary, I've considered the arguments to change the Education metric and am not convinced the alternatives are an improvement on the current version

Cheers, Matt

-----Original Message-----

From: DeYoung, Don [mailto:deyoungd@ndu.edu]
Sent: Monday, March 01, 2004 9:31 PM
To: 'Mleziva Matt Civ ESC/NI'; 'Lawrence C. Schuette'; 'Shaffer, Alan, Mr, OSD-ATL'; Goldstain Alan B Civ AEDC/CD; 'Higgins, Karen L SES'; Blake Christopher L SES HQ AFMC/XP; 'Berry, William, Dr, OSD-ATL'; 'Simmons, Brian'; 'Rohde, Robert S Dr SAALT'; 'Shah, Harshad C, Mr, OSD-ATL'; 'Ryan, George R CIV'; "Mathes, Thomas"; 'schuette@nrl.navy.mil'

Cc: 'Buckstad, Robert, COL, OSD- ATL'; 'Short, James, Dr, OSD-ATL'; Strack Gary Mr SAF/IEBB; DeYoung, Don
Subject: RE: People Metrics -- Part II

Matt,

In my last email I asked if you would share your thoughts on how the Percentage approach plays out with the Awards metric. In the interim I've done some thinking on our mutual concern about quality and how it relates to the Education and Experience metrics.

What did you think about the plausible scenario I provided last Friday? It had Site A with 100 S&Es (50 PhDs, 50 MSs), and Site B with 1,000 S&Es (500

Attachment ▷

PhDs, 200 MSs, 75 BSs, 100 ADs, and 125 with no degree. With the Percentage approach, Site B is rated 20% lower in military value -- despite having 10 times the PhDs, 4 times the MSs, as well as the talent that 300 less degreed, and generally younger, individuals have to offer. Given our shared interest in measuring quality accurately, this irrational result must be a major source of concern for you as well.

A very interesting aspect about the Percentage approach is that it negates the value of large masses of high quality talent. For example, let's take the site you hypothesized with 100 S&Es (where all 100 have PhDs), and assume another site has 1000 S&Es (all 1000 have PhDs), and just for laughs, a third site has 1,000,000 S&Es (all one million have PhDs). By the Percentage approach, each would have the same MV. And, all would be equivalent even to a site with 1 person, who is a PhD. Clearly, these are irrational results, as I'm sure you'd agree. Unlike the plausible scenario I provided last Friday, the wrinkle with this scenario is its low likelihood due to the artificially uniform percentages across the board.

I read your comments below about the 100 PhDs vs. the 100 PhDs/300 BSs, and had the following thought. What if we took the same approach with Education as we are doing with the Awards? By that I mean construct a scale with larger differentials between a PhD and a BS, like 30 points for a PhD, 10 for a MS, and 1 for a BS. Using this scale and the Absolute numbers approach, Site A with 100 S&Es (all PhDs) would score a MV of 0.9, and the site with the 1000 S&Es (100 PhDs/300 BSs) would score 1.0. That sounds rational.

We would do the same with Experience. An ST would get 30 points, persons with greater than 20 years would get 10 points, 10 to 20 years would get 5 points, and zero to 10 years would get 1 point. The STs would not be counted under the >20, 10-20, and <10 bands to eliminate double-counting.

I know you and others have said that an ST is "just a grade," but there are mountains of evidence to show that is just not the case. Those positions are for "world-class" technical experts that the DoD wants to retain as productive, innovative scientists and engineers. The ST position grants the world-class expert greater compensation without being encumbered with ill-fitting management duties. Saying these positions are "just a grade" is like saying the McKinley Climatic Chamber is just a building, and that the Edwards' test range is just a high desert plain. Their omission will elicit skepticism about the merits behind our process, and inevitable challenges on this issue will be impossible to defend.

Again, since we're both concerned about the need to judge the quality of DoD's technical talent accurately, I'd be interested in what you think about these ideas. As for me, they go a long way toward rectifying our current approach, which at present yields very irrational results.

v/r,

Don

-----Original Message-----

From: Mleziva Matt Civ ESC/NI [mailto:Matt.Mleziva@hanscom.af.mil]
Sent: Monday, March 01, 2004 2:15 PM
To: 'Lawrence C. Schuette'; DeYoung, Don; Mleziva Matt Civ ESC/NI;
'Shaffer, Alan, Mr, OSD-ATL'; Goldstain Alan B Civ AEDC/CD; 'Higgins,
Karen L SES'; Blake Christopher L SES HQ AFMC/XP; 'Berry, William, Dr,
OSD-ATL'; 'Simmons, Brian'; 'Rohde, Robert S Dr SAALT'; 'Shah,
Harshad C, Mr, OSD-ATL'; 'Ryan, George R CIV'; "Mathes, Thomas";
'schuette@nrl.navy.mil'
Cc: 'Buckstad, Robert, COL, OSD- ATL'; 'Short, James, Dr, OSD-ATL';
Strack Gary Mr SAF/IEBB
Subject: RE: People Metrics

Afraid I'll have to part company with Larry on this one - the People discussion is the difference between the "quality" and "quantity" views of education - the quality view survived a long involved vetting process - and the group has shown it is willing to change the metrics/questions when convinced it is appropriate - I for one am not convinced that (e.g.) a facility with 100 PHDs and 300 BSs should get twice as much MV as a facility with 100 PHDs - the quality of the workforce of all PHDs is, in my view, greater and therefore should get more MV - you know that you are going to get a PHD to do your work if you go there - I'm not convinced that the quantity view should prevail over the quality one - and with the vetting given to the current metric, the burden should be on those advocating a change to convince the group of the need to change - cheers, Matt

-----Original Message-----

From: Lawrence C. Schuette
To: DeYoung, Don; Mleziva Matt Civ ESC/NI; Shaffer, Alan, Mr, OSD-ATL;
Goldstain Alan B Civ AEDC/CD; Higgins, Karen L SES; Blake Christopher L SES
HQ AFMC/XP; Berry, William, Dr, OSD-ATL; Simmons, Brian; Rohde, Robert S Dr
SAALT; Shah, Harshad C, Mr, OSD-ATL; Ryan, George R CIV; "Mathes, Thomas";
'schuette@nrl.navy.mil'
Cc: Buckstad, Robert, COL, OSD- ATL; Short, James, Dr, OSD-ATL; Strack Gary
Mr SAF/IEBB
Sent: 3/1/2004 2:01 PM
Subject: RE: People Metrics

Hi Don,

I found your arguments very persuasive. A wise man once said we have permission to get smarter... (or words to that effect). I believe that when the new version of the Annex comes out (later today), we'll all be able to take 4th or 5th look at the questions. Your analysis leads me to change my mind regarding the use of % versus raw numbers in the

People category.

Regarding the list of Societies - I thought we were going to provide points only if they were Technical fellows from that list of societies, not simply members.

Again, I look forward to seeing the current wording of the metrics, questions, and scoring.

I (apparently) am also in the minority on the issue of work done in-house and out-house. I saw Matt's comments regarding the flexibility (and thus higher value) that out-house work provides. I disagree. In my mind if the % numbers that are currently in use for in-house vs out-house \$\$ aren't approved, then we should only count the in-house FTEs. Additionally, we're not going to close, or move Raytheon or M.I.T. The effect is to inflating the value of those people. We are counting their experience, education, certifications and individual output. We're also capturing the location they work at and their collective output. We're also capturing the synergy that they have achieved with their local community and their customers and tech base.

That would seem sufficient.

v/r
Larry

-----Original Message-----

From: DeYoung, Don [mailto:deyoungd@ndu.edu]
Sent: Monday, March 01, 2004 1:34 PM
To: 'Mleziva Matt Civ ESC/NI'; 'Shaffer, Alan, Mr, OSD-ATL'; Goldstayn Alan B Civ AEDC/CD; 'Higgins, Karen L SES'; Blake Christopher L SES HQ AFMC/XP; Lawrence C. Schuette; Berry, William, Dr, OSD-ATL; Simmons, Brian; Rohde, Robert S Dr SAALT; Shah, Harshad C, Mr, OSD-ATL; Ryan, George R CIV; 'Mathes, Thomas' ; schuette@nrl.navy.mil
Cc: Buckstad, Robert, COL, OSD- ATL; Short, James, Dr, OSD-ATL; Strack Gary Mr SAF/IEBB; DeYoung, Don
Subject: People Metrics

DELIBERATIVE DOCUMENT - FOR DISCUSSION PURPOSES ONLY
DO NOT RELEASE UNDER FOIA

All,

I was asked to clarify my criticism of the Percentage approach as it applies to the People metrics. Since I'm not sure what was unclear, the following may not help, but I'll try.

The scenarios I used are hypothetical, but nonetheless very likely. In those test cases, the Percentage approach yielded what most would consider absurd, or at least irrational, results, for both the Education

and Awards metrics. By contrast, the Absolute numbers approach yielded rational results.

I believe this puts the burden of proof on the proponents of the Percentage approach to show that either (a) my math was wrong, (b) the Percentage approach's results in the test cases are, in fact, rational, or (c) the Absolute approach yields results, in likely scenarios, that are also irrational and would occur on a equally or more significant level. If neither (a), (b), or (c) can be supported, then the Percentage approach surely seems invalid.

So far, the only objection I've heard to using the Absolute approach is that it would mean "bigger is always better." Without providing evidence of why that is detrimental, that's just an assertion of what shouldn't be. Not liking an approach because it creates an end state that "shouldn't be" isn't very useful, and the Commission and losing communities are not likely to find the assertion very compelling. I understand most of the subgroup leads don't like the Absolute approach, but the key matter is the degree of reasonableness in the results.

But, maybe I'm missing something. It would be far from the first time. I'd be interested in knowing if either (a), (b), or (c) can be supported. If anyone can do so, it would be important because our job is to do the right thing by the DoD, the RDT&E sites, and the country.

v/r,

Don DeYoung

DELIBERATIVE DOCUMENT - FOR DISCUSSION PURPOSES ONLY
DO NOT RELEASE UNDER FOIA



INSPECTOR GENERAL
DEPARTMENT OF DEFENSE
400 ARMY NAVY DRIVE
ARLINGTON, VIRGINIA 22202-4704

April 22, 2004

MEMORANDUM FOR DIRECTOR, BASE REALIGNMENT AND CLOSURE
(INSTALLATION AND ENVIRONMENT)

SUBJECT: Office of the Inspector General of the Department of Defense Views on
Contractor Data

The Technical Joint Cross Service Group (TJCSG) has been holding discussions to determine whether contractor data is auditable. It is our understanding that the issue concerns whether the data on contractor education, experience, and certifications are auditable and certifiable.

“Transformation Through Base Realignment and Closure (BRAC 2005) Policy Memorandum One—Policy, Responsibilities, and Procedures,” April 16, 2003, states “Section 2903(c)(5) of BRAC requires specified DoD personnel to certify to the best of their knowledge and belief that information provided to the Secretary of Defense or the 2005 Commission concerning the realignment or closure of military installation is accurate and complete.” Policy Memo One further requires that all information used to develop and make recommendations must be certified. BRAC 2005 is a DoD process, not a contractor process.

The audit community has determined that contractor information obtained through e-mails or surveys is not acceptable supporting documentation because the information is not auditable. In addition, letters from contractor officials documenting personnel education, experience, and professional certifications is also unacceptable support because that is also not auditable. However, in some instances, if the contract used to obtain contractor services identifies education, experience, and professional certification requirements as a contract line item or a deliverable and the contractor personnel with those credentials correlates to a contractor invoice, then the information would be considered acceptable support as long as the contracting officer or contracting officer representative certifies it. The issue is that the acceptability of contractor data as support would be on a case-by-case basis and not across the board for all contractors, which, therefore, does not provide the TJCSG with consistently useable data.

A similar situation occurred in the capacity analysis data call regarding questions 690 and 691. The questions pertain to personnel education levels and Defense Acquisition Workforce Improvement Act certifications, respectively. The audit community is finding that the responses to questions are not properly supported nor are the installations able to obtain this information from their contractors.

**Draft Deliberative Document-For Discussion Purposes Only
Do Not Release Under FOIA**

In some cases, contractors are requesting to be paid for their data. These particular questions and issues will be highlighted in the OIG and Service audit agency reports that are related to the data validation efforts.

If you have questions, please contact Ms. Deborah Culp at (703) 604-9335.



David K. Steensma
Assistant Inspector General
for Contract Management



Close

From: DeYoung, Don
To: 'Blake Christopher L SES HQ AFMC/XP'; 'Strack Gary Mr SAF/IEBB'; "brian.simmons@dtc.army.mil'"
Cc: "Shaffer, Alan, Mr, OSD-ATL"; "George R CIV"; "Dr. Bob Rohde"; 'Goldstaysn Alan B Civ AEDC/CD'; 'Mleziva Matt Civ ESC/NI'
Subject: RE: Signed ISG Memo and proposed draft TJCSG response - BRAC FOUO
Sent: 4/23/2004 6:15 PM

Importance: Normal

Chris,

Thanks for your candid and, as usual, clear and useful feedback. Productive debate is always easier to achieve with that kind of input. I've appreciated your consistently even-handed role throughout this debate on how to score the People Metrics.

I still disagree with your proposal to put the Percentage Approach back into play, for reasons that you already know and understand. But, for the sake of the record -- and the fact that my thinking has evolved a bit more on this subject over the last few weeks -- I'll sketch out my reasons why the goals of the TJCSG, and the DoD, are best served by using the Absolute Numbers Approach, which was recently endorsed by the ISG.

First, I want to correct your assertion that I regard the Percentage Approach, and its use, to be irrational. I do not make that claim and never have. In fact, the approach itself is quite rational. However, I have claimed, and still believe, that use of the approach risks yielding some number of irrational results when we perform our closure scenario analyses.

Second, you link two issues – the contractor data and the Percentage Approach – and raise the issue of fairness. I believe you are saying that, if contractor data is excluded and the ISG continues to endorse the Absolute Numbers Approach, then the Air Force is being treated unfairly. I don't see it that way. Instead, each issue is being debated and evaluated independently, strictly on its own merits. To me, the "linkage" of separate issues is more a feature of political negotiation than it is of quantitative analysis.

My concerns about the Percentage Approach remain for the following reasons.

The Percentage Approach depresses (to varying degrees) the value of large masses of high quality talent (whether measured by the awards, education, or experience metrics) and exaggerates the value of small masses — across the board. Take the patently absurd case where the Percentage Approach gives a one-person site with one PhD the same military value score as a 1,000-person site with one thousand PhDs. Here it is obvious that the military value of the small mass is exaggerated and the military value of the large mass is minimized. I see this absurd case as the outer boundaries of the problem "box", but every case within the box will be affected to varying degrees. And it is likely that some as yet unknown number of them will yield irrational results.

The rudimentary sensitivity analyses I've performed varies all variables, the site populations along with the assessed qualities. As the Analysis Team itself acknowledged, in the case of the Awards/Patents Metric, they gave bases of all sizes roughly the same number of awards and patents.

To illustrate my concern regarding the high probability of irrational results, the following scenarios compare a small 100-person site versus a larger 500-person site (note that this is different from the 1,000-person example I've used on previous occasions).

Absolute Numbers Approach:

Site A: 100 S&Es (50 PhDs, 50 MS) MV = 0.12

Attachment F

Site B: 500 S&Es (300 PhDs, 100 MS, 50 BS, 20 AD, 30 No degree) MV = 1.0

Percentage Approach:

Site A: 100 S&Es (50 PhDs, 50 MS) MV = 1.0

Site B: 500 S&Es (300 PhDs, 100 MS, 50 BS, 20 AD, 30 No degree) MV = 0.92

Therefore, despite having disproportionately more PhDs (6 times vice 5 times) and 2 times the number of MS, Site B's military value is almost 10% less than Site A. This is irrational. Site A lacks the greater intellectual horsepower and sheer idea generation that comes with 250 more PhDs and 50 more MS, as well as the talent that 100 less degreed, and generally younger, individuals have to offer.

Proponents of the Percentage Approach will counter that the Absolute Numbers Approach gives (under the above example) Site B a MV that is almost ten times Site A, which they find excessive. However, this is more the function of the compressed point range. For example, if a PhD was worth 10 points, an MS worth 3 points, and a BS worth 1 point, the MV for Site A increases to 0.19. My suggestion of 2 months ago to change the point range as a way to address this concern never gained traction.

The problem grows more acute when scoring for the Awards/Honors:

Absolute Numbers Approach:

Site A: 100 S&Es

No Elite awards

No Prestigious awards

5 IEEE Fellows

30 unidentified patents of unknown value in last 3 years

MV = 0.21

Site B: 500 S&Es

1 National Medal of Technology, 1 Nobel Prize, 1 Collier Trophy

1 Space Technology Hall of Famer

10 IEEE Fellows, 10 AAAS Fellows

60 unidentified patents of unknown value in last 3 years

MV = 1.0

Percentage Approach:

Site A: 100 S&Es

No Elite awards

No Prestigious awards

5 IEEE Fellows

30 unidentified patents of unknown value in last 3 years

MV = 1.0

Site B: 500 S&Es

1 National Medal of Technology, 1 Nobel Prize, 1 Collier Trophy

1 Space Technology Hall of Famer

10 IEEE Fellows, 10 AAAS Fellows

60 unidentified patents of unknown value in last 3 years

MV = 0.94

These results are irrational. A site with only 5 fellows and 30 unidentified patents of unknown value scores 6% higher in military value than a site with a Nobel Prize, a Collier Trophy, a National Medal of Technology, 20 fellows, a Space Technology Hall of Famer, and twice as many unidentified patents of unknown value. Bear in mind that Bill Gates, Steven Jobs, Admiral Grace Hopper, Norm Augustine, and David Packard all won the National Medal of Technology. The Nobel's value is self-explanatory. And the Navy and Air Force shared the Collier Trophy for inventing and developing the Global Positioning System. This would not pass the smell test with either the Commission or the communities.

Also, note how Site A, even under the Absolute Numbers Approach, scores more than 1/5 of the MV compared to Site B, which is proportionally five times larger, but far more technically accomplished. This is due to the point compression problem.

In the end, I maintain that the Absolute Numbers Approach would yield closure and realignment proposals that we can better defend to the Commission. This approach more effectively gains the necessary differentiation to properly value the Department's intellectual capital. This is a vital key to identifying the most innovative sites, and it is critical to the credibility of the BRAC process.

I believe some of this rationale is why the TJCSG chose the Absolute Numbers Approach on 1 April, and why the ISG later endorsed that decision.

On the other hand, if we fail to take this approach, we will jeopardize the success of our entire endeavor. Any community-hired BRAC consultant worth his or her salt will zero in on an irrational result as a means of discrediting the study. But either way, whatever approach we end up using, our compressed point values are vulnerable to dissection and ridicule by anyone with the statistical aptitude of an avid baseball fan and an understanding of the non-linear difference between a PhD and a BS, and a Nobel Prize and an unidentified software license.

Again, way too many words, but hopefully I've made a clearer and more logical case. We'll probably continue to disagree on this issue. Regardless of the outcome to this debate, I'll continue to respect both you and the integrity and honesty you've displayed throughout this unpleasant, disagreeable, but necessary process.

v/r,

Don

-----Original Message-----

From: Blake Christopher L SES HQ AFMC/XP

To: DeYoung, Don; Strack Gary Mr SAF/IEBB; Blake Christopher L SES HQ AFMC/XP;

'brian.simmons@dtc.army.mil'

Cc: 'Shaffer, Alan, Mr, OSD-ATL'; 'George R CIV'; "Dr. Bob Rohde"; Goldstain Alan B Civ AEDC/CD; Mleziva Matt Civ ESC/NI

Sent: 4/23/2004 3:52 PM

Subject: RE: Signed ISG Memo and proposed draft TJCSG response - BRAC FOUO

Don,

I remind you that at the AT outbrief of the sensitivity results, I asked whether the AT saw anything unexpected or irrational in the outcomes of their trials in this area. Pete made it clear they saw nothing they didn't expect to see. The only time the word irrational has been used has been by you and by the formal Navy inputs to our ISG report. Don't take that as a slap of any form, just attributing the phase to the source. I've heard no one else refer to the use of percents as irrational. As I explained in my short, but candid, note to you on your analysis, I don't consider a small but very wise workforce to be of lower Mil Value than a large generally knowledgeable one. Where you see irrational, I see logical and acceptable.

What is interesting about this debate is that it is completely service neutral, until contractor data is excluded. As peers, I believe we have made some major headway in advancing TJCSG efforts, by agreeing and disagreeing on matters. This is one of those disagreements. As I've said many times, I can accept either approach, but not because one is rational and the other isn't. They both have merits and faults. However, when a major part of the AF workforce is excluded from consideration, then I put my AF hat on and push back and am forced to normalize the size matter out of the debate, using percentages. I think it is not only rational, but FAIR.

With the respect of a peer,

Chris
Christopher L. Blake - SES
Associate Director - Plans and Programs
AFMC/XP
937-656-0308 Office
937-603-0576 Cell

-----Original Message-----

From: DeYoung, Don [mailto:deyoungd@ndu.edu]
Sent: Thursday, April 22, 2004 5:54 PM
To: 'Strack Gary Mr SAF/IEBB'; 'Blake Christopher L SES HQ AFMC/XP';
'brian.simmons@dtc.army.mil'
Cc: 'Shaffer, Alan, Mr, OSD-ATL'; 'George R CIV'; "Dr. Bob Rohde";
DeYoung, Don
Subject: RE: Signed ISG Memo and proposed draft TJCSG response - BRAC
FOUO

Gary,

The statement that, "our preliminary TJCSG sensitivity analysis...did

not produce irrational results..." is misleading. I've attached slide #6 provided by the Analysis Team for the TJCSG meeting chaired by Mr. Erb on 1 April.

If I recall correctly, the TJCSG principals decided to use absolute numbers because of the radical inversion of results between the small and large sites for both S&T and D&A. Note especially the 23 March S&T results (actuals and percent).

The Principals found this shift significant enough to convince them that the percentage approach was not the way to go.

It should also be stressed that for the PPA metric, all sites had roughly the same number of awards, patents, and publications. Only the site populations varied significantly. The scenarios I constructed, a month or so ago, varied the awards/patents along with the site populations. It was by doing so that irrational results occurred.

Finally, another problem with the percent approach that doesn't get talked about much is the compression of the results (due to the minimal range in point values). The attached slide demonstrates that compression. It's a problem because our ability to differentiate one site's talent from another is reduced. If all sites look alike people-wise, then I guess any scenario we run is as good or bad as the other when it comes to intellectual capital. That's a set up for some potentially big mistakes.

v/r,

Don

-----Original Message-----

From: Strack Gary Mr SAF/IEBB
To: Blake Christopher L SES HQ AFMC/XP; DeYoung, Don;
brian.simmons@dtc.army.mil
Sent: 4/22/2004 10:09 AM