B·O·S·C
BOARD OF SCIENTIFIC COUNSELORS

COMPUTATIONAL TOXICOLOGY SUBCOMMITTEE

Conference Call Summary
Wednesday, October 21, 2009
1:00 – 3:00 p.m. Eastern Time

Welcome and Overview
Dr. George Daston, The Procter & Gamble Company, Subcommittee Chair

Dr. George Daston, Chair of the Board of Scientific Counselors (BOSC) Computational Toxicology Subcommittee, welcomed the Subcommittee members to the teleconference. He then took roll (the list of participants is attached to this summary) and explained that the purpose of the conference call was to discuss the draft letter report prepared by the Subcommittee (the draft report is attached to this summary). He thanked the Subcommittee members for their diligence in creating a high-quality draft report so promptly. He thanked Ms. Lorelei Kowalski for her help as the Designated Federal Officer (DFO) and stated that this would be her last action before leaving her position at the U.S. Environmental Protection Agency (EPA) to take one at the U.S. General Services Administration.

Designated Federal Officer Remarks
Ms. Lorelei Kowalski, EPA/Office of Research and Development (ORD), Subcommittee Designated Federal Officer (DFO)

Ms. Kowalski, Subcommittee DFO, reviewed the Federal Advisory Committee Act (FACA) procedures that are required for all BOSC Subcommittee meetings. As the DFO, Ms. Kowalski ensures that all FACA requirements are met and that records of Board deliberations are made available to the public. She mentioned that Ms. Heather Drumm was participating on the teleconference because she will be assuming DFO duties for the Subcommittee following Ms. Kowalski's departure from EPA.

The minutes were recorded by a contractor, Ms. Kristen LeBaron of The Scientific Consulting Group, Inc., and following review by the Subcommittee members and certification by the Chair, the summary will be available on the BOSC Web Site. A notice of all public meetings of the Subcommittee must be published in the *Federal Register* at least 15 days prior to the meeting and an electronic docket was established for this conference call. The electronic docket can be accessed at http://www.regulations.gov; the docket number is EPA-HQ-ORD-2009-0688. Regarding financial conflict of interest, Ms. Kowalski worked with officials to ensure that all ethics requirements were met satisfactorily. If a Subcommittee member discovers a conflict of interest during the Subcommittee's deliberations, the individual must notify the DFO immediately. This teleconference was convened to discuss the draft letter report, which the Subcommittee members should have received via e-mail on Friday, October 16, 2009. The goal is for the Subcommittee's letter report to be vetted at the December 2009 teleconference of the BOSC Executive Committee. Although there were no advance requests to comment from the public, an opportunity for public comment was provided at 1:45 p.m.

Subcommittee Draft Letter Report

Dr. George Daston, The Procter & Gamble Company, Subcommittee Chair

Dr. Daston explained that the Subcommittee would discuss the draft letter report by examining the response to each charge question. He asked the lead writers on each charge question to highlight the main points. He stated that the recommendations within the text need to be extracted as bulleted items but there is no required number of recommendations for each charge question. He noted that the introductory section of the report contains some historical background.

Drs. M. Moiz Mumtaz and Cynthia Stokes were the lead writers for Charge Question 1. Dr. Stokes explained that the response included a description of the Computational Toxicology Research Program (CTRP) Long-Term Goal (LTG) and general conclusions regarding the progress the Program has made in attaining this goal. The Subcommittee found that the CTRP has made appropriate progress toward meeting its goals under LTG 1. The Program has not met all of its goals, but the goals were very ambitious and the Subcommittee members did not expect all of the goals to be met by this point. The response includes specific examples of tools and databases that the Program has released to support decision-making. The Program has gathered vast amounts of data, and the overall consensus was that the Program has made successful progress toward meeting its ambitious goals. Dr. Stokes added that pages 3 and 4 of the report include additional ideas from the Subcommittee members on how the CTRP can move forward.

Dr. Daston noted the following three recommendations found within the text that need to be expanded and placed in bulleted format:

(1) "These projects need to continue to build on things that are in place, drilling deeper into the data and continuing the problem of structuring, standardizing, and organizing the data so that [they] can be more easily subjected to comprehensive meta-analyses." (Page 3, line 39)

(2) "In this context, it might be worthwhile to obtain some public feedback on how people may interpret the available data." (Page 3, line 46)

(3) "This could be achieved through an annual or biannual conference by bringing together the data generators, the data users, and the risk assessors/managers, the ultimate users of these alternative methods/models." (Page 4, line 33)

Dr. Stokes agreed that these were the three specific recommendations in this section; the remaining items are general comments/concerns that could be developed into specific recommendations if necessary. Dr. Mumtaz added that the concern regarding associations and the need to rigorously examine them (page 4, line 39) should be included as another recommendation. He agreed that the Program is ambitious, but it is making substantial progress. Dr. Stokes stated that the paragraphs on page 4 that begin on lines 4 and 13 discuss the need for the CTRP to engage with stakeholders and consumers; recommendations found within these two paragraphs can be combined with others in later charge questions. Dr. Daston noted that this related to the recommendation under Charge Question 5 regarding continuing dialogue with practitioners.

Dr. Robert Kavlock offered clarifying comments regarding the response to Charge Question 1. Referring to page 2, line 37, he noted that ExpoCast™ has not been released.

Dr. James Clark stated that the Subcommittee's recommendations must be clear and specific regarding what the members would like the CTRP to achieve. Dr. Daston agreed that they need to be specific without being too prescriptive. For example, in terms of recommending a conference, there are many ways to bring together data generators and users and risk assessors/managers; a conference is only one method by which to achieve this. The recommendation should be that a formal process be put in place to

bring these stakeholders together, not necessarily through a conference. Drs. Stokes and Mumtaz agreed to clarify the recommendations for the response to Charge Question 1.

Drs. Dennis Paustenbach and Santiago Schnell were the lead writers for Charge Question 2. Dr. Paustenbach noted that the Program has used its funds appropriately to mine data. The overarching recommendation is that when EPA data are completely integrated with data from other agencies, the Program should work closely with academia to determine what data are important. The U.S. Department of Energy and some universities have begun this process, and EPA can work with them to identify what is important. Additionally, the Program should advertise its efforts so that current and future work is aligned among the Agency, academia, and international organizations to create synergy. EPA should work with the European Union's (EU) Registration, Evaluation, Authorisation and Restriction of Chemical Substances (REACH) Program to mine EU data. The Subcommittee response to the charge question also deals with efficiency, although it is a challenge to determine the true level of efficiency because of the various methods of funding and cooperative agreements. Dr. Schnell added that Dr. Lawrence Hunter's comments were incorporated into the response to this charge question. Referring to the National Institutes of Health's (NIH) successful multiscale modeling project, Dr. Schnell noted that it would be a good idea for the CTRP to link with this and other programs, such as virtual human tissue projects.

Dr. Daston summarized that the two specific recommendations under this charge question are to: (1) continue to seek and expand federal and international partnerships regarding model development and additional data streams (e.g., REACH); and (2) provide a graphical presentation of projects that addresses the efficiency and transparency of funding and full-time equivalents (FTEs), as described in the paragraph on page 6 beginning on line 9. Dr. Clark expressed some reservations regarding the second recommendation. The paragraph describes the data that the Subcommittee would like to obtain to assess efficiency; however, how these data will be used to determine and/or address efficiency is unclear. Dr. Paustenbach noted that oversight and review committees such as the BOSC are asked to evaluate efficiency, and he acknowledged that this is a difficult question to answer. The proper tools have not been identified, and this recommendation was his best attempt at getting some information that might be helpful to the BOSC in assessing efficiency. Dr. Clark stated that it is difficult to determine what constitutes a good value, particularly in a cutting-edge area such as computational toxicology. Using FTEs might be one means of determining efficiency, but ultimately it is a value judgment. Ms. Kowalski added that the BOSC Executive Committee is investigating efficiency via the program reviews, but this has not been incorporated into the standing subcommittee reviews such as this one, which differ from the program reviews.

Dr. Paustenbach commented that it would be useful for EPA and other agencies to have the tools necessary to evaluate efficiency. Dr. Daston responded that this has been discussed at the Executive Committee level, but it is not fair to place this responsibility on the CTRP. He asked whether the paragraph requesting the table (page 6, line 9) should be removed. Dr. Schnell stated that this information might shed some light on how the Program is effectively using its partnerships. Dr. Stokes commented that there is a recommendation under Charge Question 5 regarding establishing performance metrics (page 12, line 28), which appears to be related to the point that is being discussed under Charge Question 2; they could be combined, and the table may be helpful. Dr. Daston noted that there is complete transparency of funding and FTEs under the direct control of the Program, but evaluating the contributions of collaborations and external data streams is challenging. Dr. Paustenbach thought that it would be helpful for other groups within the Agency to have this attempt as an example of how to determine collaborative efficiency. Dr. Mumtaz said he thought that the purpose of the review was to determine whether the CTRP is doing a good job.

Public Comment

Ms. Kowalski called for public comments at 1:45 p.m. No public comments were offered so the discussion of the report resumed.

Subcommittee Draft Letter Report (continued)

Dr. George Daston, The Procter & Gamble Company, Subcommittee Chair

Dr. Daston asked Dr. Kavlock whether the recommendation to provide a table of collaborative outcomes would be an actionable item that Program management could effectively implement. Dr. Kavlock responded that the FTEs and overall funding levels were provided to the Subcommittee, but it is difficult to monetize the contributions of contractors. Perhaps the number of Memoranda of Understanding, Material Transfer Agreements, and Interagency Agreements for the Program could be compared to the number of such agreements entered into by other groups/programs.

Dr. Paustenbach commented that he is trying to develop a standard approach for reporting financial information so that the data can be compared inside and outside the Agency. Dr. Daston thought that such a recommendation would be more appropriate for the BOSC Executive Committee rather than for the CTRP.

Dr. Stokes commented that the third paragraph under Charge Question 2 discusses collaboration with the international scientific community. She would like the recommendation to be clarified and include mention that the Program should ensure that programming and modeling platforms are compatible. Dr. Schnell agreed that this is an important issue, but he did not think the Subcommittee should be too prescriptive in recommending a specific tool. Dr. Mumtaz added that technical personnel should be included in such collaborations. Dr. Daston summarized the discussion, stating that this section should be revised to make it clear that collaborations are not mandatory, but they are examples of other activities that will allow the Program to be aware of what is taking place internationally. Drs. Paustenbach and Schnell agreed to modify the language as discussed and prepare the recommendations for Charge Question 2.

Dr. Kavlock made the following clarifying statements regarding the response to Charge Question 2: (1) “Hubel” should be changed to Hubal in line 13 on page 5; (2) “DTRP” should be changed to CTRP in line 25 on page 5; (3) regarding line 28 on page 5, the CTRP does not collaborate directly with the National Institute of Biomedical Imaging and Bioengineering or the National Institute of General Medical Sciences; and (4) in terms of international collaboration, Drs. Kavlock and David Dix are visiting the EU to discuss working with the REACH Program.

Drs. Richard Di Giulio and Katrina Waters were the lead writers for responding to Charge Question 3. Dr. Di Giulio explained that overall they were impressed with the CTRP’s ability to integrate molecular biology and bioinformatics with high-throughput screening and the Program’s use of classic and virtual models. The Program is at the forefront of this integration and making successful progress regarding human health and the environment. Line 17 on page 7 highlights that it will be beneficial for the Program to consider how its efforts have been successful. The last full paragraph on page 7, beginning on line 39, discusses how the Program’s progress will be implemented at a regulatory level.

Dr. Ali Faqi said he thought that the CTRP is moving in the right direction to meet its goals; it is moving the field of toxicology from a descriptive state to a predictive one.

Dr. Waters noted that the sentence beginning on line 14 of page 7 on encouraging the Program to involve the statisticians and mathematical modelers in assay evaluation is a recommendation that should be highlighted. At the face-to-face meeting she saw a disconnect between the statisticians who have performed some of these evaluations and the researchers performing follow-up studies to investigate

mode of action, particularly toward the identification of new toxicity pathways. She was not aware of a defined strategy for assay evaluation or handling attribution of assays that did not contribute useful information. The Program needs to identify new toxicity pathways that lead to the creation of new assays, which will help build the databases to move toward quantitative prediction.

Dr. Daston summarized the recommendations within this charge question, one of which involves metrics, one a close collaboration with program offices and other risk assessors to develop case studies to demonstrate the applicability of methods and data streams, and one to develop a strategy to evaluate assays. Dr. Mumtaz added that there is a recommendation to continue Program outreach and publication efforts.

Dr. Stokes agreed with the concern that Dr. Waters expressed at the face-to-face meeting regarding why the Program chose certain assays rather than others. She recommended placing the specific recommendations that were just discussed with the discussion regarding the number of assays used on page 7 (line 4), so that the response is strengthened. Dr. Waters clarified that Dr. Stokes was suggesting placing recommendations from the second full paragraph on page 7 (beginning on line 10) in the first full paragraph on the same page (beginning on line 3).

Dr. Kavlock clarified that in line 4 of page 7, 467 assays from Phase I have been released thus far. In line 7 of the same page, in Phase II the CTRP is carrying out the actions that the Subcommittee just discussed, examining the assays from repeatability, interpretability, and cost viewpoints. The vendor sources that will be carried into Phase II will be determined, and new vendor sources also will be identified. Dr. Waters explained that her concern was regarding the statistical usability of the data. Dr. Dix responded that this is the exact type of analysis that the Program is performing.

Drs. Clark, Hunter, and John Quackenbush were the lead writers for Charge Question 4. Dr. Clark explained their strategy for developing the response and ensuring that the charge question was addressed. The charge question response includes the Subcommittee's findings, but not all of them are actionable items. The specific recommendations are described as such. He commended the training that the Program performs; there are many beneficial training opportunities arising from the CTRP. Dr. Clark was unsure whether the section on software quality assurance on page 10 is appropriate for a letter report. Dr. Quackenbush commended the Program on its synthesis and its training program, but he has concerns about the Program's reliance on the Science To Achieve Results (STAR) centers when there is no mechanism to renew these centers. The Program needs to consider its long-term plan in the absence of the STAR centers; this is the most significant challenge for long-term viability. Dr. Clark will expand this recommendation to include an acknowledgement of what the Program would do without the STAR centers. Dr. Kavlock clarified that the funding for the STAR centers comes directly from the CTRP's budget, and the Program has the ability to release Requests for Applications (RFAs) to solicit STAR centers that will perform research tailored to the Program's needs. The caveat is that current STAR centers cannot be re-funded, and they may not qualify to compete in future RFAs if the RFA is not relevant to their work.

Dr. Kavlock noted that the Program is hiring a communication specialist, but this was not mentioned in the draft letter report. Dr. Daston agreed that this should be added to the response to Charge Question 5.

Dr. Stokes suggested mentioning models in the section on software quality assurance and changing the title to "Software and Modeling Quality Assurance." The ideas contained in this section are good, but the section should be broadened to include modeling. Dr. Clark agreed to make these changes.

In terms of Charge Question 5, the Subcommittee agreed that the National Computational Toxicology Center should continue as an established Center within the Program. Dr. Stokes reminded the Subcommittee of the plan to incorporate into Charge Question 5 the recommendation from Charge

Question 1 regarding engaging the community, clients, and stakeholders on an ongoing basis. Dr. Daston will incorporate this recommendation into Charge Question 5.

Dr. Hunter noted the importance of computational approaches to exposure and read the verbiage that he would like included in the draft letter report. Dr. Daston commented that this could be included in the response to Charge Question 4. Dr. Daston will send the paragraph to Dr. Clark to incorporate within the response to that charge question.

Final Draft Letter Report: Next Steps
Dr. George Daston, The Procter & Gamble Company, Subcommittee Chair

Dr. Daston stated that the lead writers for each charge question will make the changes discussed during the conference call. The recommendations should be articulated in one or two sentences. The revised responses should be sent to Dr. Daston and Ms. Drumm via e-mail no later than Wednesday, October 28, 2009. Ms. Kowalski noted that the next step then would be to send the draft letter report to the contractor for final formatting and editing before it is sent to the BOSC Executive Committee. She thanked the members of the Subcommittee for their efforts during the past few years and remarked that it was a pleasure working with them. The Subcommittee members thanked Ms. Kowalski for her efforts with the BOSC and wished her well in her new endeavor.

Dr. Daston thanked everyone for their participation and adjourned the meeting at 2:34 p.m.

Action Items

- ✧ Drs. Stokes and Mumtaz will write the recommendations for Charge Question 1.
- ✧ Drs. Paustenbach and Schnell will modify the language as discussed and prepare the recommendations for Charge Question 2.
- ✧ Dr. Waters will move the recommendations from the second full paragraph on page 7 (beginning on line 10) to the first full paragraph on the same page (beginning on line 3).
- ✧ Dr. Clark will expand the recommendation regarding STAR centers to include an acknowledgement of what the Program would do without the STAR centers.
- ✧ Dr. Clark will add mentions of modeling where appropriate to the section on software quality assurance and change the title to “Software and Modeling Quality Assurance.”
- ✧ Dr. Daston will incorporate the recommendation from Charge Question 1 regarding engaging the community, clients, and stakeholders on an ongoing basis into Charge Question 5.
- ✧ Dr. Daston will send Dr. Hunter’s paragraph regarding exposure to Dr. Clark to incorporate within the response to Charge Question 4.
- ✧ All lead writers will send their revised sections to Dr. Daston and Ms. Drumm no later than Wednesday, October 28, 2009.

PARTICIPANTS LIST

Subcommittee Members

George P. Daston, Ph.D., Chair
Research Fellow
The Proctor & Gamble Company
Miami Valley Laboratories

James R. Clark, Ph.D., Vice-Chair
Distinguished Scientific Associate
ExxonMobil Research and Engineering Company

Richard T. Di Giulio, Ph.D.
Professor
Nicholas School of the Environment and Earth Sciences
Duke University

Ali Faqi, D.V.M., Ph.D., DABT
Director, Developmental and Reproductive Toxicology
MPI Research, Inc.

Lawrence Hunter, Ph.D.
Director
Center for Computational Pharmacology and Computational Bioscience Program
University of Colorado

M. Moiz Mumtaz, Ph.D.
Science Advisor
Division of Toxicology and Environmental Medicine
Agency for Toxic Substances and Disease Registry

Dennis Paustenbach, Ph.D., CIH, DABT
President and Founder
ChemRisk, Inc.

John Quackenbush, Ph.D.
Professor of Biostatistics and Computational Biology
Department of Biostatistics and Computational Biology
Dana-Farber Cancer Institute

Santiago Schnell, Ph.D.
Associate Professor of Molecular and Integrative Physiology
Brehm Investigator, Brehm Center for Type 1 Diabetes Research and Analysis
Research Associate Professor of Computational Medicine and Bioinformatics
University of Michigan Medical School

Cynthia Stokes, Ph.D.
Independent Contractor

Katrina Waters, Ph.D.
Computational Biology & Bioinformatics Group
Battelle

Designated Federal Officers

Heather Drumm
U.S. Environmental Protection Agency
Office of Research and Development

Lorelei Kowalski
U.S. Environmental Protection Agency
Office of Research and Development

EPA Participants

Karen Dean
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Computational Toxicology

David Dix, Ph.D.
U.S. Environmental Protection Agency
Office of Research and Development
Acting Deputy Director, National Center for Computational Toxicology

Richard Judson, Ph.D.
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Computational Toxicology

Robert Kavlock, Ph.D.
U.S. Environmental Protection Agency
Office of Research and Development
Director, National Center for Computational Toxicology

Other Participants

Liz Buckley
Pesticide & Toxic Chemical News

Paul LaBell
American Society of Agronomy

Aaron Lovell
Inside EPA

Contractor Support

Kristen LeBaron, M.S.
The Scientific Consulting Group, Inc.

B·O·S·C
BOARD OF SCIENTIFIC COUNSELORS

COMPUTATIONAL TOXICOLOGY SUBCOMMITTEE
AGENDA
Wednesday, October 21, 2009
1:00 p.m. – 3:00 p.m. eastern time

CONFERENCE CALL
Participation by Teleconference Only

1:00–1:05 p.m.	Welcome and Overview - Purpose of Teleconference Call	Dr. George Daston, Chair Computational Toxicology Subcommittee
1:05–1:10 p.m.	DFO Remarks	Lori Kowalski, Office of Research and Development
1:10–1:45 p.m.	Subcommittee Draft Letter Report - Overview - Draft responses to charge questions - Discussion	Dr. George Daston, Chair, Computational Toxicology
1:45–2:00 p.m.	Public Comment	
2:00–2:45 p.m.	Subcommittee Draft Letter Report (Cont.) - Discussion	Dr. George Daston, Chair, Computational Toxicology Subcommittee
2:45-3:00 p.m.	Final Draft Letter Report - Next steps	Dr. George Daston, Chair, Computational Toxicology Subcommittee
3:00 p.m.	Adjourn	

B·O·S·C
BOARD OF SCIENTIFIC COUNSELORS

October 16, 2009

Lek Kadeli
Acting Assistant Administrator
Office of Research and Development
U.S. Environmental Protection Agency

Dr. Robert Kavlock
Director
National Center for Computational Toxicology
U.S. Environmental Protection Agency

Dear Mr. Kadeli and Dr. Kavlock:

This is a letter report from the Board of Scientific Counselors (BOSC) reviewing the Office of Research and Development's (ORD) Computational Toxicology Research Program (CTRP). The Computational Toxicology Subcommittee reviewed CTRP's progress and future plans during a conference call on September 25, 2009, and a 2-day meeting held September 29-30, 2009, in Research Triangle Park, North Carolina. The BOSC Subcommittee consists of George Daston (Chair), James Clark, Richard DiGiulio, Ali Faqi, Lawrence Hunter, Moiz Mumtaz, Dennis Paustenbach, John Quackenbush, Santiago Schnell, Cynthia Stokes and Katrina Waters.

This is the fourth review of the CTRP conducted by the BOSC. NCCT first became operational in February 2005; during the 4.5 years between its establishment and this review, the CTRP has made substantial progress in establishing and meeting priorities and goals; collaborating within and outside EPA to leverage the staff's expertise; and transforming the field of toxicity testing. Many of the recommendations made by BOSC during its earlier reviews have been acted on by CTRP. Add highlights from 2008 review here This includes improved capabilities in bioinformatics through the funding of two external centers and in informatics and systems biology through staff hires; expansion of its technical approaches to even more programs within the Agency; and the formation of an extensive collaboration with the National Institute of Environmental Health Sciences (NIEHS) and the National Human Genome Research Institute (NHGRI) for its ToxCast project.

The purpose of the September 2009 review was to provide the CTRP with advice on (1) the progress the Center has made, in the past four and a half

years, in fulfilling its mission and strategic goals; and (2) whether the NCCT should continue as an established organization beyond its original five-year charter. In particular, the BOSC addressed five charge questions that focused on the progress and future of the NCCT. The BOSC's responses to these questions follow.

Charge Question 1: *What is your evaluation of the progress the Computational Toxicology Research Program (CTRP) has made in achieving its original goals and objectives, and whether it has efficiently utilized available resources?*

The mission statement of the CTRP is to integrate modern computing and information technologies with molecular biology to provide the Agency with decision support tools for high-throughput risk assessment.

The three initial long-term goals of the CTRP were as follows:

- Risk assessors use improved methods and tools to better understand and describe the linkages of the source-to-outcome paradigm,
- EPA Program Offices use advanced hazard characterization tools to prioritize and screen chemicals for toxicological evaluation, and
- EPA assessors and regulators use new and improved methods and models based on the latest science for enhanced dose-response assessment and quantitative risk assessment.

These are ambitious goals particularly the last one, the objective of which is to identify non-classical methods for assigning an approximate acute and chronic toxicological hazard rating to hundreds, if not thousands, of chemicals. The vast majority of these ratings are to be conducted without the traditional reliance on animal testing and human epidemiology.

The BOSC believes that the CTRP has made substantial progress toward meeting the original long-term goals, and that the progress is appropriate given the duration of the program's existence and the resources involved. While the goals have not been fully met, the CTRP has made significant advancements to creating and providing the methods, tools, and models that will enable risk assessors, program offices and regulators to use 21st century science and technology for their work, as indicated in the goals. Notably, a number of new tools have already been released to support decision-making, including AcToR, DSSTox, ToxRefDB, and ExpoCast.

Underpinning all of these, the CTRP has focused on integrating modern computational approaches with molecular and cellular biology and physiology to create new methods for the agency's chemical prioritization and risk assessment efforts. The activities of the CTRP in pursuit of these goals have included the assembly and integration of vast quantities of existing toxicological and toxicogenomics data; creation of new database and data warehousing tools to house the data; development of methods and tools for discerning actionable knowledge from the data; additional data acquisition on chemicals of agency interest; and the development of various types of computational models to understand biological and toxicological mechanisms and

provide predictive tools for hazard evaluation and prioritization and risk assessment. As several of CTRP's projects have matured, they have started to link together in ways that allow ToxCast and Tox21 data to be effectively captured and managed, and allow the various sources of information to be leveraged against each other in productive ways.

Overall, during its inaugural funding period CTRP has built the infrastructure necessary to bring computational tools to risk assessment, assembling the data and building the tools that are needed to collect the high-throughput screening data that the program is now facing. In the process, the program has learned the limitations of the existing data and has begun to explore how it might address the limitations of the existing systems, data, and models that will be necessary to move forward. Issues of bioavailability and bioactivity; correlations between in vitro and in vivo results; dose response relationship and toxicity are case in point.

The BOSC concludes that the CTRP has laid a strong foundation and has put forth good ideas to move forward, and that there is clearly a need to continue to build on the success of the program.

Because of the large number of projects and collaborations involved, it is difficult to assess in detail whether the CTRP has utilized the available resources efficiently. Nonetheless, it is the conclusion of the BOSC that the progress made is substantial and appropriate in light of the time the program has been in place, its budget, the number of personnel involved, and the collaborations in place within and outside the agency.

As expected with such an ambitious program, the CTRP has been faced with a number of challenges as they have strived to meet their goals. The following are the BOSC's observations about some of these challenges along with suggestions for meeting them that are relevant to continued progress.

One of the challenges that the CTRP has taken on is the assembly and integration of the vast quantities of existing, available toxicological and toxicogenomics data. Although it might seem that this should be a relatively easy problem, much of the data are not electronic, much is poorly annotated, and many ambiguities exist within the available data. The CRTP has initiated a number of data base and data warehousing projects to begin to address these issues, including AcToR, DSSTox, ToxRefDB, and ExpoCast, each of which approaches a different aspect of the problem. Each of these has made significant progress over the past few years, along the way discovering the need to do a significant amount of manual curation, but using existing tools and developing new ones as appropriate, and beginning to link together in productive ways. Despite all that has been achieved, much of the work is clearly in an early stage, and many of the current resources are more like a structured index to the underlying data rather than a comprehensive linked data resource. These projects need to continue to build on things that are in place, drilling deeper into the data and continuing the problem of structuring, standardizing, and organizing the data so that it can be more easily subjected to comprehensive meta-analyses. In addition, much of what is available is focused on research scientists and the data is presented "as is" without context. The potential problem is that as these data resources are publicly available, the public may access the information and, without context, may misinterpret what is there. For example, the data in ACToR identifies compounds that have been tested for carcinogenicity or genotoxicity, but not those that have been found to be carcinogenic or genotoxic. In this context,

it might be worthwhile to obtain some public feedback on how people may interpret the available data.

Although database development is seen as a fairly mundane task, it is the foundation on which much of the CTRP's program is being built and doing it successfully requires a significant intellectual investment. Fundamentally, a database is a model. And in building the model, it requires two components: understanding the relationship between the data elements and understanding how people will use the data. In many ways, this program is in a critical phase in that, having assembled the data into one place, building the linkages in a systematic fashion now will require understanding the value of data to the community and the manner in which those data will be used.

A large part of the success of these programs relies on the development of computational models to interpret data and make predictions. A major part of the modeling effort focuses on interrogating the databases. The subcommittee noted that a substantial part of these efforts utilize machine-learning methods. While the BOSC has no objection specifically to such approaches, it should be noted that many biostatisticians are sometimes apprehensive about such methods. Since many of the ultimate customer clients and stakeholders of the CTRPs efforts are expected to have biostatistical backgrounds, the BOSC encourages the CTRP to consult with biostatisticians early and often to assure they can meet any objections, as well as possibly attempt some additional methods.

There is also a need to interact more extensively with the broader scientific user community in the process of developing and rolling out tools and software. One of the driving principles behind the program is that the resources it produces should be both useful and used. For the resources to be useful requires that the resources address important questions and, to a large extent, it seems that there is a strong focus on addressing relevant problems in environmental toxicology and exposure and risk assessment. For the resources to be used will require understanding how risk assessors outside the CTRP and within the broader toxicological and toxicogenomics community will use these resources. This process should allow the development of appropriate use cases that can guide how the tools are created and integrated and if this were done in a systematic manner, it has the potential to rapidly advance the evolution of the resources the CTRP is developing. This could be achieved through an annual or biannual conference by bringing together the data generators, the data users and the risk assessors\managers, the ultimate users of these alternative methods\models.

There were concerns expressed by some subcommittee members that associations are not causation and this should be recognized by the EPA management, both at the CTRP level and at the level of the office of the administrator. The results of a computer generated association should be carefully examined through traditional testing and careful scientific examination. Some of the members recalled that the Ames test was once thought to serve in a similar manner that is being proposed by the various programs that are being built by the CTRP. Thus, although the BOSC fully supports providing more resources to the CTRP efforts, it offers the precautionary warning that at the best the results of these efforts will be the temporary placement of a chemical into a bin that could likely initiate a "science forcing event". In turn, the manufacturer or user of the chemical can be put on notice that this chemical appears to have

certain characteristics that give it a likelihood for being a hazard and that they will need to conduct further toxicological testing. This will bring the chemical industry on par with the pharmaceutical industry that has been conducting such cost/risk/benefit analyses for several decades.

Charge Question 2: *To what extent and how effectively has the CTRP utilized internal and external partnerships to foster its goals?*

It is clear that the CTRP has been effective at finding professional colleagues in various institutions with whom to collaborate. Certainly, they have been able to identify various research activities and data sets through EPA and other organizations in an attempt to assemble sufficient information to achieve some of their goals. For example, the exposure assessment work by Hubel et al is an example of a group who has a good idea about where various data sets reside in the government and they are committed to make more of it available. When those data are gathered, they will be accessed by other groups in the CTRP to combine with data they have on toxicity or predictors of toxicity.

It appears that CTRP has successfully engaged those in NCCT, NHEERL, NERL, NRMRL, NCEA, and other groups within the government. However, there are two major government agencies, the National Science Foundation and the Department of Energy, which support research work relevant to the Computational Toxicology program, but do not appear to be CTRP partners. The National Science Foundation has basic science programs, which are relevant not only to the biological sciences, but also in computer science and in the area of qualitative research into the activities (and information needs) of scientists. These are essential components of the DTRP research. The Department of Energy programs and data, particularly in radiation safety and environmental remediation would seem natural sources of valuable information for the Computational Toxicology program. In addition, the collaboration with NIH appears to be mediated by NIEHS, NIBIB and NIGSM. These agencies are currently funding a predictive Multi-scale Models of the Physiome in Health and Disease. Also, a more formal relationship can be established with the National Library of Medicine, particularly with the National Center for Biotechnology Information and its PubChem program. The BOSC suggests the development of these relationships would be appropriate for the CTRP.

It appears that CTRP has been less effective in developing equally strong research groups at various universities in the United States and internationally, and or other scientific and regulatory bodies outside the united states (for example, the UK, Australia, NZ, Switzerland, Germany, and a few other countries have substantial databases on exposure, toxicology, and predictive toxicology). No doubt, these relationships will follow as CTRP matures and receives additional funding. At the moment, CTRP is establishing collaborations in multi-scale modeling of developmental toxicity and virtual tissues with some US academic partners (Indiana University, University of Texas). While these collaborations are clearly of value, it also seems appropriate to expand these partnerships to include multi-scale modeling work in Europe, Asia, and elsewhere in the USA. An excellent example is the virtual physiological human initiative that is intended to support the development of patient-specific computer models and their applications in personalized and predictive medicine. This constitutes an integral part of the international Physiome Project, a worldwide public domain effort to develop a computational framework for

quantitative description of biological processes in living systems across all relevant levels of structural and functional integration, from molecule to organism, including the human. The Physiome Project has established standards, which are widely used in the multi-scale modeling community. In parallel, the mathematical and computational oncology community has established similar initiates for assessing cancer progression and treatments. CTRP will benefit substantially in establishing partnerships with these multi-scale modeling enterprises in the US, Europe and Asia.

As an oversight group, we would very much like to have seen a table which presented the various relationships with the FTEs from each organization committed to a particular "joint" collaboration. In that table, we would like to have seen an indication of the level of financial resources from CTRP that were dedicated to the various projects, and a timeline for various milestones. Like so many government initiatives, it is very difficult to rate the efficiency or productivity of these programs because it is unclear if the CTRP is devoting 0.25 FTE for 3 years to the initiative and NHEERL is devoting 5 FTE for 4 years; and thus it is predominantly the result of NHEERL resources that makes the project a success. Another layer of complexity occurs when some portion of the projects is conducted by post-docs, summer or more permanent interns, and graduate students (who may or may not be called FTEs). Then, beyond that, there are sub-contractors who often participate. Building such teams is not to be discouraged and, indeed, they are to be promoted. But it is not possible to determine which person or persons is "leading the project" and providing the driving force to resolution. If one were to interview each team, it would be clear who was truly providing the laboring ore.

Charge Question 3: *What evaluation can you provide relative to the contributions of the CTRP to the advancement of transforming the field of toxicity testing?*

The CTRP appears to be at the cutting edge of transforming the field of toxicity testing. Its original goals, defined at its creation in 2005, were highly consistent with recommendations subsequently described by the National Academy of Science in 2007 (Toxicology in the 21st Century: A Vision and a Strategy). Particularly important contributions in this area include major advances in high throughput screening (HTS), advances in approaches for data-mining from various data sources including HTS and other ToxCast efforts, advanced model development of virtual tissues, and the incorporation of uncertainty analysis into model development and, ultimately, risk assessments.

The incorporation of modern computing with molecular biology as developed by the CTRP is necessary to move the discipline of toxicology from the current stage, which is primarily descriptive science to a more predictive one. The concept of utilization cell-based in vitro testing will assist the understanding of the key biological pathways by which chemicals induce adverse effects. The development of knowledge bases of toxicity pathways, toxicological responses, and key information on biological networks will lead to the use of solid science in the risk assessment. The HTS screening will lead to a more cost effective testing; which will further save money and reduce the use of animal testing. In addition, virtual tissues (liver and embryo) that link across levels of biological organization from molecular to cellular to tissue level responses will be good predictive tools for general and developmental toxicity. The NCCT is establishing collaborations with other institutions across the world with similar goals, which may lead to the expansion of the number of predictive virtual tissues. Moreover, we expect the

attrition rate for pharmaceutical compounds to be reduced as the computational toxicology tools will provide a better prediction of human toxicity.

The CTRP has demonstrated rapid progress in the development of HTS. To date, during Phase I of ToxCast, 409 assays have been employed with nine platforms. These assays have been applied to an initial set of 309 chemicals for which relatively extensive toxicological information is available (largely active chemicals in pesticides). In Phase II, this number of assays has grown to 467 so far, which will be applied to a new set of 700 chemicals for which less information is available. This phase will provide for a critical evaluation of the large array of assays under consideration for their relative utility of chemical prioritization.

The BOSC applauds the efforts of the NCCT to embark on the challenges of the Tox21 paradigm, showing us the strengths and limitations of the 'toxicity pathway' approach and the challenges of generating truly predictive HTS platforms. The challenge for this group is to apply the lessons learned from the Phase I ToxCast efforts and iterate with Phase 2 to add new assays and define a strategy for attrition of those that provide limited or inconsistent information. We encourage them to keep the statisticians and mathematical modelers involved in assay evaluation so that they can move from qualitative prediction to quantitative prediction of outcomes from exposure data. To achieve their future milestones, the group needs to define metrics by which they can measure their success (such as specificity and sensitivity goals) and declare victory for specific classes of compounds or cell types. The identification of new toxicity pathways, and therefore new assays, will be essential to gain the predictive power necessary to predict outcome from exposure beyond a single class of compounds.

The initial evaluation of Endocrine Disruptor profiling with ToxCast data is extremely promising in its ability to identify new MOA that aren't traditionally considered in addition to known developmental or reproductive endpoints. Likewise, the ToxCast data display good sensitivity for predicting neoplastic liver lesions, and statistical models have demonstrated a data gap for new assays to be developed for non-neoplastic lesions. Although there is some suspicion of these statistical approaches because "they don't make sense based on what we know", there is value in an unbiased evaluation of the usefulness of particular assays. This is a huge step forward for NCCT to demonstrate the potential utility of the ToxCast approach, and it justifies continued funding to not lose the momentum of this research team.

The engagement of collaborators in NCEA to transform risk assessment into a 'NexGen' paradigm using tools and databases coming out of the NCCT is certainly impressive, although daunting. Translating the predictions from the HTS assays to human population risk will require strong connections between the MOA data and statistical genetic diversity, such as those being provided by the Carolina Center for Computational Toxicology, led by Dr Ivan Rusyn. The ability to put real uncertainty factors into exposure limits that protect 99% of the population will truly revolutionize human health risk assessment.

It is very difficult to determine how the work of CTRP will be implemented from a regulatory standpoint. Their work will likely help organizations identify the chemicals most deserving of significant study and the "type" of additional testing that needs to be conducted.

Ultimately, the CTRP research will establish a methodology which can be relied upon by

researchers around the world for quickly identifying those chemicals which "have a red flag". This could help significantly in optimizing the resources of the various organizations such that they will not be conducting routine toxicology "screening tests" on chemicals which raise too many red flags after being run through the ultimate program offered by CTRP.

Charge Question 4: *To what extent do the ORD intramural projects, the extramural STAR centers, and the five stated CTRP management priorities described in the FY09-12 implementation plan combine to efficiently support the goal of providing high throughput decision support tools for screening and assessing chemical exposure, hazard and risk to human health?*

Ultimately the implementation of the CTRP's goals is going to require an iterative approach to developing models as the data available grows in both its quantity and complexity. In order to do this, CTRP must develop methods that can be used to make verifiable predictions and to generate appropriate data to test those predictions. This is clearly going to require developing a strong partnership that reaches beyond the boundaries of the CTRP program and takes advantage of other existing programs, including ORD intramural projects, other intramural EPA projects, the extramural STAR centers, and partnerships with other agencies.

The FY09-12 implementation plan lays out these interactions and the role that they will play in helping to direct and develop the CTRP program. This document serves as a useful reference and guide as to the roles and activities of all parties contributing to the computational toxicology program. The new start programs outlined for the ORD laboratories and centers are well leveraged with the CTRP. Results from those projects have the potential to be quickly incorporated into ongoing CTRP activities, providing for effective and efficient use of R&D efforts. Continuing cooperation with the STAR centers and the innovations that are taking place at these university sites will enhance the development and robustness of the decision support tools under development. In total, NCCT interactions with other ORD laboratories and the STAR centers are essential as they will provide not only the starting data needed to fully develop the CTRP program, but also some sampling of the potential users of the systems that will be essential for providing feedback and ideas for the next iteration of tools.

In their implementation plan, the CTRP lays out five priority management areas:

1. Toxicity Predictions and Chemical Prioritizations Incorporating Exposure
2. Strengthening Cross-ORD Collaborations
3. Tox21: A Federal Partnership Transforming Toxicology
4. Communicating Computational Toxicology
    a. EPA Program Office Training and Implementation of Computational Tools
    b. Communities of Practice for Chemical Prioritization and Exposure Science
5. Developing Clients for Virtual Tissues

Finding: On the whole, it is evident that addressing the five priority management areas will further support program efficiency and effectiveness, and help sustain the progress the CTRP has achieved to date. The BOSC agrees that the combined programs outlined in the implementation plan are key contributions that are needed to maintain the Computational Toxicology Research Program on a path to achieve the stated goals.

Finding: The CTRP needs to be more integrative, both internally and externally, to ensure all parties are working from common assumptions, data development schedules, and deliverable planning.

As an example, at the review, there were instances where the theoretical and modeling groups noted that they needed access to quantitative data to inform certain aspects of their models while in the same session the experimental groups were presenting precisely those data. Some of this speaks to the stage of the various projects, but as the ultimate goal is to have computational tools that are useful for informing risk assessment, these groups need to begin to work more closely together. Similar, the CTRP needs to continue to expand its outreach to the broader community, both within the EPA and in the extramural community. This is not to say that the CTRP has not been effective in building a strong outreach program, but only that this needs to be a priority, and possibly a higher priority

Recommendation: It will be important to detail specific roles for the STAR Centers as part of the integrated approach to managing the program's mission.

One of the omissions in this management plan is the role that the STAR Centers will play in the future. Although the STAR Centers and their importance in the CTRP's future plans are laid out elsewhere in the implementation plan, it will be important to include these into a more integrated approach to managing the program's mission. Our understanding is that this omission is deliberate and related to the fact that the STAR program funding is independent of the CTRP budget and that consequently management cannot rely on it.

Recommendation: Incorporation of Ecological receptors into the overall CTRP should become a higher priority as the program progresses.

The BOSC noted an absence of ecological health as an endpoint for the high throughput decision support tools for screening and assessing chemical exposure, hazard and risk. To become fully integrated and supportive of the Agency's regulatory activities, the CTRP will have to move into the field of ecological risk assessment at some point. Acknowledging this need and developing a forward plan to incorporate it as part of the CTRP should be part of the longer term plan.

Training

Finding: The review committee was universally impressed with the quality of the postdoctoral fellows and their work.

Recommendation: One area that deserves particular comment in regard to the management plan is its training component, and in particular the training of postdoctoral fellows. The BOSC was universally impressed with the quality of the postdoctoral fellows and their work. We would encourage the CTRP to continue its emphasis on training postdoctoral fellows since these scientists have the potential to be ambassadors to the rest of the community to help extend the understanding and acceptance of the types of computational tools the CTRP is trying to develop and in doing so, ultimately help to improve those tools and their efficacy.

Software Quality Assurance

Recommendation: Highlight software quality assurance with a software testing approach augmented with a sophisticated evaluation approach that probes how the systems produced work in the hands of users.

As the CTRP continues and expands its efforts to develop complex software systems, such as the virtual tissues, software testing and quality assurance becomes ever more important. Software testing is the practice of probing a program for errors (or "bugs"), typically by using a structured set of manually constructed inputs to generate a list of specific performance errors. Static software testing involves inspection of the source code, usually in a structured fashion called a walk-through. Dynamic testing involves executing the code on a set of test inputs. Structured input sets can test how systems perform in the face of boundary conditions (e.g. null or very long inputs), and systematically vary combinations of representative inputs. These methods are called "black box" tests, since they do not require any knowledge of the implementation. Dynamic methods that make use of knowledge of the implementation are called "white box" approaches, for example, code coverage metrics that test what proportion of the source code is reached while processing an input suite. White box methods can provide information about how to improve test suites themselves. Open source approaches to development of software can be exploited as an extension of white box methods bringing large communities of software engineers to the evaluation of code.

A frequent cause of software failures is a lack of compatibility with another application, an operating system, or web browser. Compatibility testing is particularly important in a distributed software development plan, where the separately developed components must be compatible with each other. Test suites that exercise all aspects of each component's interface to the others address this issue in distributed software development.

Testing in each of the proposed computational research areas is complicated by the fact that structured input sets and code coverage metrics have blind spots in these advanced software systems. The range of possible inputs to a tissue model or natural language processing system cannot be exhausted by a structured test set. For these reasons, a software testing approach must be augmented with a sophisticated evaluation approach that probes how the systems produced work in the hands of users.

User-centered Design

Recommendation: Promote "User-centered design", an approach that grounds the process of design in information about the people who will use the product.

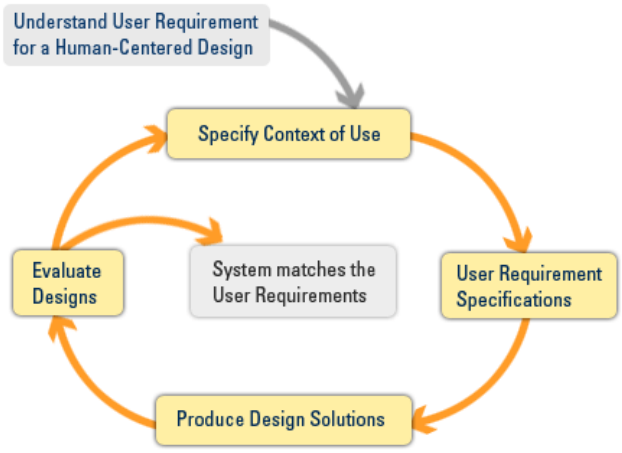
Developing computer systems to support complex and incompletely specified activities (such as modeling toxicity, assessing risk and, prioritizing exposures) is a difficult task. Users are generally unable to specify *a priori* what would be most useful for a computer system to do, although generally do not have trouble describing what they like and don't like about any particular implementation. Furthermore, customer communities are heterogeneous, with needs that vary with analytical goals, methodological approaches, the types of data being analyzed, the amount and quality of relevant background knowledge available, and a wide variety of other factors. Customers and developers both can be frustrated by this seemingly circular need to produce software before defining requirements, which can themselves change in different

circumstances and as new databases and software are produced.

User-centered design is an approach that grounds the process of design in information about the people who will use the product. There is an international standard (ISO 13407: Human-centered design process) that defines the general approach. Figure 1shows the iterative nature of this process: starting with an understanding of the context of use (who will use it, under what conditions, to what ends) leads to a set of requirements that must be met, which in turn leads to a design solution, which is then evaluated through usability testing with actual users, which may lead to additional understanding of the context of use, and so on.

Qualitative methodology is a proven approach for effectively characterizing and explaining such issues as how scientists and policy makers make meaning while proceeding through complex analyses and how they inscribe visualized representations of knowledge into these problem-solving practices (Neressian, 2008). Field observations can reveal the ways in which users generate preferred, new, or augmented analytical practices as they progressively use a new technology (Mirel, 2009; Vicente, 2002). Additionally, qualitative ethnographic methods can reveal ambiguities that scientists negotiate amid biological uncertainty and incomplete data and show the processes by which they disambiguate them. The National Science Foundation recently issued a report stressing the importance of the study of scientists' research processes through qualitative methods (Lamont & White, 2005).

Figure 1: The ISO user centered design process



In general, trained observers using longitudinal fieldwork are better than domain experts themselves at discerning discrete steps in their own reasoning and the role that artifacts such as software play in facilitating their work (Schon, 1983; Hutchins, 1996). Field observations and semi-structured post hoc interviews can importantly complement focus groups or structured interviews or surveys for elucidating longitudinal views of scientists' and policy makers' thinking and behaving (Schon, 1983; Cresswell and Plano, 2006).

Charge Question 5: *The NCCT was established as an organization with a five-year charter ending in February 2010, which would continue dependent on: 1) meeting established goals; and 2) having continuing mission-critical goals and objectives. What recommendation(s) can you provide the Agency regarding continuation of the NCCT as an established organization, and the criticality of its goals and objectives to EPA?*

The BOSC strongly supports action by EPA to make the NCCT permanent. It is clear from EPA's strategy (cite strategy document) that computational toxicology will be integral to the future of toxicology, risk assessment, and regulatory decision-making by the Agency. The Agency will not be able to fulfill its strategy, or indeed its mission, without significant expertise

and an active research program in computational toxicology. The NCCT has made significant contributions during the short time it has been in existence. NCCT work products have had an impact on Agency activities. Work products such as DSSTox, Actor, and ToxRefDB have been of great assistance to program offices and to the toxicology community at large. The longer-term projects underway at the Center have been productive and have demonstrated their potential value to EPA. The staff of the NCCT has proved that the structure of the center, a core of strong expertise that leverages its expertise through collaborations outside the Center, is an ideal organizational structure in a resource-scarce environment. We recommend in the strongest terms that NCCT be made permanent.

The program has been enormously successful in developing the tools and resources necessary to bring computational predictions to the science of toxicology and risk assessment. In the process, they have learned many of the lessons they need to move forward, including many of the reasons why simply building the framework they have in place has been so challenging. In moving forward, the program cannot rest on its laurels, but must continue to address these fundamental infrastructure problems and questions, building on what they have already done. The challenge will be to do this, which can become all-absorbing, while continuing to build the resources that can leverage these resources effectively.

The program has also assembled a tremendous intellectual resource, ranging from the more senior personnel involved in the program to the staff and postdoctoral trainees. It is this intellectual infrastructure as much as the data, databases, and software that represent the real value of the program The various scientists who are the backbone of the program are extraordinarily focused on the end goals and while there is always room for improvement in any program, it is these scientists and their commitment that will assure that the program retains its focus.

One recommendation would be the establishment of performance metrics. While this group has been publishing at a reasonable rate, the primary goal is not academic publication but rather the development of tools and resources for informing risk assessment, and there are potential objectives that can be used to assess the relative impact of these tools on the field, including web hits, software downloads, and citation rate for these tools in publications and grants. While some of this was presented, the program should establish a subset of these as benchmarks and provide some measure of the historical change in these metrics during future funding periods.

In conclusion, the BOSC believes that NCCT is making exceptional progress toward its mission. We are pleased to provide advice on this important Center and look forward to future opportunities to provide timely advice to guide and improve NCCT and its programs.

Sincerely,

Gary S. Sayler, Ph.D.
Chair, BOSC

References:

Cresswell JW, Clark VLP. (2007) Designing and Conducting Mixed Methods Research. California: SAGE Publications, Inc.

Hutchins E. (1995) Cognition in the Wild. Massachusetts: MIT Press.

Lamont L, White P. (2005) Workshop on Interdisciplinary Standards for Systematic Qualitative Research. National Science Foundation Workshop Report, May 19-20, 2005.

Mirel B. (2001) Usability for dynamic work: interface strategies that make or break support for complex problem solving. In: Proceedings of the 2001 Usability Professionals' Association Conference, June 25-29, 2001, Lake Las Vegas, Nevada.

Neressian N. (2008) Creating Scientific Concepts. Massachusetts: MIT Press, 272 pp.

Schon DA. (1983) The reflective practitioner: how professionals think in action. Basic Books, Inc.

Vicente K. (2002) Ecological interface design: progress and challenges. *Human Factors* 44(1):62-78.