ENVIRONMENTAL WORKING GROUP

1436 U Street NW, Suite 100
Washington, DC 20009
T: 202.667.6982
F: 202.232.2592

January 25, 2008

Dr. Michael D. Shelby
Director
Center for the Evaluation of Risks to Human Reproduction
National Institute of Environmental Health Services
Department of Health and Human Services
P.O. Box 12233
MD EC-32
Research Triangle Park, NC 27709

Re: Comments on the Bisphenol A (BPA) Expert Panel Report

Dear Dr. Shelby:

We are writing to provide comments to the National Toxicology Program (NTP) as the agency develops its final position on the extent to which the toxic plastics chemical bisphenol A (BPA) poses a risk to human reproduction and development. As you conduct your review, we urge you to consider fully the following important data and information that is highly relevant to a determination of BPA's potential impacts on human health:

- **The objectivity of CERHR's review of BPA toxicity remains in question.** Sciences International, a contractor who was subsequently fired by the National Institutes of Health for potential conflict of interest, prepared the initial BPA review document for the Center for the Evaluation of Risks to Human Reproduction (CERHR). This document continued to be used by the expert panel despite the fact that several prominent scientists and public health advocates questioned its objectivity. We understand that CERHR's final version of this review document (CERHR 2000a) will be used as a basis for NTP's determination. The objectivity of the findings in this document remain in question and should be reviewed in full by NTP.
- **The CERHR final expert panel report contains many errors and inconsistencies:** In their review of the CERHR expert panel interim draft (CERHR 200b), independent BPA experts identified hundreds of errors and inconsistencies (EWG 2007a); review of the final expert panel draft finds that many of these errors and inconsistencies were not adequately addressed. NTP must ensure that its determination is based on accurate information, and must not rely on the inconsistent and incorrect findings that still plague the CERHR assessment.
- **The CERHR expert panel failed to consider the significant, high exposures to BPA for formula fed infants** (EWG 2007b). NTP's consideration of these exposures is essential if the agency is to reach an accurate determination on BPA's risks.
- **New data confirms the relevance of BPA studies that used non-oral routes of administration.** These studies were categorically excluded by the expert

panel in reaching their final decision. In light of the new data, NTP should incorporate findings from these excluded studies in making its determination.

- **BPA experts raise serious concerns about potential human health impacts from BPA exposures**. A NIEHS-sponsored panel of 38 BPA experts, which convened in Chapel Hill, North Carolina in November of 2006, published a comprehensive consensus statement regarding BPA toxicity and determined that BPA exposure is a risk to human health (vom Saal 2007); the NTP's thorough consideration of this expert panel's findings is critical.

Each of these points is described in detail below.

The objectivity of CERHR's review of BPA toxicity remains in question. In March of 2007, the National Institutes of Health fired the contractor (Sciences International) that was hired to prepare the initial BPA review document, citing potential conflicts of interest when information became available that showed that Sciences International staff had previously worked with BPA manufacturers. The document that was prepared by this contractor, however, continued to be used by the expert panel, despite the fact that several prominent scientists and public health advocates had questioned the objectivity of the Sciences International review. CERHR's final report is derived from this original, conflicted document. The objectivity of findings in this final report remains in question and must be thoroughly reviewed by NTP.

The CERHR final expert panel report contains many errors and inconsistencies. In April of 2007, CERHR released an interim draft expert panel report (CERHR 2007b); instead of addressing the issues that were brought up regarding the objectivity of the initial review prepared by Sciences International, the interim draft was even more error-riddled and even less objective that the initial review. In fact, BPA experts who reviewed this interim draft noted hundreds of errors in documentation, analysis, and interpretation; they submitted these findings in written comments to the expert panel in June of 2007 (EWG 2007a). These BPA experts found the following:

- 297 potential errors in documentation and analysis of study results, and in interpretation of the study findings and their significance that are in conflict with the peer reviewed literature
- 195 instances where the panel assessment is incomplete, including incomplete documentation of relevant test results or missing justifications for panel assertions
- 48 instances in which the panel inconsistently applied criteria for study evaluation

Our detailed evaluation of these errors and inconsistencies in the April 2007 draft in included as an attachment to this letter (attachment 1). These issues with documentation, analysis, and interpretation resulted in an assessment that heavily favored industry studies over government and independent studies. In this interim draft, the expert panel rejected government and independent studies at 3 times the rate of industry studies.

Only some of these issues have been addressed in the final expert panel report that was released in November of 2007 (CERHR 2007a). Significant inconsistencies and errors remain within this report.

In just one example, the panel reviewed a study from Cagen et al in section 3.2.3.2 in which they noted "the lack of much effect with diethylstilbestrol treatment is a weakness", but they go on to conclude "the panel considered this study adequate and of high utility". However, in a review of a second study by Cagen et al in section 3.2.5.1 in which there were also problems with the positive control, the panel noted "this paper is inadequate for the evaluation process due to absence of response of the positive control group". It is unclear why two studies in which there are serious issues with the positive control are judged so differently.

The expert panel had also noted in the interim draft that they had specific concerns with the use of DMSO as a vehicle for BPA because of its biological activity: in response to this, we had noted in our public comments from June of 2007 that while the panel singled out DMSO, they did not raise any objections to the use of oil vehicles which have been shown to often have background estrogenicity. This issue does not appear to have been addressed at all by the expert panel.

The fact that there are still inconsistencies in the expert panel's final draft illustrates how the CERHR expert panel continues to apply arbitrary standards throughout this evaluation.

The CERHR expert panel failed to consider the significant, high exposures to BPA for formula fed infants. We would also bring your attention to our recent report on the presence of BPA in canned infant formula (EWG 2007b). Laboratory studies of canned infant formula conducted by the Food and Drug Administration (FDA) and a certified commercial laboratory commissioned by EWG reveal that BPA leaches from metal can linings into formula. EWG analysis of these results revealed the following:

- One of every 16 infants fed ready-to-eat canned formula would be exposed to BPA at doses exceeding those that altered testosterone levels, affected neurodevelopment, and caused other permanent damage to male and female reproductive systems (2.0 and 2.4 ug/kg/day- Howdeshell et al 1999, Honma et al 2002- studies cited as 'adequate' by the expert panel)
- At the highest BPA levels found in formula (17 parts per billion), nearly two-thirds of all infants fed ready-to-eat formula would be exposed above doses that proved harmful in animal tests (2.0 and 2.4 ug/kg/day- Howdeshell et al 1999, Honma et al 2002)

Laboratory studies have consistently shown that the most sensitive periods of exposure to BPA are during pregnancy and early life (Maffini 2006). These infant formula findings reveal that millions of formula fed infants may have daily, sustained exposures to BPA at levels that have been shown to cause harm in lab animals. These exposures could be relatively continuous throughout their first 6 months of life and should be

fully considered by NTP as the agency reaches a determination on BPA's potential impacts on human health.

New data confirms the relevance of BPA studies that used non-oral routes of administration. At the close of the second expert panel meeting in August of 2007, the CERHR expert panel issued conclusions regarding the potential reproductive and developmental toxicity of BPA. They expressed "some concern" that exposure to BPA may cause neural and behavioral effects in the developing fetus, infants, and children, but expressed "minimal concern" about other potential health effects. The panel came to these conclusions by discarding a large number of independent, peer-reviewed studies linking BPA exposure to mammary and prostate gland lesions, impaired fertility, and ovarian dysfunction in lab animals.

The panel made a decision to include data from only those studies in which BPA was administered to lab animals via oral routes of administration. This decision by the expert panel resulted in the exclusion of many well-conducted studies from academic labs from across the United States; most importantly, the studies that were excluded had all passed through the rigors of peer review and had been published in a number of prestigious scientific journals. In addition, this decision to exclude data from studies in which BPA was administered to lab animals via non-oral routes of administration is not backed up by the scientific literature and is highly unusual in public health evaluations.

In fact, a recent study published in the journal Reproductive Toxicology finds that non-oral routes of BPA administration are completely valid in assessing potential health effects (Taylor 2008). In this study, scientists administered BPA to neonatal mice by both oral and subcutaneous routes and found no significant difference in plasma levels of unconjugated BPA, leading study authors to concluded "the large numbers of BPA studies that used non-oral administration at very low doses during the neonatal period should not be dismissed by scientists or the regulatory community based on route of administration".

This definitive scientific study refutes the faulty reasoning that the expert panel used in discarding many valid and scientifically sound studies that linked low dose BPA exposure with adverse health effects such as breast and prostate cancer, infertility, and early puberty. In fact, the decision by the expert panel to disregard studies that used a non-oral route of administration is just one more in an long list of missteps that has plagued this review and leads us to question the validity of the conclusions.

BPA experts raise serious concerns about potential human health impacts from BPA exposures. Lastly, we would like to draw your attention to a series of papers that were published in the journal Reproductive Toxicology in 2007. These papers were published by a group of 38 BPA experts from around the world who systematically reviewed over 700 BPA related scientific papers.

In contrast to the CERHR expert panel, none of whom were BPA experts, this group of 38 scientists included many of the world's most published BPA experts from top

academic universities and government institutions. These scientists conducted a highly structured and organized review of the BPA literature that focused on consensus building; their findings were condensed into a consensus statement in which they concluded:

- **The similar effects observed in wildlife and laboratory animals exposed to BPA predict that similar effects are also occurring in humans** and
- **Much evidence suggests that these adverse effects are occurring in animals within the range of exposure to BPA of the typical human living in a developed country, where virtually everyone is exposed to measurable blood, tissue and urine levels of BPA that exceed the levels produced by doses used in the low dose animal experiments"** (vom Saal 2007)

The final decision rendered by NTP regarding the reproductive and developmental toxicity of BPA will have repercussions on the public health, and as such, should not be based on the flawed and biased CERHR evaluation. We urge NTP to recognize the merits of the review conducted by the 38 BPA experts, in comparison with the CERHR evaluation. The scientific data clearly backs up their concerns about the reproductive and developmental toxicity of BPA and we urge NTP to recognize the low dose developmental and reproductive toxicity of BPA.

Sincerely,

Anila Jacob, M.D., M.P.H.
Senior Scientist
Environmental Working Group

References:

1) CERHR 2007a. Expert panel report on Bisphenol A. Bisphenol A evaluation. Center for the Evaluation of Risks to Human Reproduction.
2) CERHR 2007b. Interim draft expert panel report. Bisphenol A evaluation. Center for the Evaluation of Risks to Human Reproduction.
3) EWG (Environmental Working Group) 2007a. Failure of CERHR assessment of BPA to meet basic scientific standards. Attachment 1.
4) EWG (Environmental Working Group) 2007b. Toxic plastics chemical in infant formula. Available online at: http://www.ewg.org/node/22233.
5) Honma S, Suzuki A, Buchanan DL, Katsu Y, Watanabe H, Iguchi T. (2002). Low dose effects of in-utero exposure to bisphenol A and diethylstilbestrol on female mouse reproduction. Reproductive Toxicology 16: 117-22.
6) Howdeshell K, Hotchkiss AK, Thayer KA, Vandenbergh JG, vom Saal FS. 1999. Plastic bisphenol A speeds growth and puberty. Nature 401: 762-64.

7) vom Saal FS, Belcher SM, Guillette LJ, Hauser R, Myers JP, Prins GS, Welshons WV et al. 2007. Chapel Hill Bisphenol A Expert Panel Consensus Statement: Integration of mechanisms, effects in animals and potential impact to human health at current exposure levels. Reproductive Toxicology 24: 131-38.
8) Maffini MV, Sonnenscheim C, Soto AM. 2006. Endocrine disruptors and reproductive health: the case of bisphenol A. Molecular and Cellular Endocrinology 8: 254-55.
9) Taylor JA, Welshons WV, vom Saal FS. 2008. No effect of route of exposure (oral; subcutaneous injection) on plasma Bisphenol A throughout 24 hr after administration in neonatal female mice. Reproductive Toxicology 25(2): in press.

ATTACHMENT 1

ENVIRONMENTAL WORKING GROUP

1436 U Street NW, Suite 100
Washington, DC 20009
T: 202.667.6982
F: 202.232.2592

August 6, 2007

Dr. Michael D. Shelby
Director
Center for the Evaluation of Risks to Human Reproduction
National Institute of Environmental Health Services
Department of Health and Human Services
P.O. Box 12233
MD EC-32
Research Triangle Park, NC 27709

Re: **Failure of CERHR Assessment of BPA to Meet Basic Scientific Standards**. Supplemental Comments on the Interim Draft NTP-CERHR Report on the Reproductive and Developmental Toxicity of Bisphenol A.

Dear Dr. Shelby:

You must be aware of the publication last week of a consensus statement on BPA signed by 38 independent specialists in BPA toxicity from around the world. These scientists concluded that BPA presents a clear risk to human health. The statement and the comprehensive review papers that accompany it underscore, by way of contrast, the hopeless corruption of the ongoing review of BPA being conducted at your Center (the NIH Center for the Evaluation of Risks to Human Reproduction, or CERHR).

The Environmental Working Group has conducted a detailed review of the comments by 9 scientists conducting BPA research at 6 laboratories in the U.S. and E.U., submitted to you in response to CERHR's interim draft BPA assessment (Vandenberg et al. 2007; Schonfelder 2007; Prins 2007; vom Saal 2007; Welshons 2007; Zoeller 2007). Our review shows that the CERHR panel's review of BPA utterly fails to meet basic, universally understood standards for scientific reviews and data quality, including those laid out in NIH policy and federal law. These standards require that assessments be accurate, unbiased, consistent, complete, and conducted by those with the necessary expertise and independence to ensure objectivity. Instead, our review of scientists' comments reveals that the CERHR assessment may contain nearly 300 errors of fact and interpretation; is biased, inconsistent, incomplete; and clearly fails to meet the most basic scientific standards. Among our findings, which are detailed in the attached table, are that the CERHR assessment is:

- **Inaccurate. 297 errors:** Reviewers identified 297 potential errors in documentation and analysis of study results, and in interpretation of the study findings and their significance, in conflict with the peer reviewed literature.

- **Incomplete. 195 instances of incomplete study reviews:** Reviewers documented 195 instances where the panel assessment is incomplete, including incomplete documentation of relevant test results or missing justifications for panel assertions.

- **Inconsistent. 48 basic inconsistencies:** Reviewers documented 48 instances in which the panel inconsistently applied criteria for study evaluation.

- **Biased.** The assessment heavily favors industry studies over government and independent studies. In its most recent assessment, the Panel rejected government and independent studies at 3 times the rate of industry studies (Vandenberg et al. 2007).

Consider also the following, striking differences between the CERHR panel and the BPA panel which released the consensus statement last week (this panel convened in Chapel Hill, NC, and is referred to as the "Chapel Hill panel" for purposes of this document). Both panels are funded by NIH, but are different in almost every other aspect:

- **The objectivity of the CERHR assessment is compromised by the panel's lack of specialization in BPA science**. The CERHR panel contains few with advance knowledge BPA through their own study, and none for whom the chemical has been a primary focus of their work. The panel has just 12 members to assess over 500 BPA-related papers. The Chapel Hill panel includes 38 of the world's most published BPA experts from top universities and government institutions.

- **The accuracy and consistency of the CERHR assessment is compromised by the panel's organizational structure and their failure to communicate**: Within the CERHR panel study reviews were conducted independently by each scientist, prompting on panel member to state in a recent article in Risk Policy Report that "one thing that has plagued this review is that each reviewer was assigned a bunch of papers, and they reviewed them without any other input." Additionally, in the middle of the panel review process CERHR fired the consulting firm managing the project over concerns about potential conflicts of interest, and a director of the Center was replaced for reasons not disclosed to the public, creating significant changes in management in the midst of the review. In contrast, the Chapel Hill review was conducted in a highly structured, organized manner: 4 breakout groups were each asked to address 4 critical issues related to BPA, and only if there was consensus among all 4 groups were responses incorporated into the final consensus statement.

- **The objectivity of the CERHR assessment is compromised by not having been subjected to a standard peer review**: The assessment of the CERHR panel has not been subjected to standard peer review, and contains nearly 300 errors of fact and interpretation according to BPA specialists. The initial draft was prepared by the contractor mentioned above who was fired over concerns about possible conflict of interest, calling into question the validity of the contractor's work on this assessment. In contrast, the work of the Chapel Hill panel was subjected to standard and comprehensive internal and independent external peer review.

We question the Center's ability to produce a scientifically sound document from this process when the comments you have received from BPA experts include statements calling into question the ability of the panel to meet the most basic scientific standards:

- "*Is the panel purposefully misrepresenting data or grossly misunderstanding it?*" (Vandenberg et al. 2007)
- "*There are two general aspects [of the assessment] which to me represent the antithesis of valid science.*" (Welshons 2007)
- "*The criteria established by the panel are arbitrary*." (Vandenberg et al. 2007)
- "*highly curious*" [comment on an uninformed critique] (Zoeller 2007)

- "*If one were seeking to establish a mechanism that would be virtually certain to underestimate the potential for harm to be caused by a chemical, the CERHR mechanism is exactly the process that they would want to establish to achieve that objective*." (vom Saal 2007).

The public has now paid for two assessments of BPA toxicity, the one conducted by your panel, which has failed to meet basic standards for the conduct of scientific reviews; and a peer reviewed assessment by a panel of BPA specialists (the Chapel Hill panel), which issued its final assessments last week. If you proceed with the CERHR panel process the public will have to pay for this assessment four times all told, because your assessment will require both a thorough peer review, and a complete revision from top to bottom of the current, corrupted document.

Instead of trying to salvage what you have, we urge you to dissolve your current panel and adopt the recommendations of the Chapel Hill panel issued last week, which is peer reviewed and meets established scientific standards. We would also urge you to invest your Center's resources in conducting studies and forwarding policies that will help reduce the public's exposures to this chemical that so clearly poses risks to human reproduction.

EWG has recently requested data from formula manufacturers on BPA levels in their products, to help fill data gaps left by FDA's meager, 14-sample study. We have also completed a new analysis of infant formula showing, based on available data, that babies who drink ready-to-feed formula can easily be exposed to BPA in amounts that exceed those found harmful in laboratory studies, and are exposed to BPA at greater levels than any other segment of the population. Our correspondence to formula manufacturers and our new infant formula analysis are available on our website at www.ewg.org.

BPA poses risks to human reproduction and is an urgent public health issue. We urge you to lead on this issue instead of spending energy to rectify a corrupt product from a corrupt process.

Sincerely,

Anila Jacob, M.D., M.P.H
Senior Scientist

Jane Houlihan
Vice President for Research

References

All references listed below except CHCS (2007) are currently available in pdf form at http://cerhr.niehs.nih.gov/chemicals/bisphenol/pubcomm-bisphenol.html.

CHCS (Chapel Hill Consensus Statement). 2007. vom Saal FS, Akingbemi BT, Belcher SM, Birnbaum LS, Crain DA, Eriksen M, Farabollini F, Guillette LJ, Hauser R, Heindel JJ, Ho SM, Hunt PA, Iguchi T, Jobling S, Kanno J, Keri RA, Knudsen KE, Laufer H, LeBlac GA, Marcus M, McLachln JA, Myers JP, Nadal A, Newbold RR, Olea N, Prin GS, Richter CA, Rubin BS, Sonnenshein C, Soto AM, Talsness CE, Vandenbergh JG, Vendenberg LN, Walser-Kuntz DR, Watson CS, Welsons WV, Wetherill Y, Zoeller RT. Integration of Mechanisms, Effects in Animals and Potential to Impact Human Health at Current Levels of Exposure. Reproductive Toxicology 24(2): 2007, in press.

Prins GS. 2007. Letter from Dr. Gail S. Prins, College of Medicine, Department of Urology, University of Illinois at Chicago, to Dr. Michael Shelby, National Institute of Environmental Health Sciences. June 19 2007.

Schonfelder G. 2007. Letter from Dr. Gilbert Shonfelder, Institute of Clinical Pharmacology and Toxicology, Campus Benjamin Franklin, Charity University Medicine Berlin, to Dr. Michael Shelby, National Institute of Environmental Health Sciences. Received June 20 2007.

Vandenberg LN, Maffini MV, Rubin BS, Soto AM. 2007. Response to the Interim Draft of the NTP-CERHR Report on the Reproductive and Developmental Toxicity of Bisphenol A. Tufts University School of Medicine. Department of Anatomy and Cellular Biology. Boston, MA. June 20, 2007.

Vom Saal 2007. Comments on the Interim CERHR Report on Bisphenol A (April 2007). Comments from Dr. Freferick vom Saal, Division of Biological Sciences, University of Missouri, Columbia, MO, to the National Institute of Environmental Health Sciences. Received June 20, 2007.

Welshons WV. 2007. Comments from Dr. Wade V. Welshons, Department of Biological Sciences, Verterinary Medicine, University of Missouri, Columbia MO, to Dr. Michael Shelby, National Institute of Environmental Health Sciences. June 20 2007.

Zoeller RT. 2007. Letter from Dr. R. Thomas Zoeller, Biology Department, Morrell Science Center, University of Massachusetts Amherst, to Dr. Michael Shelby, National Institute of Environmental Health Sciences. May 19 2007.

ATTACHMENT

Table. CERHR Assessment of BPA fails to meet basic scientific standards **for data quality and objectivity.**

A summary of comments to CERHR from independent BPA scientists

Scientific standards for objectivity and data quality	CERHR assessment fails to meet the standard in this area...	Failure of CERHR assessment according to BPA expert review comments	Reviewer (reference)
Accurate	Errors in interpreting significance of study findings	Panel mistakenly concluded that study effect is of unknown relevance (Ho et al, Cancer Research 2006), when a substantial literature demonstrates its relevance, as described in Prins (2007).	Prins 207
Accurate	Errors in conclusions on statistical analysis	Panel deemed a study "inadequate based on inappropriate statistics" with no justification for this conclusion, even though the study relied on accepted, rigorous statistical analysis (one-way analysis of variance) unlikely to merit criticism if it had been assessed by the panel (Zoeller et al. 2005).	Zoeller 2007
Accurate	Errors in understanding study implications	Panel criticized a study for failing to include a positive control (Zoeller et al. 2005), but failed to recognize that the study provided first-ever data on the effect measured, so no positive control compound would be available. The reviewer calls the panel's uninformed critique "highly curious."	Zoeller 2007
Accurate	Errors in understanding study implications	Panel failed to understand a major finding of a critical thyroid study, which showed that BPA produced a profile of effects that were consistent with the interpretation that BPA acts as a selective indirect antagonist on the beta thyroid hormone receptor (Zoeller et al. 2005).	Zoeller 2007
Accurate	Failure to correct errors of fact noted by commenters	Factual errors in prior draft of document, noted in review comments, were not corrected in this draft.	vom Saal 2007
Unbiased	Apparent bias of errors to favor industry-funded assessments	Reviewer notes apparent bias in distribution of errors in panel assessment that would favor industry-funded assessments and would lead to an assessment underestimating potential for harm from BPA exposures.	vom Saal 2007
Expertise, Accurate, Complete, Reliable	Exclusion of BPA specialists from the panel	Reviewer notes that errors and lack of attention to critical issues are likely related to the panel's relative lack of expertise in BPA research and lack of familiarity with BPA-related scientific literature.	vom Saal 2007
Consistent, Complete, Accurate	Failure to establish and document defensible criteria for assessing the BPA literature	Reviewer notes that criteria established by the panel for assessment of study utility are arbitrary, many without explanation or merit.	Vandenberg et al. 2007, pg 2
Accurate, Expertise, Reliable	Failure to follow accepted scientific standards for review and study evaluation	Reviewer notes that two fundmental aspects of the assessment are the "antithesis of valid science:" Failure to institute standard peer review of panel findings that would ensure accuracy; failure to apply established criteria for inclusion of and interpreting results from control animals that are critical to understanding BPA study findings.	Welshons
Consistent, Accurate	Inaccurate and incomplete assessments of study findings and	Reviewer notes that developmental toxicity data section of the panel's assessment is "filled with inconsistencies and inaccurate statements."	Vandenberg et al. 2007, page 44

	implications		
Accurate	Inaccurate representation of study findings	Reviewer notes that the panel is either "purposefully misrepresenting data or grossly misunderstanding it," and that summaries provided for some studies are "completely inaccurate and do not represent the experiments conducted and/or the results obtained."	Vandenberg et al. 2007, page 8
Accurate, Reliable	Inaccurate interpretations resulting from review of studies by panel without the requisite specialized knowledge.	Reviewer notes that the panel makes "fundamental errors due to ignorance" in its assessments of BPA studies, and comments that the panel reviews do not meet journal standards for peer review.	Welshons
Accurate	Lack of understanding of endocrine research	Panel has failed to understand the basic requirement in endocrine research for study controls and the implications of results in control animals. The reviewer calls this failing "beyond my comprehension" given the widely understood need for negative controls in positive experiments, and positive controls in negative experiments in endocrine research. The reviewer notes that in modern endocrine research, many of the studies without controls that the panel considers acceptable would, in fact, be unpublishable.	Welshons
Complete, Accurate	Lack of scientifically sound reasoning in discounting important study observations	Reviewer notes that observations in Zoeller et al. (2005) are discounted by the panel for reasons that are unclear or do not appear to be valid.	Zoeller
Accurate	Misrepresentation of study significance	Reviewer states that panel systematically misrepresented significance of metabolic studies.	Vandenberg et al. 2007, page 8
Accurate	Errors in understanding of need for consistency in control studies	Panel mistakenly called an industry-funded study (Tyl et al. 2003) a "replication" of an independent study (Nagel et al. 1997) which found effects, even though the Tyl study used a feed known to be contaminated with estrogenic compounds that mask the effects of BPA. It is standard scientific practice in replication to use identical feed to the study being replicated. This study found no BPA-related effects.	vom Saal 2007
Accurate	Errors in interpreting scientific literature on animal feed contamination and influence on BPA study findings	The panel failed to consider that two industry-funded studies of BPA which failed to find effects used animal feed known to be contaminated with estrogenic substances that can completely mask the effects of BPA (Thigpen et al. 2003 (Comp Med)). The panel also raised concerns about feed contamination in a study relying on feed proven to be free of the contaminants at any levels that would affect study findings (vom Saal et al. 1997, PNAS; Timms et al. 2005, PNAS; Richter et al. 2007, EHP; Howdeshell et al. 1999, Nature; Palanza et al. 2002, EHP).	vom Saal 2007
Accurate	Errors in interpreting the role of positive controls in study validation	The panel failed to note concerns that an industry-funded studies showing no effects from BPA also found no effects in its positive control animals. Positive control animals are exposed to a substance known to produce the same effect scientists are seeking to explore with the study's test substance. It is standard scientific practice to consider a study as having failed if the positive control animals fail to show a response, because this failure means that the study design or test conditions would not allow the study to reveal effects from the test substance, either.	vom Saal 2007; Schonfelder 2007

Reliable, Accurate	Errors in interpretation stemming from lack of panel members who specialize in BPA research	Reviewer notes that failure to integrate the range of avialable data on animal feed contamination, and other related failings of the panel, likely stems from the fact that panel composition largely excludes members who specialize in BPA research.	vom Saal 2007
Accurate	Errors in interpreting study results as "false positives."	The panel categorized as potential "false positives" findings replicated in many experiments.	vom Saal 2007
Accurate	Errors in interpreting relevance of exposure route.	The panel categorized studies with continuous instead of episodic exposures useless in their evaluation of BPA toxicity in the complete absence of data to support that conclusion, and without recognizing studies which support the continuous sources of exposure for humans, including BPA contamination in dust and air. The panel justified exclusion of these studies in part by expressing concerns that injection results in excess unmetabolized BPA in the bloodstream relative to oral exposures, but failed to recognize that this form of exposure mimics human fetal exposures, and also failed to note the substantial body of literature showing unmetabolized BPA in human tissues and fluids that further support the relevance of continuous exposure studies.	vom Saal 2007
Accurate	Errors in interpreting relevance of exposure route and potential for "false positives."	The panel categorized as potential "false positives" findings replicated in many experiments. The panel categorized studies with continuous instead of episodic exposures useless in their evaluation of BPA toxicity in the complete absence of data to support that conclusion, and without recognizing studies which support the continuous sources of exposure for humans, including BPA contamination in dust and air. The panel justified exclusion of these studies in part by expressing concerns that injection results in excess unmetabolized BPA in the bloodstream relative to oral exposures, but failed to recognize that this form of exposure mimics human fetal exposures, and also failed to note the substantial body of literature showing unmetabolized BPA in human tissues and fluids that further support the relevance of continuous exposure studies.	vom Saal 2007
Accurate	Errors in analysis of study findings.	The panel notes a "lack of clarity" in mouse strain in a study deemed of limited usefulness even though the mouse strain is clearly stated in the study ("CF-1 mice were purchased from Charles River Laboratories...", vom Saal et al 1998).	vom Saal 2007
Accurate	Errors in analysis of study findings.	The panel notes a need for consideration of testis weight in a study deemed of limited usefulness, even though the study clearly documents the effect of testis weight and uses it as the basis for analysis (vom Saal et al 1998).	vom Saal 2007
Accurate	Errors in analysis of study findings.	The panel inaccurately characterizes study findings, stating that studies did not find statistically significant effects on the prostate at 0.020 mg/kg bw/day, when the studies clearly report statistically significant effects at that dose (Nagel, 1997 #6; vom Saal, 1998 #187). [unclear - did both studies show effect at this dose?]	vom Saal 2007
Accurate	Errors in analysis of study scope	Panel mistakenly concluded that study implications are exclusive of estrogenic endpoints when study authors clearly state otherwise (Nagel et al. 1997,	vom Saal 2007

		vom Saal et al. 1998).	
Accurate	Errors in analysis of the nature of the study	Panel fails to understand the study analysis (Howdeshell et al. 1999, Nature), which relied on a litter-based analysis consistent with a significant literature on the effects of intrauterine position on pup response, but which was confused by the panel reviewer as potentially a pup-based analysis and deemed of marginal utility.	vom Saal 2007
Accurate	Errors in analysis of study design	The panel noted confusion on whether the pup or litter was used as the statistical unit for analysis for a study (Gupta 2000 PSEBM) when the author clearly states that 15 individual pups from 15 separate litters were used in the analysis.	vom Saal 2007
Complete	Exclusion of important review findings	The panel included findings from a published critique of Gupta (2000, PSEBM) but failed to note that the model used as the basis for the critique (Elswick et al. 2000) was deemed "misleading", "illogical" and "flawed" by the NIH Low Dose Review Panel.	vom Saal 2007
Accurate	Lack of understanding of statistical significance in scientific studies	Panel rejected studies for evaluation based solely on judgment that the studies provided "an insufficient number of animals for rigorous statistical analysis," reducing "confidence in the results." The number of study animals required is based on power analysis and is small when the expected magnitude of the effect is great, where small numbers of animals can yield statistically significant results. NIH guidance requires the use of the fewest animals possible to achieve statistical significance. With their conclusion the panel ignores basic, widely accepted, NIH-endorsed statistical principles behind the design of toxicological studies.	vom Saal 2007; Vandenberg et al. 2007
Accurate	Errors in interpretation of toxicological implications and validity of study design.	The panel inappropriately excluded some studies because of their use of DMSO as a vehicle to administer BPA to test animals, failing to note that DMSO exposures are far below those that are known to be biologically active, that DMSO did not produce effects in control animals, that DMSO is not associated with effects related to those under investigation for BPA, and that DMSO has been the vehicle of choice for a wide range of related studies.	Vandenberg et al. 2007, pg 2
Consistent	Inconsistent evaluation of study design	The panel erroneously dismissed studies using DMSO as an administration vehicle, but does not similarly discuss or assess a wide range of other potentially problematic administration vehicles in other studies, including corn oil that can be contaminated with estrogenic compounds that might mask the effects of BPA.	Vandenberg et al. 2007, pg 3
Consistent	Inconsistent application of evaluation criteria	The panel proposed a hypothesis that findings in a BPA study they reviewed could be due to interactions between BPA and the administration vehicle (olive oil), but fails to note the same concern with respect to the other 13 studies evaluated that use olive oil as the administration vehicle.	Vandenberg et al. 2007, pg 3
Consistent, Complete	Inconsistent application of evaluation criteria	The panel found fault with injection as dosing vehicle but failed to note concerns with other dosing vehicles, including stress induced by oral gavage that has been shown to mask the effects of low doses of hormones, and inaccuracies stemming from non-uniform feed and water exposures.	Vandenberg et al. 2007, pg 5
Accurate	Arbitrary and capricious choice of	The panel arbitrarily chose 7 as an appropriate sample size for animal experiments without	Vandenberg et al. 2007, pg 5

	study evaluation criterion	justification. The reviewer notes that this "capricious choice... is contrary to the understanding of statistical power and sample size analysis, which should be done by the experimenter a priori, i.e. before conducting the experiment."	
Accurate	Errors in interpreting non-monotonic dose responses	The panel criticized 12 studies as having "non-dose related" results, demonstrating a lack of understanding of the widely observed non-monotonic responses characteristic of many endocrine studies and widely reported in the scientific literature, having been observed in 40% of the studies with a design that would allow for its detection.	Vandenberg et al. 2007, pg 5
Consistent, Complete	Inconsistent application and inadequate documentation of evaluation criteria	The panel dramatically changed study evaluation findings between drafts of the assessment but failed to document changes in evaluation criteria that resulted in these alterations, and failed to note if evaluation criteria were established at all in advance of the initial review.	Vandenberg et al. 2007, pg 3
Consistent	Inconsistent application of evaluation criteria	In Section 3 of the assessment the panel criticized 43% of all studies reviewed for what the panel perceived as an inadequate sample size (n<7), but then inexplicably failed to note this same concern in 11% of other studies reviewed that also had an n<7.	Vandenberg et al. 2007, pg 7; Schonfelder 2007
Consistent	Inconsistent application of evaluation criteria	The panel rejected a number of studies from consideration because of their use of DMSO as an administration vehicle, but inexplicably accepted two other studies that also used DMSO, and for one, supplied by the plastics industry, failed to even note it as a concern.	Vandenberg et al. 2007, page 8
Accurate, Complete	Inaccurate interpretation of significance of pharmacokinetic studies	The panel failed to note concentrations of BPA in critical tissues in pharmacokinetic study as the determining factor for toxicity, or significance of study for fetal exposures.	Vandenberg et al. 2007, page 8
Accurate	Inaccurate representation of study findings	In review of a study of anogenital distance, the panel noted that "study authors concluded that the endpoint was not affected by prenatal, lactional and/or post-wearning exposures to bisphenol A," when, in fact, the study authors did not draw this conclusion (Rubin et al. 2001).	Vandenberg et al. 2007, page 9
Accurate	Lack of understanding of effect of dose and timing on endocrine studies	The panel noted a lack of reproducibility associated with a studies of LH serum levels, but fails to recognize that BPA scientists would not attempt to reproduce endocrine effects in female rats from fetal development until weaning by studying the same endocrine effects in male rats during puberty (Rubin 2001, Akingbemi 2004).	Vandenberg et al. 2007, page 9
Accurate	Lack of understanding of dose response	The panel cites concerns about a "lack of dose response releationships" in a study that found effects only at the highest dose (Rubin et al. 2001), a situation in which consideration of dose response is irrelevant.	Vandenberg et al. 2007, page 9
Accurate	Misrepresentation of study doses	The panel erroneously lists doses included in a study of BPA effects in mammary glands (Murray 2007).	Vandenberg et al. 2007, page 9
Accurate, Complete	Misrepresentation of studies and failure to document	In a discussion of a study for which the panel failed to provide a reference, the panel noted a mammary gland finding that has never been observed in a study (mammary gland alterations in pubertal and adult mice). They appear to be referencing a study conducted on gestational day 18 mice, but if so, mistakenly note that the study used subcutaneous silastic implants for administration (Vandenberg et	Vandenberg et al. 2007, page 9

		al. 2007).	
Consistent	Inconsistent application of evaluation criteria	Four studies were considered adequate by the panel even though their positive controls failed, a clear, widely accepted sign of a failed experiment.	Vandenberg et al. 2007, page 10
Accurate	Failure to recognize key components of study design	The panel failed to adhere to standard scientific practice in considering the need for and results from negative controls in endocrine studies, and in multiple instances ignored common potential sources of contamination such as phytoestrogens in feed, estrogenic activity of oil vehicles, plastics in the animal environment, and contamination in tissue culture experiments.	Vandenberg et al. 2007, page 10
Consistent	Inconsistent application of evaluation criteria	The panel inconsistently treated the lack of information about feed in study documentation, listing it as a weakness in only 14% of the studies reviewed that failed to document it.	Vandenberg et al. 2007, page 10
Accurate, Complete	Failure to recognize key components of study design	In studies with oil vehicles the panel failed to note as a weakness failure to include negative controls, and the panel failed to discuss findings in negative controls when they were included. (Note that for none of the oil vehicle studies including negative controls were the results in those animals discussed by the panel.)	Vandenberg et al. 2007, page 10
Accurate, Complete	Failure to recognize key components of study design	The panel failed to recognize the documented potential for contamination from polycarbonate cages, bottles, and other plastics used in BPA experiments.	Vandenberg et al. 2007, page 10
Unbiased	Bias in study evaluations	In this most recent draft of the assessment, the panel changed many of their evaluations of study adequacy, and found many more studies inadequate than had been proposed in the previous draft, but in making these changes they rejected independently funded studies at three times the rate of industry funded studies.	Vandenberg et al. 2007, page 11
Unbiased, Complete	Bias in study inclusion, Incomplete list of studies for assessment	In the most recent assessment draft the panel stripped from the report, without explanation, the results description for 38 studies that were included in the previous draft. Only one of these was funded by industry, and 32 of these were originally considered adequate.	Vandenberg et al. 2007, page 8
Unbiased	Potential bias in study interpretation based on perception of quality of the research group	Panel review notes strengths of studies to include "the expertise of the group" and "well conducted by a respected lab," indicating reviewers' preferences for labs and calling into question the objectivity of the review	Vandenberg et al. 2007, page 36, 56
Reliable	Reliance on upublished industry studies and translations not subjected to peer review	Reliance on unpublished industry studies not subjected to peer review process, and reliance on translations of select parts of studies published in Japanese, provided to the panel by the American Plastics Council	Vandenberg et al. 2007, page 32, 36, 41, 53, 57
Accurate	166 instances of errors in documentation, analysis, interpretation, and evaluation of significance	Reviewer identified 166 potential errors in documentation of study results, and in interpretation of the study findings and their significance, in conflict with the peer reviewed literature. These are described in the detailed section of this reviewer's comments. Data presented in the comment's summary section are, in contrast, documented individually in this table.	Vandenberg et al. 2007, pg 16-60
Complete	95 instances of incomplete documentation and consideration of study	Reviewer documented 95 instances where panel assessment appears to be incomplete, including incomplete documentation of relevant test results or justification for panel assertions. These are	Vandenberg et al. 2007, pg 16-60

	findings, interpretations, and conclusions	described in the detailed section of this reviewer's comments. Data presented in the comment's summary section are, in contrast, documented individually in this table.	
Consistent	39 instances of inconsistent application of evaluation criteria	Reviewer documented 39 instances in which the panel inconsistently applied criteria for study evaluation. These are described in the detailed section of this reviewer's comments. Data presented in the comment's summary section are, in contrast, documented individually in this table.	Vandenberg et al. 2007, pg 16-60
Accurate, Complete	88 instances of panel failing to consider absence of testing for background contamination	Reviewer documented 88 instances of the panel failing to consider the potential effect of contamination in studies which did not test for estrogenicity of administration vehicle. These are described in the detailed section of this reviewer's comments. Data presented in the comment's summary section are, in contrast, documented individually in this table.	Vandenberg et al. 2007, pg 16-60
Accurate	Misunderstanding of reliable measure for reproduction status in test animal	Panel criticized study for not documenting an indicator of mating that is not considered reliable for the test animals (the researcher subsequently provided the panel with the appropriate measure).	Schonfelder 2007
Accurate	Failure to find basic information in study documentation	Panel criticized study for not documenting animal numbers when the study documentation clearly stated this information.	Schonfelder 2007
Accurate	Failure to find basic information in study documentation	Panel noted that they estimated study findings from a graph, when the study findings were provided in full in the study documentation.	Schonfelder 2007
Accurate	Error in understanding standard statistics used in study design	Panel criticized a study as having "too few animals to reach a conclusion with certainty" when the study identified statistically significant effects and used the appropriate number of animals needed given the expected magnitude of the effect.	Schonfelder 2007