Report of the Integrity Review Panel

Scientific and Research Misconduct Allegation 2012-001

Executive Summary

This report summarizes the recommendations of the NOAA Integrity Review Panel regarding allegations of Scientific and Research Misconduct by a NOAA employee. The allegations concern conduct that occurred in 2010 prior to the adoption of NOAA's current Scientific Integrity Policy, NAO 202-735D (Policy). Consequently, the Integrity Review Panel (Panel), established by the Deputy Under Secretary for Operations (DUSO), applied the standard for evaluating misconduct that was in effect when the alleged scientific misconduct occurred. Under this standard, misconduct is defined as, "fabrication of results, plagiarism, or clear misstatement of facts."[1]

Allegations of scientific and research misconduct were filed by Mr. Jeff Ruch on behalf of the Public Employees for Environmental Responsibility (PEER), hereinafter referred to as the Complainant, against Dr. William Lehr, a Senior Scientist with the Emergency Response Division of the National Ocean Service's Office of Response and Restoration (NOS/ORR), hereinafter referred to as the Respondent. (For administrative purposes, this allegation is identified as SRMA 2012-0001.) The four specific allegations by the Complainant, filed on January 27, 2012, were as follows:

1. Falsification of scientific findings
2. Failure to objectively consider conflicting findings
3. Prevention of conflicting views from being reported to key decision makers
4. Fabrication of findings and failure to provide traceability of data

All of the allegations are associated with activities that followed the aftermath of the Deepwater Horizon oil spill in 2010. The Respondent led the Plume Calculation Team (Plume Team), one of four teams formed under the auspices of the Flow Rate Technical Group (FRTG). This group reported its findings to the National Incident Command.

To evaluate the allegations, the Panel considered the written testimony of the parties, documents provided by the parties, and relevant information compiled by the Panel members through its own search. (A complete list of information considered is attached.) No oral testimony was solicited or received. The Panel did not reach agreement on Allegations 1 and 2. Two panelists concluded that the available evidence for these allegations do not have sufficient substance to warrant investigation. The third panelist concluded that these allegations have sufficient substance to warrant investigation. The Panel agreed Allegations 3 and 4 did not have sufficient substance to warrant investigation. The Panel worked collaboratively to prepare the main body of the report (Sections 1 to 4) but each panelist prepared independent statements of their recommendations as found in sections 5.1, 5.2, and 5.3.

[1] This standard was informed by the White House Office of Science and Technology Policy guidance on the appropriate standard of conduct applicable to Federal scientists, 65 Fed. Reg. 76260 (December 6, 2000).

The Panel respectfully submits these recommendations to the Determining Official on [date].

1.0 INTRODUCTION

This report addresses allegations of Scientific and Research Misconduct by a NOAA staff member who led the Plume Calculation Team (Plume Team), of the Flow Rate Technical Group (FRTG). The FRTG comprised four separate teams, using a variety of methods to develop estimates of the flow rate of the Deepwater Horizon oil spill in 2010. The alleged misconduct relates to the process of providing information about the estimates to senior decision makers.

The names and affiliations of the Integrity Review Panel (IRP or Panel), all NOAA employees, are given below.

Paul J. Rago, Ph.D.
Integrity Review Panel Chair
Chief, Population Dynamics Branch
Northeast Fisheries Science Center
National Marine Fisheries Service

James E. Hoke, Ph.D.
Director, Hydrometeorological Prediction Center
National Centers for Environmental Prediction
National Weather Service

Marian B. Westley, Ph.D.
NOAA Physical Scientist
Geophysical Fluid Dynamics Laboratory
Office of Oceanic and Atmospheric Research

This report contains the first application of the National Oceanic and Atmospheric Administration's (NOAA) Scientific Integrity Policy NAO 202-735D (Policy). This report does not represent a full application of the Policy since the allegations deal with activities that occurred before the Policy was formally adopted. No basis exists to apply a detailed scientific policy standard retroactively. Instead, the Panel applied the standard of conduct in place at the time of the alleged misconduct. The Panel did, however, follow the procedures and timelines specified in the Procedural Handbook accompanying NOAA's current Scientific Integrity Policy. These procedures establish a rigorous and well-defined approach to evaluating allegations of misconduct. At the conclusion of the inquiry, the Panel did not reach consensus on all of the allegations. As a result, each panelist prepared a separate statement of his/her evaluation of the evidence and summary findings. The individual statements, contained herein, are standardized with respect to description of the allegations, core source material used, and basis for evaluating misconduct. The independent recommendations of each panelist are summarized in Section 4.0. Details of each panelist's recommendations are provided in Sections 5.1, 5.2, and 5.3.

2. COMPLAINANT ALLEGATIONS—SMRA 2012-001 (PEER)

2.1 Allegation 1. Falsification of Scientific Findings: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, and §7.02 [Respondent] intentionally falsified the Scientific Product of the Plume Team by naming his Final Report "*Deepwater Horizon Release Estimate of Rate by PIV*" and by reporting that the majority of the thirteen members of the Plume Team used a technology called Particle Image Velocimetry (PIV) and estimated an oil leak rate of 25,000 to 30,000 bpd. The truth is that only three of the thirteen members of the Plume Team used PIV for their official estimates of the oil leak rate.

2.2 Allegation 2. Failure to Objectively Consider Conflicting Findings: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, and §7.02, [Respondent] intentionally omitted any discussion in his Final Report and Final Presentation about the use of a different technology called FTV by three other members of the Plume Team. The accurate estimates by FTV were in the range of 50,000 to 60,000 bpd, but [Respondent] did not report the estimates to key decision makers or to the public. [Respondent] failed to "objectively consider conflicting data" and failed to "accurately report results" to key decision makers.

2.3 Allegation 3. Prevention of Conflicting Views from Being Reported to Key Decision Makers: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, §7.02 and NOAA's Code of Ethics for Science Supervision and Management, [Respondent] prevented members of the Plume Team who used FTV from communicating their findings to key decision makers. On July 30, 2010, [Respondent] gave the Plume Team's Final Presentation to the team of key decision makers (including DOE Sec. Chu, DOI Sec. Salazar, USGS Dir. McNutt, the Directors of three DOE National Labs, etc.) who were determining the government's final estimate of the oil leak rate. Only the three members of the Plume Team who used PIV and underestimated the oil leak rate were informed of the Final Presentation and allowed to meet with the key decision makers. Members of the Plume Team using FTV were not informed of the Final Presentation. Thus, [Respondent] prevented the members using FTV from meeting with the key decision making team, and prevented "the timely communication of scientific findings" to key decision makers.

2.4 Allegation 4. Fabrication of Findings and Failure to Provide Traceability of Data: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, §7.02, [Respondent] added an additional estimate by PIV from a scientist who was not a member of the Plume Team to his Final Report and Final Presentation. [Respondent] did not reveal to the Plume Team's members, to peer reviewers, to key decision makers, or to the public that he added an estimate from a scientist who was not a member of the Plume Team. It appears that [Respondent]also altered the values of the estimates by PIV to make them appear identical and more "consistent."

3. FACTUAL AND PROCEDURAL BACKGROUND

3.1 Scientific Standards for Evaluation

This matter involves allegations regarding conduct that pre-dates NOAA's current Scientific Integrity Policy (Policy) which was put into effect on December 7, 2011. When the allegations were referred for inquiry, the DUSO explained that the current Policy would not be applied retroactively; rather, the inquiry would be based on standards of conduct in effect at the time of

the alleged misconduct. The following section describes the pre-existing standards applicable for the current inquiry.

Claims against NOAA scientists regarding conduct that occurred prior to December 7, 2011, were evaluated under general principles of scientific misconduct long-understood to apply to professional scientists. Specifically, misconduct is defined as "fabrication of results, plagiarism, or clear misstatement of facts." This standard was confirmed by the DUSO in an email to the Panel Chair on June 6, 2012. The White House Office of Science and Technology Policy (OSTP) provides additional guidance on the appropriate standard of conduct for federal scientists.[2] In 2000, OSTP published guidance on the establishment of research misconduct policies for federally-funded research. That guidance defined a number of terms useful in the administration of research misconduct complaints, and the entirety of the OSTP guidance was useful in informing the current inquiry.

3.2 Evaluation of Allegations

The allegations against the Respondent were evaluated in the context of the unprecedented nature of the oil spill and the government's response to the spill. According to McNutt *et al*. (2011), "At the time of the *Deepwater Horizon* blowout, there were no proven methods for directly measuring the deep sea discharge of hydrocarbons at the relevant pressures and temperatures." The Flow Rate Technical Group (FRTG) was chartered on May 19, 2010, by the National Incident Command Interagency Solutions Group. McNutt et al. (2011) summarized its purpose as follows: "Experts from many scientific disciplines were brought together to perform the FRTG's two primary functions: *(i)* as soon as possible, generate a preliminary estimate of the flow rate, and *(ii)* within approximately 2 months, use multiple, peer-reviewed methodologies to generate a final estimate of flow rate and volume of oil released."

The Panel's understanding of the Respondent's role in the Plume Calculation Team – one of four teams that made up the Flow Rate Technical Group – is as follows. The Respondent was responsible for the following:

- Formed the initial Plume Team and consulted with additional experts to identify appropriate additional members
- Arranged for British Petroleum (BP) to provide video to the team members for analysis
- Convened meetings of the team to develop consensus estimates of the oil spill flow rate for rapid communication to high-level decision makers in the emergency response effort
- Led the writing of the interim report of the team
- Wrote the draft of the body of the final report of the Plume Team (The Plume Team report is dated July 21, 2010)
- Solicited appendices from members of the Plume Team
- Coordinated peer review of the body of the report and its technical appendices

More than seven months later (on March 10, 2011) the FRTG released a final report, which assessed the work of its teams by comparing the results of multiple approaches to the measured flow rate before the well was capped. The Respondent was an author of the FRTG's final report.

[2] 65 Fed. Reg. 76260 (December 6, 2000) "Federal Policy on Research Misconduct."

Subsequently several members of the FRTG, including two members of the Plume Team and the FRTG leader, Marcia McNutt, were co-authors on a paper published in the Proceedings of the National Academy of Science on October 28, 2011. This paper comprehensively compared alternative methodologies to estimate oil flow at the Deepwater Horizon site, "including work not conducted under the auspices of the FRTG" (McNutt et al. 2011). The Respondent is acknowledged among the "numerous experts who read and improved… the manuscript."

3.3 Materials Considered for Evaluation

In its review, the Panel considered the detailed allegations submitted by the Complainant, as well as the written testimony of the parties, documents provided by the parties, and relevant materials identified by the Panel pursuant to its own search (including, among other things, electronic messages published on the U.S. Geological Survey Freedom of Information Act website).

The Panel relied on published documents, the written allegations, and the written testimony to understand the timeline of events. The Respondent and Complainant provided additional written testimony regarding the allegations. There were no oral communications about the substance of the inquiry with Respondent or Complainant. A detailed listing of the evidence is found in Section 6.

4. SUMMARY FINDINGS AND RECOMMENDATIONS OF THE INTEGRITY REVIEW PANEL

As noted in Section 1.0, the Panel did not reach consensus on all of the allegations. Three separate statements, Sections 5.1, 5.2 and 5.3 were prepared and standardized with respect to description of the allegations, core evidence, and standard for evaluating misconduct. Although the sections were written as part of a collaborative process, the statements are the individual views of the Panel members. The summary recommendations of the Panel Members are provided below.

	Summary Recommendations on SRMA 2012-001		
Allegation	***Rago***	***Hoke***	***Westley***
Allegation 1. Falsification of Scientific Findings	No investigation warranted.	No investigation warranted.	Investigation warranted (in part)
Allegation 2. Failure to Objectively Consider Conflicting Findings:	No investigation warranted.	No investigation warranted.	Investigation warranted (in part)
Allegation 3. Prevention of Conflicting Views from Being Reported to Key Decision Makers:	No investigation warranted.	No investigation warranted.	No investigation warranted.
Allegation 4. Fabrication of	No investigation warranted.	No investigation warranted.	No investigation warranted.

Findings and Failure to Provide Traceability of Data			
Summary Recommendation	Insufficient grounds to merit further investigation. Allegations should be dismissed.	Insufficient grounds to merit further investigation. Allegations should be dismissed.	Investigation warranted on part of Allegations 1 and 2.

5. DETAILED FINDINGS BY INTEGRITY REVIEW PANEL MEMBERS

5.1 Recommendation of Integrity Review Panel Chair Paul Rago

I have reviewed the Complainant's allegations and documentation available to the NOAA Integrity Review Panel. The following sections describe my analyses of the allegations. Only the title of the allegation is included herein; a complete listing of the allegations is provided in Section 2.0. Each allegation consists of several assertions. I have chosen to paraphrase the key assertions and respond individually to each. I then follow with a summary statement on each allegation individually and conclude with a summary statement (5.1.5) on all allegations by the Complainant.

5.1.1 Allegation 1. Falsification of Scientific Findings

5.1.1.1 Assertion—Respondent intentionally falsified scientific product by naming it "PIV"

Response: The report was the product of the Plume Team. While it appears that the Respondent wrote the introduction and executive summary, all of the team members had an opportunity to comment on the title and its contents prior to release. Moreover, each team member had the opportunity to prepare individual appendices that allowed each member to clarify the methodology and appropriate nomenclature for their methods. External experts prepared written reviews of each of the Team member's reports but not every External Reviewer had time to comment on each Appendix report. The evidence shows that Team members were afforded the opportunity to provide comments or rebuttals to the assertions of the External experts. Hence there was ample opportunity for the Plume Team to clarify the title of the final report and its organization.

An email from Mr. Shaffer, a member of the Plume Team, to all members of the Plume Team on June 23, 2010, expressed reservations about the use of the automatic PIV methodology but did not make any suggestions about changing the terminology in the final report. Mr. Shaffer did not suggest any wording changes to the report that would have expanded its scope to other related approaches or a more restrictive usage of the term "PIV".

The term PIV is used as a general term to describe the use of video images of turbulent flow to deduce the overall rate of fluid flow. The Final report distinguishes the methodology used in this report from "true PIV" by noting the reliance on "interrogation spots" and a sampling methodology to determine flow. Numerous assumptions are required to estimate flow including

the ratio of gas to liquid, temporal variation in these ratios, and the behavior of gases and liquids under pressure. Throughout the written record, PIV is sometimes used in a restrictive sense to refer to automated software. In other instances, PIV is used in a generic sense to characterize manual analysis of features that may not be as easily identified by an automated computer algorithm. One of the alternative methods, Feature Tracking Velocimetry, essentially uses manual tracking of larger "interrogation spots" which are called features. Mr. Shaffer, in an email response to Dr. Savas on May 26, 2010, notes that

> "the various "PIV" variants we are using are not strictly what engineers would call PIV. But given that the audience will be the general public, I think calling this "PIV" is close enough. Myself, I'm using a new "PIV" technique that does not use any correlations methods such as those used in traditional double-pulse PIV."

5.1.1.2 Assertion—Only three of the thirteen members of the Plume Team used PIV for their official estimates of the oil leak rate.

> *Response:* A review of those Appendices in the Final Report reveals that four of the five authors used PIV or a variant thereof. The identities of the scientists are not given in the final report. An earlier report (June 8, 2010) of Possolo and Espina, statisticians from the National Institute of Standards and Technology, lists estimates from six named experts. The Respondent reported in his written testimony (July 23, 2012) that the identification of individuals was not acceptable to some members. In the final report, Appendix 1, Possolo and Espina do not reveal the identities of the individuals supplying estimates. However, they do reveal that Expert G's estimates arrived on July 13. I also note that McNutt et al. (2011, p. 4 and Table 1) identifies four (not three) experts (A, B, C, and E) who applied PIV analyses. Thus, of the seven distinct estimates of oil flow rate prepared by the Plume Team, four of them used PIV. In normal usage, 4 of 7 constitutes a majority and justifies the use of the word "most."

5.1.1.3 Recommendation on Allegation #1:

I do not believe that the Respondent intentionally falsified the report by using the term PIV in the title of the Plume Team's Final Report. Distinctions among the methods, as described in the Appendices, are important but not critical for conveying the uncertainty among the various methods and investigators.

The available evidence suggests this allegation does not have sufficient substance to warrant investigation.

5.1.2 Allegation 2. Failure to Objectively Consider Conflicting Findings

5.1.2.1 Assertion—Respondent intentionally omitted any discussion in his Final Report and Final Presentation on the use of a different methodology.

Response: Overall, the information in the Final Report was thorough and allowed for a full exposition of the alternative methods by the individual authors in the appendices. The Final

Report was prepared and peer-reviewed by six external experts. Plume Team members had the opportunity to provide comments to the reviewers before preparation of the final draft. It appears that the report was prepared between June 13 when the final estimates were submitted to the NIST statisticians and July 21 when the final report was issued.

The final report and the appendices were reviewed by external experts. None of their comments, including those critical of the main body of the report, were edited. Team authors also had the opportunity to respond to the external experts. The Final Report was made available to decision makers before the "Final Presentation" was made on July 30, 2010. Most of the appendices summarize methods that depart from the more restrictive use of the term PIV. It should be noted that even within the methods labeled as FTV (Feature Tracking Velocimetry) the exact methodology differs among investigators. This is to be expected since the identification of features relies on "the human brain as an expert system to painstakingly choose large and fast features to track" (McNutt et al. 2011). Most of the Final Report is devoted to highlighting the range of scientific viewpoints rather than quashing them.

5.1.2.2 Assertion—The accurate estimates by FTV were in the range of 50,000 to 60,000 bpd, but Respondent did not report the estimates to key decision makers or to the public.

Response: The "accuracy" of the FTV estimates was not known at the time the report was prepared. The accuracy of the higher estimates was first documented in the primary scientific literature with the publication of the McNutt et al (2011) report, a full five months after the Plume Team completed it work.

The key decision makers, supported by FRTG of which the Plume Team was a part, released an official estimate of 35,000 to 60,000 on June 15, 2010. This was more than a month before the Final Report of the Plume Team was released. The key decision makers referred to in the allegation had already made a decision on the magnitude of leak, and the final report of the Plume Team had been available for more than a week before the Final Briefings on July 30 and July 31.

5.1.2.3 Assertion—Respondent failed to "objectively consider conflicting data" and failed to "accurately report results" to key decision makers.

Response: This assertion is not supported by the evidence. The final report of the Plume Team had all of the estimates, including those derived by the Plume Team, and those subsequently derived by individual members in their separate reports in the Appendices. The Final Report of the Plume Team was made available to key decision makers.

5.1.2.4 Conclusion on Allegation #2: The July 21, 2010, final report of the Plume Team incorporated a summary of the consensus views of the team, individual appendices prepared by individual investigators, a complete set of unedited reviews by six independent experts, and rebuttals to the reviewers by individual team members. The scientific debate was fully documented. The process reflects a high degree of scientific transparency uncommon in the peer-review literature. The full documentation of the Plume Team's findings was available in a written form to decision makers and their staff before the verbal briefings on July 30-31, 2010.

The available evidence suggests this allegation does not have sufficient substance to warrant investigation.

5.1.3 Allegation 3. Prevention of Conflicting Views from Being Reprorted to Key Decision Makers.

5.1.3.1 Assertion—Respondent prevented members of Plume Team who used FTV from communicating their findings to key decision makers

Response: The evidence suggests that Mr. Shaffer was permitted to participate in a high level briefing to Secretaries Chu and Salazar on June 14, but Shaffer was not chosen to serve as the lead for this presentation. Selection of the leads for this presentation was made by Dr. Marcia McNutt, not the Respondent. Details on justification for this decision are provided in email correspondence between McNutt, the Respondent, Shaffer, and other Plume Team members on or about June 17, 2010. McNutt argued that Shaffer had not met the standards of openness, documentation, and peer review within the Plume Team to justify his serving as lead presenter of the Team's findings. The direct presentations of Lasheras and Savas represented the two methodologies – PIV and FTV.

The July 30 and July 31 meetings were a combination of face-to-face meetings with key decision makers and a conference call for individuals who could not attend. The meeting was organized by Sandia National Labs not the Respondent. The actual list of participants is not known to the Panel.

There is a factual error in Allegation 3. The Respondent was not present at the meeting. The Respondent contributed to the presentation, but it was delivered by other Plume Team members. Omission of FTV estimates in the July 30 presentation appears to be unintended because the report of the Plume Team was already a matter of public record, verifiable by all in attendance.

5.1.3.2 Assertion—Members who used FTV were not informed of the final presentation.

Response: The referenced meeting was organized by staff at Sandia Lab, not the Respondent, so the Complainant's attribution of this conduct to the Respondent is not supported by the record.

5.1.3.3 Conclusion on PEER Allegation #3

Participation in the presentations to various meetings with decision makers was not controlled by the Respondent. E-mail evidence provided by the Respondent confirms that Mr. Shaffer did not make the presentation on June 14 for the Team because other Team members, representing both PIV and FTV methods, were more prepared to report the findings of the Plume Team.

The available evidence suggests this allegation does not have sufficient substance to warrant investigation.

5.1.4 Allegation 4. Fabrication of Findings and Failure to Provide Traceability of Data

5.1.4.1 Assertion—Respondent added an additional estimate by PIV from a scientist who was not a member of the plume team, coerced members to alter their data and did not inform other Team members or public of this decision.

Response: Individuals who made estimates in the Table on 15 are not identified because several Team Members objected to an earlier report (June 8, 2010) that provided the first names of each individual. The Respondent reported in his written statement (July 23, 2012) that the identification of individuals was not acceptable to some members. The principle of anonymity was also applied in the McNutt et al (2011) publication in PNAS, a publication co-authored by Savas and Shaffer, both of whom were on the Plume Team.

In Appendix 1 of the July 21, 2010 Final Report of the Plume Team, the NIST statisticians Possolo and Espina analyzed data from from a total of 7 individuals. Five of these individuals provided estimates at the June 13th meeting of the Plume Team in Seattle. Two additional estimates that arrived after that. As noted in Appendix 1 of the Final Report (p. 16), Expert F's estimate was extracted from his contribution to the final report. Possolo and Espina refer to Expert F'smethod as manual imaging velocimetry. Expert G's estimate arrived in a separate transmittal on July 13, 2010. Contrary to Allegation 4, the origin of the estimates is provided in the Final Report. No addition was made by the Respondent or NIST Plume Team members. It should be noted that the June 8 report of Possolo and Espina contained six estimates, not five. Possolo and Espina note that estimates from Shaffer were not included because other experts were estimating average volume of oil spilled while Shaffer was estimating maximum volume.

A primary piece of evidence by the Complainant is the difference in the number of rows in table presented to decision makers on July 30 from the table included in the Final Report of the Plume Team published on July 21, 2010. The Respondent acknowledged that he made a cut and paste error in the slide presentation and failed to use the final table from page 15 of the report. I note that there is no difference in the values for the first 5 rows of the tables, supporting the Respondent's claim that these values were a cut and paste error. It is also noted that both of the estimates in question for Experts F and G were not available on June 13, a critical date corresponding to a meeting of the Plume Team that preceded a teleconference call and meeting with Secretaries Chu and Salazar on June 14. Hence, the chance that an earlier tabular summary could have replaced the Appendix 1 table seems probable. Finally, I note that the Final Report of the Plume Team was available to all decision makers well in advance of the July 30 and 31 meetings.

5.1.4.2 Assertion—Respondent altered the estimates by PIV to make them appear identical and more consistent.

Response: The methodology to analyze the Plume Team's estimates is clearly specified in Appendix 1 of the Final Report. Appendix 1 was prepared by Plume Team statisticians Possolo and Espina. They used a well-established method to combine separate estimates into a composite value. The methodology takes into account the mean and variance of the individual estimates and the qualitative "degree of confidence" the authors had in their estimates during a June 14 conference call. An additional assumption of Possolo and Espina was that the estimates were normally distributed. They also adjusted the estimates for conversions related to thermodynamics

and oil to gas ratios. The composite distribution of estimates is polymodal and highly skewed with a heavy right hand tail with values ranging up to 123,000 bpd. The lower tail of the composite distribution excludes estimates that fall below the volume of oil actually being recovered from the Top Hat (24,000 bpd). Hence the statisticians, not the Respondent, concluded that truncation of estimates below 24,000 bpd was justified. McNutt et al (2011) also included this truncation principle in their evaluation of the flow rates. I also note that left truncation of the composite distribution results in an overall estimate of discharge that is higher than would be obtained otherwise.

5.1.4.3 Conclusion on PEER Allegation #4

The decision to not identify the individuals associated with each estimate appear to be a joint decision by the Plume Team. There was adequate time to have raised this issue within the team prior to publication of the Final Report. No additional estimates advocating the PIV method were included. All of the Plume Team members who made estimates are identified on Page 1 of the Final Report. Differences in estimates attributable to the Respondent were in fact the results of decisions made by Team Members from NIST (Possolo and Espina).

The available evidence suggests this allegation does not have sufficient substance to warrant investigation.

5.1.5 Summary Statement

The available evidence suggests these allegations in total do not have sufficient substance to warrant investigation.

The OSTP Executive Order (Federal Register Vol 65, No. 235 pp 76260-76264) instructs that a finding of misconduct requires that 1) There be a significant departure from accepted practices of the relevant research community; and 2) The misconduct be committed intentionally, or knowingly, or recklessly; and 3) The allegation be proven by a preponderance of evidence. The preponderance of the evidence does not suggest deliberate manipulation of facts or attempts to mislead decision-makers. In fact, all of the Plume Team statements and reports are explicit in their characterization of the findings as preliminary and as not official. The timing of events does not support intentional actions to mislead. All official statements of the Plume Team were highly qualified with respect to the uncertainty of the estimates and their intended usage by decision makers.

In hindsight, the Plume Team's final report could have been improved by resolving the disparity between the details of the main report and appendices. The addition of several paragraphs in the final report could have clarified the process of consensus building followed by the Chairman and the Plume Team. In particular, the Final Report would have benefited from a description of the relationship between the range of estimates listed in the Executive Summary (p. 3), Conclusions (p. 13), and the individual Appendices. External reviewers 2 and 5 also noted this disparity. The Respondent's response (p. 200) to those reviewers was instructive. The Respondent notes:

> "The introductory background material was not intended to be a summary of the individual appendices, but rather to provide a basic explanation of the PIV method and to record the consensus results. This consensus was reached prior to the documenting material in the appendices and changing the section to include them would not accurately represent the process."

Hence, the introductory section summarizes the proceedings and conclusions of the Plume Team at the time they occurred. The Appendices were prepared after the Plume Team had met and were not necessarily reflective of the Team's plenary deliberations. The evidence suggests that interactions among Team were primarily with the reviewers rather than with other Team members on the contents of the main report. Inclusion of a timeline of the plenary and conference call meetings and the timing of events for preparation of the Final Report would have identified the major decision points.

The PowerPoint presentation for the July 30, 2010, meeting was prepared after the Final Report had been made available to the public and senior managers. The presentation was labeled as "Predecisional Drafts" and for "Official Use Only." Deliberate contradiction of the written record in a verbal presentation would not be prudent or logical since its basis could be established by any number of participants, including members of the Plume Team. Moreover, the Secretaries of Energy and Interior had already issued a joint statement on the government's official estimate of 35,000-60,000 bpd on June 15, 2010, about 45 days prior to the July 30 briefing. Following the July 31 conference call, the Secretary of Energy released a final estimate of 62,000 bpd. Thus, it appears that the high-level decisions had been made before the Complainant alleges the misconduct occurred.

Finally, I note the consistent pattern of qualifying statements in the written reports and statements of the Plume Team and its Chair, the Respondent. This is reflected in the 1) Interim Report of the FRTG on May 27, 2010, 2) an intermediate statement by Respondent on June 10, 2010, and 3) the final report of the Plume Team on July 21, 2010.

5.2 Recommendation of Integrity Review Panelist James Hoke

I have studied the four allegations by the Complainant, as well as the relevant documentation available to the NOAA Integrity Review Panel at this time. The available evidence suggests none of these allegations has sufficient substance to warrant investigation.

Before reviewing the individual allegations, I would like to discuss a number of points that are fundamental to my findings and recommendation on this matter.

5.2.1 Fundamental Principles

Standard for judging misconduct. As described in the introduction above, the standard used by the panel in judging misconduct regards fabrication of results, plagiarism, and clear misstatement of facts.

Standard of proof. The standard I used in this inquiry was "preponderance of the evidence" and not "beyond reasonable doubt." This is the standard set forth in "Federal Policy on Research Misconduct," 65 Fed. Reg. 76262.

Emergency situation. Deepwater Horizon failed catastrophically on April 20, 2010, and the Macondo well was capped on July 15, 2010. The Final Report of the Plume Team was issued just six days later. Thus, the vast majority of the work leading to the Final Report was conducted during extreme emergency conditions requiring rapid responses to questions and other requests with short deadlines and necessitating critical decisions without the luxury of a great amount of time. The foremost role of the Plume Team was to advise the National Incident Command and not to conduct a thorough scientific investigation.

Particle Image Velocimetry (PIV). In general, PIV is a technique for determining fluid flow based on the analysis of sequential video images. There appeared to be two very different definitions of PIV, however, used in the evidence. In one definition, PIV was used generically to refer to both automated and manual techniques. (An example of a manual technique is Feature Tracking Velocimetry – FTV.) Alternately, the term PIV was used to refer specifically to the automated techniques only and not the manual techniques. Thus, one definition was a subset of the other. The Plume Team used PIV in both contexts without always clarifying which definition was being used. As a result, there was a great deal of opportunity for misunderstanding and for confusion among the Plume Team members and others as to the meaning of PIV.

Definition of consensus. In his statement of July 23, 2012, to Dr. Rago, Chair of the Integrity Review Panel, the Respondent outlines the process by which the Plume Team reached consensus on its estimates of flow rate.

> "The team was tasked by the National Incident Command to produce estimates of the flow rate at specific times, often while the members were still trying to process that data. The approach that I and Dr. McNutt adopted to generate these numbers was to call a meeting or teleconference of the flow experts and allow them to

> reach a consensus estimate that at least a majority of the team could support. I would then draft the language describing their conclusions for release and get approval from the team before submission to the FRTG Head."

A key point is that consensus was defined as the majority for the purposes of the Plume Team. It would have been helpful if the Final Report had explicitly explained that process. Also, in retrospect, removed in time from the stress of the international emergency, one might ask whether it would have been preferable to have used a higher level of agreement.

Final Report Disconnect. The body of the Plume Team's Final Report and the report's appendices do not complement each other. In typical reports, the appendices often provide additional information deemed too detailed for inclusion in the body of the report. In the Plume Team's Final Report, however, there were appendices involving FTV for which there was no clear connection to the body of the report. This disconnect is explained on page 200 of the Final Report in the following statement attributed to the Respondent.

> "The introductory background material was not intended to be a summary of the individual appendices, but rather to provide a basic explanation of the PIV method and to record the consensus results. This consensus was reached prior to the documenting material in the appendices and changing the section to include them would not accurately represent the process."

5.2.2 Analysis of Allegations

There were four allegations of scientific misconduct presented by the Complainant. The following is my analysis of each.

Allegation 1. *Falsification of Scientific Findings: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, and §7.02, [the Respondent] intentionally falsified the Scientific Product of the Plume Team by naming his Final Report "Deepwater Horizon Release Estimate of Rate by PIV" and by reporting that the majority of the thirteen members of the Plume Team used a technology called Particle Image Velocimetry (PIV) and estimated an oil leak rate of 25,000 to 30,000 bpd. The truth is that only three of the thirteen members of the Plume Team used PIV for their official estimates of the oil leak rate.*

As pointed out in the Fundamental Principles above, the term PIV was used in two contradictory ways by the Plume Team. The resultant confusion between the two definitions in my opinion rendered irrelevant the distinction between the two definitions to decision makers and the public. Also, the evidence indicates the majority of the Plume Team supported PIV as defined in its limited context as specifically an automated technique. Results from FTV (a type of PIV in the broader context) were presented in the appendices of the Final Report. Based on all those considerations, I believe referring to PIV in the title of the Final Report may have been confusing, but it was not deceptive. Also, I did not find evidence in the Final Report or the Final

Presentation indicating the Respondent reported the Plume Team "used" PIV. The available evidence suggests this allegation does not have sufficient substance to warrant investigation.

Allegation 2. *Failure to Objectively Consider Conflicting Findings: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, and §7.02, [the Respondent] intentionally omitted any discussion in his Final Report and Final Presentation about the use of a different technology called FTV by three other members of the Plume Team. The accurate estimates by FTV were in the range of 50,000 to 60,000 bpd, but [the Respondent] did not report the estimates to key decision makers or to the public. [The Respondent] failed to "objectively consider conflicting data" and failed to "accurately report results" to key decision makers*.

Conflicting findings were discussed during the deliberations of the Plume Team. The consensus of the Plume Team was presented in the body of the Final Report and in the Final Presentation. Given the definition of consensus used in determining Plume Team conclusions and the process used to reach that consensus discussed in the Fundamental Principles above, it does not seem unreasonable for the body of the Final Report and the Final Presentation to focus on the consensus results. The Final Report did include appendices providing results from FTV, so the information was available to decision makers and the public who might have wanted to dig into those details. The available evidence suggests this allegation does not have sufficient substance to warrant investigation.

Allegation 3. *Prevention of Conflicting Views from Being Reported to Key Decision Makers: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, §7.02 and NOAA's Code of Ethics for Science Supervision and Management, [the Respondent] prevented members of the Plume Team who used FTV from communicating their findings to key decision makers. On July 30, 2010, [the Respondent] gave the Plume Team's Final Presentation to the team of key decision makers (including DOE Sec. Chu, DOI Sec. Salazar, USGS Dir. McNutt, the Directors of three DOE National Labs, etc.) who were determining the government's final estimate of the oil leak rate. Only the three members of the Plume Team who used PIV and underestimated the oil leak rate were informed of the Final Presentation and allowed to meet with the key decision makers. Members of the Plume Team using FTV were not informed of the Final Presentation. Thus, [the Respondent] prevented the members using FTV from meeting with the key decision making team, and prevented "the timely communication of scientific findings" to key decision makers.*

On June 14, 2010, a member of the Plume Team who was an FTV advocate; Dr. Marcia McNutt, Flow Rate Technical Group (FRTG) lead; and several others met with Department of Energy Secretary Steven Chu and Department of the Interior Secretary Kenneth Salazar. The Respondent was not in attendance. I am aware of no reason why the FTV advocate from the Plume Team would not have had the opportunity at that meeting to communicate his thoughts to the key decision makers present. In addition, I see no reason why the FTV advocate from the Plume Team could not have conveyed his thoughts at any time to Dr. McNutt in her capacity as the lead of the FRTG, which had as a subgroup the Plume Team in which the FTV advocate served.

Contrary to the allegation, the Respondent did not attend the meeting of July 30, 2010. According to the Respondent's written testimony (dated August 14, 2012), the meeting was organized by Sandia National Laboratories staff, who among other things were responsible for establishing the agenda. According to the Respondent, the organizers requested the Plume Team presentation "give the essence of the model or experimental technique." As the Respondent did not have the time to prepare the presentation and was unable to attend, he requested Plume Team members Dr. Steven Wereley and Dr. Alberto Aliseda be responsible for preparing and delivering the presentation, which focused on the consensus results of the Plume Team. (The Respondent provided several slides for the presentation for that meeting.) Given the time constraints on decision makers at that meeting, it seems reasonable the presentation would focus on the consensus results of the Plume Team, that persons supporting the majority opinion of the team would give the presentation, and that attendance at the meeting would be limited to those deemed essential to the decision-making process.

As mentioned in the preceding paragraph, the Respondent provided several slides for the presentation of July 30, 2010. In his statement of July 23, 2012, the Respondent says he accidentally made a mistake in preparing one of the slides. He said he unintentionally missed the estimates from an FTV advocate and one other Plume Team member, perhaps by the Respondent's use of an old version of a document. Thus, the Respondent's intent was to include FTV results in the Final Presentation, but it did not happen because of the error.

The available evidence suggests this allegation does not have sufficient substance to warrant investigation.

Allegation 4. *Fabrication of Findings and Failure to Provide Traceability of Data: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, and §7.02, [the Respondent] added an additional estimate by PIV from a scientist who was not a member of the Plume Team to his Final Report and Final Presentation. [The Respondent] did not reveal to the Plume Team's members, to peer reviewers, to key decision makers, or to the public that he added an estimate from a scientist who was not a member of the Plume Team. It appears that [the Respondent] also altered the values of the estimates by PIV to make them appear identical and more "consistent."*

On page 11 of the Complainant's allegations, there is reference to a table on page 15 of the Plume Team Final Report. That table is used as the basis for the allegation that the Respondent added an estimate from a scientist not on the Plume Team. That table, however, resides in the Final Report on the first page of Appendix 1: "NIST Statistical Analysis" and therefore was developed by NIST and not by the Respondent.

Additional detail regarding the essence of the last sentence of Allegation 4 above is found on page 12 of the allegations. The Complainant states "it appears that the actual numerical values of estimates in the table presented [*sic*] by [the Respondent] in his Final Presentation have been changed to make them appear more 'consistent'." That sentence appears to be in error because the results for the first five of the seven experts in the NIST-prepared table on page 15 of the Final Report are identical to those of the five experts in the table in the Final Presentation. (As explained in the preceding paragraph, the results for the sixth and seventh experts were accidentally not included in the table of the Final Presentation.)

The available evidence suggests this allegation does not have sufficient substance to warrant investigation.

Section 5.3 Recommendation of Integrity Review Panelist Marian Westley

I have studied the four allegations by the Complainant, as well as the relevant documentation available to the NOAA Integrity Review Panel at this time. To me, the available evidence suggests that parts of two of these allegations have substance, and therefore I recommend an investigation.

Before reviewing the individual allegations, I would like to discuss a number of points that are fundamental to my position on this matter.

5.3.1 Fundamental Principles

The purpose of the inquiry. As described in the *Procedural Handbook for NOAA Administrative Order (NAO) 202-735D: Scientific Integrity* Section 5.02(a), "The purpose of the inquiry phase is to assess whether a Scientific and Research Misconduct allegation has substance and to determine whether an investigation is warranted."

Standard for making a decision. The Panel agreed to apply the standard described in the Federal Policy on Research Misconduct that was in place at the time of the alleged misconduct (see 65 Fed. Reg. 76262): "Research misconduct is defined as fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results… Falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record... Research misconduct does not include honest error or differences of opinion." The standard of proof is "a preponderance of the evidence," not "beyond a reasonable doubt," and I note that in this inquiry phase, we are reviewing the preponderance of the evidence that the allegation has substance and warrants investigation, not necessarily the preponderance of the evidence that scientific misconduct occurred. Applying this standard, my choices for each allegation were: "To me, the available evidence suggests the allegation of misconduct has no substance" or "To me, the available evidence suggests the allegation of misconduct has substance, and therefore I recommend an investigation."

Context. The allegations against the Respondent must be evaluated in the context of the unprecedented nature of the oil spill and the government's response to the spill. According to McNutt *et al.* (2011): "At the time of the *Deepwater Horizon* blowout, there were no proven methods for directly measuring the deep sea discharge of hydrocarbons at the relevant pressures and temperatures." The Flow Rate Technical Group (FRTG) was chartered in May 2010: "Experts from many scientific disciplines were brought together to perform the FRTG's two primary functions: *(i)* as soon as possible, generate a preliminary estimate of the flow rate, and *(ii)* within approximately 2 months, use multiple, peer-reviewed methodologies to generate a final estimate of flow rate and volume of oil released." I note that these two functions are

somewhat different: one is primarily an emergency response function and the other is primarily a research function, and modes of behavior that might facilitate the first function (e.g. the use of expert judgment, an emphasis on timeliness over accuracy, and the reliance on established techniques wherever possible), may not be well-suited to the second function. The Plume Calculation Team was part of the FRTG and shared these two functions. I interpret the interim flow rate statements made by the Plume Calculation Team as the Team's execution of the first function. The last flow rate statement by the Plume Calculation Team was included in a June 15, 2010, press release. I interpret the July 21, 2010, release of the report, *Deepwater Horizon Release Estimate of Rate by PIV* (hereinafter referred to as the *Final Report*), as the Plume Calculation Team's contribution to the second function of the FRTG, which was fulfilled with the March 10, 2011, release of the FRTG's report, *Assessment of Flow Rate Estimates for the Deepwater Horizon /Macondo Well Oil Spill* (hereinafter referred to as the FRTG *Assessment*).

The Plume Calculation Team *Final Report* contains a 12-page main report, with flow rate estimates provided in the Executive Summary and Conclusions sections, and 200 pages of appendices. As my fellow Panelists have noted, there are clear disparities between the material summarized in the main report and the full range of estimates and approaches included in the appendices. The FRTG *Assessment* includes a section on the Plume Calculation Team's work that appears to be based on the material in the Plume Calculation Team main report, and includes the Plume Calculation Team main report as an appendix. Therefore, decisions concerning what material to include and exclude from the Plume Calculation Team main report affected the representation of the Plume Calculation Team's findings in the FRTG *Assessment*.

As oil and gas exploration reaches into ever more extreme and remote environments (e.g. the Arctic continental shelf), NOAA must be ready to respond to environmental emergencies in which standard approaches are likely to prove inadequate and technologies essential to the response and recovery effort may be developed and refined in real time. It is likely that future emergency responders will be faced with the same dual charge that the FRTG faced: to develop meaningful information as quickly as possible while also performing the research necessary to facilitate future emergency response efforts. NOAA and our federal partners will require the ability to enlist expertise from other government agencies and from academia (as was done in this case), to conduct an effective response using blended teams with experts from multiple sectors representing a wide range of experiences, and to understand and communicate uncertainty and minority viewpoints. The work of the FRTG is likely to form a precedent for the government's response to future environmental emergencies, and it is my conviction that NOAA would be well-advised to study the experience of the Plume Calculation Team and learn from the strengths and weaknesses of the processes used by the Team during the Deepwater Horizon/Macondo oil spill.

5.3.2 Analysis of Allegations

There were four allegations of scientific misconduct presented by the Complainant. The following is my analysis of each.

Allegation 1. *Falsification of Scientific Findings: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, and §7.02, [Respondent] intentionally falsified the Scientific Product*

of the Plume Team by naming his Final Report "Deepwater Horizon Release Estimate of Rate by PIV" and by reporting that the majority of the thirteen members of the Plume Team used a technology called Particle Image Velocimetry (PIV) and estimated an oil leak rate of 25,000 to 30,000 bpd. The truth is that only three of the thirteen members of the Plume Team used PIV for their official estimates of the oil leak rate.

This allegation results from the disparities between the material summarized in the main body of the Plume Calculation Team *Final Report* and the full range of estimates and approaches included in its appendices. I take the allegation concerning the title of the report seriously for two reasons: 1. naming a report is a very powerful way of communicating the content of a report and 2. the name then propagated into other communications referencing the work of the Plume Calculation Team, such as FRTG *Assessment*, which includes a section named "Video PIV Analysis" and an Appendix named "Plume Calculation Team 2010; Particle Image Velocimetry Report."

There are ten appendices in the Plume Calculation Team *Final Report*, only one of which includes the term PIV in its title, and even in this case, the author is careful to state that his approach "would be more properly classified as correlation-based feature tracking" (*Final Report* p. 57). Of the five appendices dedicated to estimates of the oil flow rate made by analyzing video of the plume, three include lengthy documentation of the failure of Particle Image Velocimetry when applied to analysis of the Deepwater Horizon oil leak and instead provide estimates using alternative approaches. Thus the title *Deepwater Horizon Release Estimate of Rate by PIV,* does not seem appropriate to the work of the Plume Calculation Team.

In his testimony to the Panel, the Respondent justifies the use of the term PIV in the title by quoting from an email by team member Franklin Shaffer: "I agree that the various "PIV" techniques we are using are not strictly what engineers would call PIV. But given that the audience will be the general public, I think calling this "PIV" is close enough." I find this defense unsatisfactory since the email was sent on May 26, 2010, and cannot therefore refer to a draft of the *Final Report* (which was submitted for review in late June). Furthermore, the email was included in a chain of emails in which several other members of the Plume Calculation Team appear to express reservations about the term PIV, including this statement from team member Alberto Aliseda: "I think we should clarify that we are not doing PlV, but rather using PlV algorithms to obtain velocity measurements from the features on the surface of the jet" (see email from Steven T. Wereley to Alberto Aliseda and others, "RE: draft report - two notes from Omer", sent on 26 May, 2010.) The Plume Calculation Team Scientific Product associated with the time period of the cited email did not include the term PIV in its title (see *Estimated Leak Rates and Lost Oil from the Deepwater Horizon Spill, May 27, 2010, Interim Report to the Flow Rate Technical Group*). Therefore, I find the use of the term PIV in the title of the final report to be inaccurate, and I find it difficult to believe that the Plume Calculation Team agreed to this wording. Further investigation would be necessary to determine if including the term PIV in the title of the Plume Calculation Team final report constituted scientific misconduct by the Respondent.

The Complainant further alleges that the Respondent intentionally falsified the Scientific Product of the Plume Calculation Team "by reporting that the majority of the thirteen members of the

Plume Calculation Team used a technology called Particle Image Velocimetry (PIV) and estimated an oil leak rate of 25,000 to 30,000 bpd." In my view, this part of the allegation does not warrant investigation under the standards that we are using since PIV is referred to in the report as "the main method used," a statement that contains no information on the number of people who used the method, and the diligent reader can trace the 25,000 to 30,000 bpd leak rate estimate to a June 10, 2010, press release from the National Incident Command.

In summary, for the part of Allegation 1 that refers to the title of the Final Report, the evidence suggests to me that the allegation has substance and therefore I recommend an investigation.

Allegation 2. *Failure to Objectively Consider Conflicting Findings: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, and §7.02,[Respondent] intentionally omitted any discussion in his Final Report and Final Presentation about the use of a different technology called FTV by three other members of the Plume Team. The accurate estimates by FTV were in the range of 50,000 to 60,000 bpd, but [Respondent] did not report the estimates to key decision makers or to the public. [Respondent] failed to "objectively consider conflicting data" and failed to "accurately report results" to key decision makers*.

Falsification includes "changing or omitting data or results such that the research is not accurately represented in the research record" (see OSTP Policy of Research Misconduct, 65 Fed. Reg. 76262). A major finding after the Deepwater Horizon oil spill was that "The method of automated PIV, used by several groups of experts during the spill to analyze video segments, was inappropriate for the application and resulted in oil flow rates that were biased too low by a factor of two" (McNutt *et al.* 2011). Much of the information needed to reach this finding was available in the Appendices of the Plume Calculation Team *Final Report*. In a report dated June 15, 2010, and included in Appendix 4 of the Plume Calculation Team *Final Report*, Ömer Savaş wrote: "During the teleconference on June 10, 2010, it was clear that I must provide details for my concerns and bases for my opinions." His report described in detail the limitations of PIV as applied to the oil spill (see page 47 of the *Final Report*). His statement provides evidence that the team was discussing concerns with the PIV approach during team meetings as early as June 10, 2010. While the body of the Plume Calculation Team *Final Report* lists some of the issues with PIV analysis that could apply to any video analysis of the plume (the flow was not spatially or temporally uniform, the ratio of oil to gas in the plume was not known with certainty, some of the video was of excessively poor quality), there is no indication that three of the "PIV experts" had tried PIV, realized that its shortcomings were insurmountable, and developed alternative approaches. I regard the failure to fully represent the findings of the team to be a major defect in the Plume Calculation Team *Final Report*.

The flow rate estimates provided in the main body of the Plume Calculation Team *Final Report* do not encompass the full range of values provided by the Plume Calculation Team. The Conclusions section of the *Final Report* states that for the video provided after the riser was cut, "The best estimate of the PIV experts was for a flow of 35,000 to 45,000 bbl with the possibility that the leak could be as high as 50,000 bbl day." This range does not capture the flow rate estimates provided in two Appendices: 62,500-68,000 bbl/day (with uncertainties larger than ± 50%) in Appendix 6 and 61,000 bbl/day ± 15,000 bbl/day in Appendix 7. It is not clear how the estimate cited in the Conclusions of the *Final Report* was developed or when. While the full

range of values is available to diligent readers in the Appendices to the Plume Calculation Team *Final Report*, it is the estimates provided in the main body of the Plume Calculation Team *Final Report* that appear in the March 10, 2011, report of the full FRTG as part of the heading on the Plume Calculation Team's work: "Video PIV Analysis (Plume Calculation Team 2010, Appendix D): 25,000 to 30,000 BPD (pre-riser cut), 35,000 to 50,000 BPD (post-riser cut)" (see page 11 of the FRTG *Assessment*). Appendix D of the FRTG *Assesssment* is the Plume Calculation Team *Final Report* minus appendices, which are mentioned in this note on page 21 of the FRTG *Assessment*: "Due to the length of the full Plume Calculation Team report, this appendix includes only the summary section. The full report can be downloaded at: http://www.usgs.gov/oilspill/ and http://www.doi.gov/deepwaterhorizon/index.cfm." In other words, a reader who begins by reading the FRTG *Assessment* would have to be diligent indeed to find the full range of estimates provided by the Plume Calculation Team since they are not included in the "summary section."

Reviewer 5 of the Plume Calculation Team *Final Report* called attention to the report's opacity: "I found the report itself to be very weak. Basically, it relies on the reader to sort through all of the appendices to understand the numbers presented. Specifically, it is not clear where the numbers presented in the Executive summary and in the body of the report come from nor what they mean…Had the report presented a coherent summary of the work presented in the appendices it would have been a lot easier to read and more believable." (See page 193 of the *Final Report*).Reviewer 2 provided several concrete suggestions for improving the final report, concluding, "I would like each set of results presented in one final table in the Overview" (see page 180 of the Final Report).

Rather than accept this constructive criticism and edit the report, the Respondent replied that he took "exception to some comments by reviewers 2 and 5. The introductory background material was… [intended to] record the consensus results. This consensus was reached prior to the documenting material in the appendices and changing the section to include them would not accurately represent the process" (see page 200 of the Plume Calculation Team's Final Report). While respect for team process is laudable, the reader is still left with no understanding of what that process entailed. Was the "best estimate" a consensus statement, and if so, what was the definition of consensus and how were views that fell outside of the consensus represented in the report? At what point in time was each consensus reached, considering that team members were changing their estimates not only as new video became available to them, but also as they reached new understanding of the strengths and weaknesses of the methods they were applying to their analyses? Because the Respondent received clear feedback from two reviewers concerning the weakness in the report, and because the Respondent had time to respond to these reviewers, I am left with the impression that the Respondent chose to leave the report in its current, confusing state in order to marginalize the contributions of certain team members and to hamper the communication of higher flow rate estimates to key decision makers and to the public.

In support of this allegation, the Complainant provides a copy of the slides prepared for a presentation of the Plume Calculation Team's findings given to Secretaries Chu and Salazar on July 30, 2010. The first slide lists the Respondent as the first author, with two other members of the Plume Calculation Team also listed as authors. The presentation contains a single results

slide in which the results from the authors of Appendices 6 and 7 are missing. The Respondent has explained to the Panel that he was not present at the July 30, 2010, meeting and that the presentation was given by two other members of the Plume Calculation Team. He admits that he prepared the results slide and asserts that the omissions were unintentional, perhaps the result of working under excessive time pressure. Since scientific misconduct implies knowing and willful falsification of results and not genuine error, I do not believe that the flawed results slide is evidence of scientific misconduct.

In summary, for the part of Allegation 2 that refers to the Plume Calculation team *Final Report*, the available evidence suggests to me that the allegation has substance and therefore I recommend an investigation.

Allegation 3. *Prevention of Conflicting Views from Being Reported to Key Decision Makers: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, §7.02 and NOAA's Code of Ethics for Science Supervision and Management,[Respondent] prevented members of the Plume Team who used FTV from communicating their findings to key decision makers. On July 30, 2010, [Respondent] gave the Plume Team's Final Presentation to the team of key decision makers (including DOE Sec. Chu, DOI Sec. Salazar, USGS Dir. McNutt, the Directors of three DOE National Labs, etc.) who were determining the government's final estimate of the oil leak rate. Only the three members of the Plume Team who used PIV and underestimated the oil leak rate were informed of the Final Presentation and allowed to meet with the key decision makers. Members of the Plume Team using FTV were not informed of the Final Presentation. Thus, [Respondent] prevented the members using FTV from meeting with the key decision making team, and prevented "the timely communication of scientific findings" to key decision makers.*

This allegation and the supporting documentation focus on two issues: the first is attendance at a July 30, 2010, meeting with Secretaries Chu and Salazar. The second is alleged efforts by the Respondent "to discredit and remove members from the Plume Calculation Team who reported that PIV was making mistakes and underestimating the leak rate."

According to the Respondent, the July 30, 2010, meeting was organized by the Department of Energy, and the Respondent was not at liberty to invite additional members of his team. While I feel it is unfortunate that the full Plume Calculation Team was not informed of the existence of the meeting and the process by which the Respondent chose those individuals who would represent the Plume Calculation Team at the meeting, I do not believe that this constitutes scientific misconduct.

To support Allegation 3, the Complainant argues that the Respondent tried "to discredit and remove members from the Plume Calculation Team who reported that PIV was making mistakes and underestimating the leak rate" and provides, as corroborating evidence, an email conversation showing that Dr. Marcia McNutt, the leader of the FRTG, encouraged Mr. Shaffer to resign from the team on June 8, 2010. The Respondent has provided the Panel with further evidence of Dr. McNutt's desire to remove Mr. Shaffer from the team. While there is ample evidence of discord in the Plume Calculation Team, and perhaps much could be learned from the Team's dynamics about leadership during crisis, I note that Dr. McNutt is not the subject of this inquiry.

In summary, absent clear indication that the Respondent actively tried to remove members of the Plume Calculation Team who were reporting that PIV was underestimating the flow rate, the available evidence suggests to me that Allegation 3 has no substance and does not require further investigation.

Allegation 4. *Fabrication of Findings and Failure to Provide Traceability of Data: In violation of NAO 202-735D, §6.01(a), §6.01(b), §6.01(c), §7.01, §7.02, [Respondent] added an additional estimate by PIV from a scientist who was not a member of the Plume Team to his Final Report and Final Presentation. [Respondent] did not reveal to the Plume Team's members, to peer reviewers, to key decision makers, or to the public that he added an estimate from a scientist who was not a member of the Plume Team. It appears that [Respondent] also altered the values of the estimates by PIV to make them appear identical and more "consistent."*

The basis of this allegation is the Table of estimates shown in Appendix 1 of the Plume Calculation Team *Final Report* (see page 15). This appendix was written by statisticians from the National Institute of Standards and Technology (NIST), not by the Respondent, and was based on estimates provided to the statisticians by members of the Plume Calculation Team. The confusion as to the identity of the experts appears to derive from the fact that four members of the Plume Calculation Team – Alberto Aliseda, Oscar Flores, Juan Lasheras and James Riley – co-wrote a single Appendix report. However, as Plume Calculation Team members and flow rate experts, they were entitled to provide independent estimates of the flow rate to the statisticians. An earlier report by the NIST statisticians listed the following experts by name: Alberto, Ira, Jim, Juan, Omer and Steve. (This report, *Pooling Expert Assessments: June 8 2010,* by Antonio Possolo and Pedro Espina is available as an attachment to an email from Matt Lee-Ashley to Marcia McNutt and others, with the subject: "FOR IMMEDIATE REVIEW - close hold - draft release on updated Plume Team estimates" sent on June 8, 2010.) The estimates of Alberto, Jim and Juan were indentical on June 8, 2010, which leads me to surmise that experts A, B and C in the *Final Report* are the same three people.

In summary, the available evidence suggests to me that Allegation 4 has no substance.

6. EVIDENCE CONSIDERED

- Written allegations of Scientific Misconduct
  - Letter of Jan 23, 2012 from Complainant to Office of Deputy Undersecretary for Operations
  - Letter of Feb 2, 2012 from Complainant to Office of Deputy Undersecretary for Operations
- Correspondence from DUSO to Parties
  - Letter of February 27, 2012 from Deputy Undersecretary for Operations, Charles S. Baker to Complainant.
- Emails from DOI FOIA site. www.usgs.gov/foia/FRTG_emails/

Steven T. Wereley. "Re: Sen. Boxer." Email to Juan Lasheras, copied to Bill Lehr and Marcia McNutt. June 9, 2010. Available at: http://www.usgs.gov/foia/FRTG_emails/06-07-2010...Senator%20Boxer.pdf, accessed October 21, 2012.

Bill Lehr. "Re: meeting Monday 6/7/2010 Noon PDT." Email to Mark Sogge and others, June 7, 2010. Available at: http://www.usgs.gov/foia/FRTG_emails/06-07-2010...Meeting%20Monday%206-7-2010%20Noon%20PDT%20R.pdf, accessed October 21, 2012.

Bill Lehr. "Re Sample Conclusion Template." Email to Franklin Shaffer and others, June 7, 2010. Available at: http://www.usgs.gov/foia/FRTG_emails/06-07-2010...sample%20conclusion%20template.pdf, accessed October 21, 2010.

Juan Lasheras. No subject. Email to Bill Lehr and Marcia McNutt, June 8, 2010.Available at: http://www.usgs.gov/foia/FRTG_emails/06-08-2010...FRTG%202.pdf, accessed October 21, 2010.

Vic Hines. "Final draft FRTG Qs and As." Email to Marcia McNutt, June 9, 2010. Available at: http://www.usgs.gov/foia/FRTG_emails/06-09-2010...Final%20draft%20FRTG%20Qs%20and%20As.pdf, accessed October 21, 2010.

Steven T. Wereley. "RE: draft report - two notes from Omer." Email to Alberto Aliseda and Franklin Shaffer, May 26, 2010. Available at: http://www.usgs.gov/foia/FRTG_emails/05-26-2010...draft%20report%20-%20two%20notes%20from%20Omer%202.pdf, accessed October 21, 2010.

- Written testimony of Respondent

Written Statement of June 23, 2012 by Respondent to IRP Chair Dr. Paul Rago entitled "Answer to the false allegations of Frank Shaffer and Public Employees for Environmental Protection (PEER)"

Written Statement of August 14, 2012 by Respondent to IRP Chair Dr. Paul Rago. Untitled

- Official Reports

McNutt, M., R. Camilli, G. Guthrie, P. Hsieh, V. Labson, W. Lehr, D. Maclay, A. Ratzel, M. Sogge. 2011a. Assessment of Flow Rate Estimates for the Deepwater Horizon / Macondo Well Oil Spill, National Incident Command, Interagency Solutions Group, Flow Rate Technical Group http://www.doi.gov/deepwaterhorizon/loader.cfm?csModule=security/getfile&PageID=237763

Plume Calculation Team. 2010a. Estimated leak rates and lost oil from the Deepwater Horizon Spill, May 27, 2010. Interim Report to Flow Rate Technical Group.

Plume Calculation Team. 2010b. Deepwater Horizon Release, Estimate of Rate by PIV. PlumeTeam report to the Flow Rate Technical Group. July 21, 2010.
http://www.usgs.gov/oilspill/
and http://www.doi.gov/deepwaterhorizon/index.cfm

Possolo and Espina. June 8, 2010.Deepwater Horizon—Pooling Expert Assessments.
http://www.usgs.gov/foia/FRTG_emails/06-08-2010...Draft%20release%20on%20updated%20Plume%20Team%20estimates.pdf

- Press releases and statements in the press

Flow Rate Group Provides Preliminary Best Estimate Of Oil Flowing from BP Oil Well. May 27, 2010 http://www.restorethegulf.gov/release/2010/05/27/flow-rate-group-provides-preliminary-best-estimate-oil-flowing-bp-oil-well

Flow Rate Technical Group's Continues Analysis of Data from BP Well. June 8. Press release.
http://www.usgs.gov/foia/FRTG_emails/06-08-2010...Draft%20release%20on%20updated%20Plume%20Team%20estimates.pdf

Statement of the Flow Rate Technical Group, Plume Calculation Team. June 8, 2010.
http://www.usgs.gov/foia/FRTG_emails/06-08-2010...Draft%20release%20on%20updated%20Plume%20Team%20estimates.pdf

Press Briefing by National Incident Commander June 9, 2010
http://www.restorethegulf.gov/release/2010/06/09/press-briefing-national-incident-commander-june-9-2010

Admiral Allen, Dr. McNutt Provide Updates on Progress of Scientific Teams Analyzing Flow Rates from BP's Well. June 10, 2010 Press release http://www.doi.gov/news/pressreleases/Admiral-Allen-Dr-McNutt-Provide-Updates-on-Progress-of-Scientific-Teams-Analyzing-Flow-Rates-from-BPs-Well.cfm

U.S. Scientific Team Draws on New Data, Multiple Scientific Methodologies to Reach Updated Estimate of Oil Flows from BPs Well. June 15, 2010. http://www.restorethegulf.gov/release/2010/06/15/us-scientific-team-draws-new-data-multiple-scientific-methodologies-reach-updated

U.S. Scientific Teams Refine Estimates of Oil Flow from BPs Well Prior to Capping. August 2, 2010. http://www.restorethegulf.gov/release/2010/08/02/us-scientific-teams-refine-estimates-oil-flow-bps-well-prior-capping

- Journal publications

McNutt, M.K, R. Camilli, T. J. Crone, G. D. Guthrie, P. A. Hsieh, T. B. Ryerson, O. Savas, and F. Shaffer. 2011. Review of flow rate estimates of the Deepwater Horizon oil spill. Proc. National Academy of Science. PNAS Early Edition. 8 p. www.pnas.org/cgi/doi/10.1073/pnas.1112139108