Exhibit D: Peer Review and Information Quality Breakdown in an Endangered Species Act Decision: The Case of the Greater Sage Grouse

Rob Roy Ramey II[1,3]

John D. Wehausen[2]

Laura M. Brown[1]

[1]Wildlife Science International, Inc.
P.O. Box 386
Nederland, CO 80466
USA

[2]University of California (retired)
White Mountain Research Station
3000 East Line St.
Bishop, CA 93514
USA

[3]Corresponding author
robroyrameyii@gmail.com

16 October 2013

ABSTRACT
This paper discusses issues with the implementation of Information Quality Act guidelines in U.S. Endangered Species Act (ESA) listing decisions. These issues are illustrated by the key scientific paper and peer review processes that figured prominently in the decision to list the greater sage grouse (*Centrocercus urophasianus*) as "warranted" under the ESA as a threatened or endangered species in 13 states and provinces. We examine limitations of the data, errors and bias in the analyses and inferences based upon those analyses, and then explore why and how questionable data and analyses were used as the basis for such a far-reaching decision, even when independent peer review did not support the conclusions. We discuss policy implications and potential policy solutions, and how these checks and balances could reduce opportunities for various types of error and bias in the ESA decision-making process.

INTRODUCTION

The conservation of biodiversity is a worldwide concern, especially the conservation of threatened and endangered species and the habitats they depend upon. In the United States the protection of species threatened with potential extinction is provided by the Endangered Species Act (ESA) of 1973. The Act requires that the U.S. Fish and Wildlife Service (USFWS) make decisions to list species as threatened or endangered, *"solely on the basis of the best scientific and commercial data available"*. The USFWS must determine whether a species, subspecies, or distinct population is likely to become threatened or in danger of extinction (endangered) in the reasonably foreseeable future, throughout all or a significant portion of its range. In making such decisions (and those that follow to aid the recovery of species), the USFWS is afforded substantial judicial deference in interpreting what constitutes best available scientific and commercial data, sometimes refered to as *best available science* (Hickey 2009). Although the ESA refers to *data*, the USFWS actually relies on published and unpublished studies, and professional opinion, rather than the underlying *data*. The USFWS assures the quality of the information which is used for its decisions by relying on the Information Quality Act (IQA), the bulletin (OMB 1999, 2002) implenting IQA, and the Department of Interior's Scientific Integrity policies (DOI 2011).

For many rare or declining species, there are only limited data available, and those data may be incomplete or inadequate for the purposes of assessing population numbers and trends. The problem is particularly acute in species that are not of commercial value. For example, data may have been collected over many years for other purposes and now applied to answer questions that were not originally anticipated. Or, the agencies monitoring the species may have been reluctant to change and adopt superior methods of data collection. Therefore, listing decisions and recovery actions may be made on the basis of limited or sub-optimal data, which can hinder the types of discriminating analyses and the inferences that can be drawn from them.

In other cases, underlying data used in studies may not be made public because agencies or researchers have witheld access to them. This may be because agencies or researchers consider the data proprietary, or they may not want to reveal the locations of endangered species. In either case, when data are not made public, it prevents independent reanalysis and review (Fischman and Meretsky 2001).

In this paper we explore these issues by examining the highly influential scientific paper by Garton et al. (2011), that figured prominently in the decision to list the greater sage grouse as "warranted but precluded" for threatened or endangered status under the ESA (USFWS 2010). We examine limitations of the data used by Garton et al. (2011), the analyses, inferences based upon those analyses, and then explore why and how such an important decision as an ESA listing could have been based on such questionable analyses of questionable data. This is of particular concern given that there was considerable independent peer review that did not support the conclusions of that analysis. We also discuss potential policy solutions to these shortcomings, and how these checks and balances could benefit the conservation of species by reducing opportunities for various types of error in the research and decision making process.

SHORTCOMINGS OF THE APPLIED METHODOLOGY

The species in question, the greater sage grouse, is a large ground-nesting bird dependent upon sagebrush habitat in western North America. Each spring, sage grouse congregate at traditional sites (leks) where the males display in order to attract and mate with females. Thirteen states and provinces began counting the number of adult male sage grouse at prominent leks in the 1940's and 1950's as a potentially useful index of population size. Initially, male counts were made at a few large and easily located leks. Then, from 1965 to 2001, the number of counted leks increased approximately ten-fold. The data collection, however, continued to be a non-random sample of leks, but included no information on the number of leks that were not included in these counts.
Concern and repeated litigation over the status of sage grouse, and a desire to quantitatively estimate population sizes and trends, has motivated three different research groups to conduct analyses of male lek count data (Connelly et al. 2004; WAFWA 2008; Garton et al. 2011). The most recent and most ambitious of these studies, Garton et al. (2011), used 42 years of male lek count data (from 1965-2007) to estimate population trends, reconstruct estimates of past population sizes, and forecast population sizes and probabilities of persistence 30 and 100 years into the future, to 2037 and 2107 respectively.

The male lek count data used by Garton et al. (2011) and previous authors were collected by different states and provinces - some of which used different methods - and by many different individuals at thousands of locations. Data from different states and provinces were combined for analysis in Sage Grouse Management Zones (SMZs) and metapopulations. The authors claim that they carefully examined all data prior to analysis to ensure that they were obtained following appropriate procedures, but the authors also acknowledged that they *"had to assume that the data were collected properly."* However, the number of cases where this assumption had to be made was not reported, nor did they report the number of leks that were deleted from the raw data.

After filtering the data, the analytical approach had multiple procedures:

(a) Male lek count data were used to develop annual estimates of the rate of change from the previous year to the present year for each lek with successive counts, and these were then averaged across each population.

(b) The reciprocal of those estimates was then used to back-calculate (reconstruct) breeding population sizes prior to 2007 (the terminal year in which the largest number of leks was counted). This effectively estimated how many male sage-grouse would have been counted in earlier years, if the maximum number of leks counted had been counted every year. A formula for estimating the compounding error of such a procedure was applied to their reconstructed population data and 90 percent confidence intervals (CI) were reported.

(c) The reconstructed population sizes were then used to find "best fit" stochastic population models by considering 26 exponential and density-dependent growth models with varying numbers of parameters (including year, two time periods (1969-1987 and 1988-2007), and time lags). Model selection procedures were employed to evaluate models relative to each other. Additionally, the data were grouped in 5-year blocks, using averages and associated statistics for each block.

(d) The models developed in (c) were used for 30 and 100-year population forecasts as part of a population viability analysis (PVA). Extinction predictions were based on the proportion of replicate trajectories where the estimated effective population (N_e) sizes fell below 50 or 500, in which case populations were deemed "quasi-extinct."

Garton et al. (2011) reported that 44% of their models indicated declining carrying capacity through time, ranging from -1.8% to -11.6%. In other words, their results found that 56% of populations were stable, increasing, or had no significant trend. Also, 18% of the models incorporated lower carrying capacities from 1987 – 2007, compared to 1967-1987. Again, this could also be viewed as 72% of populations being stable, increasing, or having no significant trend. They also reported that 13% (3) of 24 populations for which they had sufficient data, had a high likelihood of declining below $N_e = 50$, and 54% (13) had a likelihood of declining below $N_e = 500$ within 30 years. On a 100-year time horizon, 75% of the populations and 29% of the SMZs were projected to decline below effective population sizes of 500. For 2007 they estimated a minimum of 88,816 male grouse. They assumed a ratio of 2.5 adult females per lekking male, yielding a minimum population estimate of 310,856 adult sage grouse. This number contrasts with an estimated population size of approximately 535,542 sage grouse, based on estimates provided by states and provinces (USFWS 2010).

The authors acknowledged the inherent inaccuracy of lek counts and several limitations of the data for inferring population abundance and trends, and conceded that they made no attempt to estimate true population abundance using leks counts. Yet, despite this caveat, Garton et al. (2011) subsequently used lek count data to create an index of historical abundance, population reconstructions, and probability of extinction forecasts for 30 and 100-year time horizons. They concluded by proposing that: *"these forecasts will be useful in guiding decisions concerning the future of sage-grouse and the sagebrush communities upon which they depend."*

Data limitations in the conservation of endangered species can lead to a policy dilemma analogous to the challenge of minimizing Type I and Type II statistical errors. Type I error occurs when conservation actions are based on an erroneous or exaggerated conclusion that a biologically meaningful and statistically valid risk threatens a species. Type II error may occur if conservation actions are *not* taken, based on the mistaken belief that little or no biologically meaningful and statistically valid risk threatens a species, when one actually does. Minimizing both types of error can be difficult, because attempts to minimize one type of error can increase the probability of the other type of error.

In practice, the situation is more complex than this simple dichotomy for two reasons. First, Type I and II error scenarios assume that the basic data are sound, a condition that can be difficult to meet with endangered species. Because scientific uncertainty is anathema to government, scientists are encouraged to fill these information gaps as best they can with new analyses of existing data, or new data and analyses. Second, when one type of error is viewed as having more serious consequences than the other, the standard of proof becomes asymmetrical (MacCoun 1998). For the USFWS, one of the consequences of a decision that might result in a species decline (or extinction) is the threat of costly lawsuits brought by environmental groups. And once listed, the USFWS and other agencies have an additional consequence to consider: in 1978 the U.S. Supreme Court interpreted language of the ESA to conclude that *"the value of endangered species is incalculable"* and that a listed species must be protected *"whatever the cost."* Such interpretations naturally lead to a precautionary approach and to increased potential for Type I error in listing decisions. Other errors, including errors of omission, selective interpretation, or confirmation bias (Nickerson 1998; Robertson 2009), may also contribute to either Type I or II error.

Known issues with lek count data

Numerous published papers have pointed out *why* male lek count data are unreliable and inappropriate for inferring population abundance and trends. These include: Jenni and Hartzler (1978), Emmons and Braun (1984), Walsh et al. (2004), Connelly and Schroeder (2007), Garton (2007), and Western Association of Fish and Wildlife Agencies (2008). There were also six publicly available peer reviews commissioned by the Colorado Division of Wildlife that specifically pointed out methodological issues with Garton et al. (2011). These include Conroy (2009), Noon (2009), Runge (2009), and three anonymous peer reviews (CDOW 2009). (Note: The version of Garton et al. that was reviewed in 2009 by Conroy, Noon, and Runge was the peer reviewed and accepted version that the USFWS relied upon in making its ESA listing decision in 2010 (Garton et al. 2009). The 2011 version of Garton et al. that we discuss here is virtually identical to the 2009 version, with just minor edits to text.)

Briefly, the issues identified by the authors and reviewers listed above include:

1) No demonstrated correspondence between male lek counts and actual population number or trends.
2) Data collection procedures were not standardized among states and sometimes varied within states over time.
3) Personnel monitoring leks and individual differences in methods and detection ability change over time, leading to *observer bias*.
4) Data sets from multiple states and provinces (i.e. data from two to six states) were combined for analysis of SMZs. (This problem is exacerbated by the fact

that some states supplied data summaries while others provided raw lek count data.)

5) Data were not randomly collected by any state or province, and there are an unknown number of unsampled leks in each population. Therefore, it is impossible to know the extent to which sampling effort is representative of the distribution of sage grouse within populations or SMZs. This also affects the definition of dispersal distances which, in turn, are used to determine whether populations are isolated.
6) Only males were counted; there is no accounting for the number of females or juveniles in the populations sampled, their sightability, nor how these differ across different sagebrush habitats or decades.
7) The number of grouse counted at a lek depends upon the spatial definition of a lek: a more inclusive definition includes nearby satellite leks and results in a higher count, while a more restrictive definition results in more leks with fewer birds counted in each lek. Previous authors provided quantitative criteria for what constituted a lek. Connelly et al. (2004) considered all males within 2.5km of a lek to be part of that lek, while the Western Association of Fish and Wildlife Agencies (2008) used 0.5km as a cut-off. Garton et al. (2011) did not specify any cut-off distance.
8) A disregard for estimating the number of unknown leks makes it is impossible to use male lek count data to estimate population number or trends.
9) A lek is not reported in databases until two or more male grouse are found using it. Consequently, counts at a lek start with a positive number and any lek that has become inactive or merged with another lek is followed by zero counts. This leads to negatively-biased trends.
10) The assumption that lek-attendance rates of adult male greater sage-grouse are high and constant is not supported by the data.
11) The number of sage grouse leks being counted has increased over time, but the non-random sampling of leks has not yet changed.
12) Small sample sizes and variation in sample sizes across years at each lek increases the statistical unreliability of reconstructed population estimates.

The low resolution of population reconstructions

Plots of population reconstructions and their 90% confidence intervals in the study by Garton et al. (2011) are so wide that no trend can be supported at that confidence level for many populations. (At 95%, the confidence intervals would be so wide that there would be nothing to discuss about the results.) The following illustrates the magnitude of the problem: First, the 90% CI for the Dakotas (Figure 2 in Garton et al. 2011): about 950 male sage grouse were estimated for 2005, but the 90% CI for 1968 was 400 - 9,250, thus a trend ranging between a 90% decrease and a 150% increase over that time period. Second, the east-central Idaho population, with only two leks counted in 1965-1969 and four leks in 2000-2007, had 90% CIs between zero and no upper limit across all years. Yet despite the enormous uncertainty surrounding these and other population reconstructions, Garton et al. (2011) were willing to make several remarkably *precise* predictions about the future of some populations. For example, they stated that the Powder River Basin, Wyoming population *"will fluctuate around carrying capacity which will decline from 3,042 males attending leks in 2007, to only 312 males attending leks in 2037, to going extinct with only two males attending leks in 2107 if this trend continues at the same rate in the future."* That population had a 90% CI of 0 - 180,000 in 1968, 5,000 - 40,000 in 1987, and an estimate of about 8,000 in 2007.

Lack of accounting for error in population growth models and negative trend bias

It is important to recognize that the population growth models in Garton et al. (2011) were not fitted to observed lek count data but instead to *reconstructed population estimates*. These were calculated in such a way that the input and output variables share data, and therefore cannot be considered independent (i.e. the population reconstruction method depends upon quantities that appear on both the "prediction" and "predictor" side of the equation). One reviewer (Conroy 2009) reported that this resulted in "built in patterns" in the reconstructed population estimates, which in turn affected the population growth models and led to erroneous inferences. Similarly, one of the anonymous CDOW reviewers reported a negative trend bias when Garton et al's (2011) method was applied to simulated input data that deliberately had *no* trend. That reviewer reported that 34-40% of the simulated populations produced a statistically significant negative trend using Garton et al.'s (2011) methods. These reviewers also pointed out that sampling variation and statistical uncertainty from reconstructed population estimates were not carried over by Garton et al. (2011) into subsequent models of population growth and persistence.

These assessments are supported by results in Appendix 1 of Garton et al. (2011) where they list results for best models of their reconstructed population data: the 26 adjusted r^2 values range from 0 to 0.682, the highest of which is for a population with data for only 1996-2007, and the next closest value was 0.498, and average r^2 was only 0.257. This indicates that the models, on average, did not explain 75% of the variation in the data sets (i.e. low resolution).

The low statistical resolution of the reconstructed populations for which the models were developed suggests that a great deal of error accompanies the PVA forward projections. Similar to the issues with estimating population reconstructions in reverse time, errors will compound and grow exponentially. Garton et al. (2011) discuss this potential, but ultimately emphasize the literature that better supports their analyses. In reality, given the poor resolution of the reconstructed population data base and the growth models based upon it, the PVA projections incorporate a great deal of compounded error that renders projections at even 30 years meaningless. This leaves almost no clearly useful analytical results in what Garton et al. (2011) produced.

Mathematical error(s)

Garton et al's (2011) use of 20 males and 2.5 times that number of females to achieve an N_e of 50, is in error and should result in an N_e of 57.14 (using Wright's 1938 equation). Instead, only 17.5 breeding males would be needed for an N_e of 50 (assuming a ratio of 2.5 females per lekking male). Likewise, 175 males rather than 200 would be required for an N_e of 500. In other words, extinction risk was overestimated across all populations by setting the minimum number of breeding males *higher* than necessary for maintaining an N_e of 50 or 500. Although these differences may seem slight, they do establish different thresholds for generating extinction probabilities across all populations. (This was not a result of the formula error noted below.)

Garton et al. (2011) presented an incorrect equation for estimating effective population size: $N_e=1/((1/Nm) + (1/Nf))$, where Nm is number of breeding males and Nf is the number of breeding females in a population. The correct equation, from Wright (1938) is: $N_e=4NmNf/(Nm+Nf)$. The two equations would have been mathematically equivalent if Garton et al. (2011) had used a four instead of a one in the numerator. It is unknown whether this mistake carried over into the population viability analysis (in which case it would have overestimated extinction risk), or whether it was a typographical error in their paper. This question cannot be answered because the code and data used to perform the analysis are not publicly available.

Reliance on the 50/500 rule of thumb: an obsolete concept

The basic concept underlying minimum viable population size (MVP) and population viability analysis is that there must be some "minimum conditions for the long term persistence and adaptation of a species or population" (Soule 1987). An effective population size (N_e) of 50 was suggested as the minimum in the short term to limit the loss of heterozygosity through genetic drift and potential resultant inbreeding depression that could lead to a risk of population extinction (Soule 1980). An N_e of 500 was proposed as the minimum necessary to maintain the long-term adaptive potential of a population (Franklin 1980) based on a handful of studies of quantitative genetic variation in highly inbred lines of mice, maize and *Drosophila* (summarized by Lande 1976). None of those studies actually compared extinction risk with genetic variation or N_e.

Although the 50/500 rule of thumb is widely cited, field data, laboratory studies, and theory show that this rule of thumb is not a reliable predictor of extinction. Successful populations have been founded by few individuals, and populations with a much lower N_e than 50 have persisted long past when they should have gone extinct under the 50/500 rule of thumb (Krausman et al. 1993, 1996; Goodson 1994; Luikhart and Cornuet 1997; Wehausen 1999; Ramey et al. 2000; Frankham 2005). Criticism of the 50/500 rule of thumb was succinctly summarized by Boyce (1997): *"Unfortunately, the 50/500 rule does not have a sound genetic or demographic basis. And there is no theoretical or empirical justification for basing MVP on an estimate of N_e… until such evidence becomes available, reliance on rules of thumb, such as the 50/500 rule is arbitrary and capricious."*

In practical terms, the predictions of future sage grouse population sizes by Garton et al. (2011) are not falsifiable because they are simply probability statements about what might happen if environmental conditions are unchanged. While long-range predictions based on models are potentially useful heuristic tools, they are also notoriously inaccurate and can be easily over-applied (Pielke, Jr. and Conant 2003). Their lack of potential falsifiability effectively places decisions based upon them outside the realm of science.

Hunting mortality: an error of omission in model development

Garton et al. (2011) ignored the effects of sport hunting in their models, although it is the largest documented source of sage grouse mortality: 207,433 sage grouse harvested in the U.S. during 2001-2007 (Reese and Connelly 2011). We find it curious that Garton et al. (2011) ignored hunting mortality, while suggesting that other human activities must have reduced carrying capacity, specifically: expansion of cheatgrass and conifer woodlands, increased fire frequency, energy development, and spread of West Nile Virus.

If one accepts population estimates (88,816 male grouse in 2007 or a total population size of 310,856), then hunters removed 28,180 sage grouse or approximately 9 percent of the species in 2007 alone. In four of the six pervious years, the take was even higher (up to 37,607 in 2006). These numbers do not include the number of grouse that were wounded and not recovered by hunters.

Regionally, the estimated percentage of sage grouse hunted may have been even higher in some years. For example, in 1992 an estimated 34,388 sage grouse were harvested by sport-hunters in Wyoming (Reese and Connelly 2011). Using the upper and lower 90% CI values of the estimated number of males in the Wyoming Basin SMZ and Powder River population in 2007 (and 2.5 adult females per male counted at leks), hunting loss would have amounted to 12 - 29% of the estimated adult population. This is the same SMZ where Garton et al. (2011) estimate a rate of decline between 3.4% and 10.5% annually. With this level of hunting mortality occurring annually, we question the assumption that there is no (additive) demographic effect (Gibson et al. 2011). The difficulty in establishing a link is in part due to the fact that sage grouse lek counts, the basis of hunting harvest, are not a reliable indicator of population number or trends (see discussion above). Clearly, more refined data and methods are needed to address this question.

DISCUSSION

Once a ESA listing is final, compliance is a costly endeavor. Compliance with regulations associated with listings usually involve a substantial allocation of conservation resources in order to be effective (Government Accountability Office 2006; Ferraro et al. 2007). Compliance can lead to secondary costs to local communities and regional economies (Wanger 2010), and is imposed with no regard to cost based on the

Supreme Court's admonishment that ESA listed species must be protected *"whatever the cost*" (*TVA v. HILL, 437 U.S. 153 (1978)).*

Independent and detail-oriented peer reviews are important for prudent decision makers. Equally important is the availability of data and methods used to ensure the replicability of results and allow identification of errors, methodological biases, and potential for falsification of hypothesized population trends (Fischman and Meretsky 2001). This is recognized and required by IQA Guidelines issued by federal agencies. However, in the case of the greater sage grouse, the failure was not of the guidelines themselves, but of the agencies' failure to apply them.

In the case of the sage grouse decision, the question is: what were the checks and balances in the ESA listing process, and why did these fail to detect and filter out a study with numerous limitations, errors, and unfalsifiable predictions? We argue that the reason is largely due to reliance on an ineffective peer review process and acceptance of "scientific" information that has not been sufficiently scrutinized (e.g. due to data being withheld or reliance on population predictions with unreasonable margins of error).

Peer Review

Science is a human activity, therefore errors can and do occur, and peer review exists as a filter on information quality. However, there is no guarantee that papers being peer reviewed will be examined in depth, results replicated, or reviewer comments fully addressed and made public. Unless peer reviewers are provided the original data along with sufficient time and resources to adequately investigate the analyses, the reviewers are forced to assume that the data are sound.

Currently, the USFWS does not require that the data used in research that it cites be made publicly available, nor do they actively engage in or encourage replication of results in peer review. Since 2002 however, IQA guidelines set a higher standard for federal agencies, including the USFWS. They require that studies be reproducible and provide a rebuttable presumption that peer-review of the studies was adequate (OMB 2002). Additionally, the U.S. Department of Interior's information quality guidelines (US-DOI 2002) require that reproducibility *"shall generally require sufficient transparency about data and methods that an independent reanalysis could be undertaken by a qualified member of the public."* And USFWS (2007) guidelines state that, *"higher levels of scrutiny are applied to influential scientific, financial or statistical information, which must adhere to a higher standard of quality."* It is apparent that these requirements were not applied to their full extent by the USFWS in its consideration of Garton et al. (2011) because the raw data were unavailable, and valid criticisms of the data and methods made by reviewers outside of the production of this monograph series were clearly ignored by both the editors of the volume and the USFWS in its decision. This raises questions about the efficacy of the peer review process in the production of this highly influential paper, and with the peer review of the USFWS decision that cited the paper 62 times.

It also raises issues with the efficacy of the peer review of the recent USFWS and State-sponsored Conservation Objectives Team Report (COT 2013), which cited Garton et al. (2011) 61 times and based their population threats analyses, population definitions, current and projected numbers of males in each population, and probability of population persistence on Garton et al. (2011).

As long-time students of the ESA and peer reviewers of USFWS recovery plans and proposed rules, it has been our experience that peer reviewer and public comments on proposed rules are typically combined into broad categories, paraphrased, and summarized by the USFWS. Responses are then prepared to these summaries. Many valid criticisms and details are potentially lost in this process, diminishing the value of reviews and public comments. For example, valid issues raised in outside peer reviewer comments of Garton et al. (2011) were only discussed in a brief paragraph in the USFWS's "warranted but precluded" decision (USFWS 2010):

> *"We received these reviews and have reviewed them in the context of all other data we received in preparation of this finding. Their primary concern was about the applicability of analyzing and presenting future population projections in the manner done by Garton et al. (in press), based on the limitations of the data, the assumptions required, and uncertainty in the estimates of the model parameters. Garton et al. (in press) acknowledged these concerns, as several of the reviewers pointed out, and their analyses underwent peer review via the normal scientific process prior to acceptance for publication."*

The last sentence of this summary also illustrates a key false assumption in the ESA decision-making process: that the *"normal scientific peer review process"* leading to publication is automatically a good filter on information quality. Empirical evidence and the collective experiences of many authors renders this assumption disputable (Mahoney 1977; Roy and Ashburn 2001; Hilborn 2006; McCook 2006; Sandström and Hällsten 2008; Casadevall and Fang 2009; Fang et al. 2012; and Ramey 2012). While traditional peer review is a useful tool, it is clearly an imperfect tool and applied with great variation. As a result, proposals have come forth on how to improve its effectiveness or adopt innovative alternatives (Weicher 2008; Suls and Martin 2009).

Despite variation in how peer-review is conducted, there are at least two well-justified standards that distinguish a rigorous peer-review process from a less than rigorous one. One is: required preparation of a detailed response to each of the peer review criticisms, and discussion of why the criticisms might not be considered valid and should be ignored. While the extent to which this occurred in production of the *Studies in Avian Biology* monograph (of which Garton et al. 2011 is one of 25 chapters) is unknown because reviews were confidential (itself a violation of the Information Quality Guidelines), the USFWS's response to outside peer reviewer's criticisms (see previous paragraph) is illustrative of a process that deviates from this standard.

The second standard is: the role of editorship and authorship need to be independent so that editors are not in a position to review and approve articles that they have authored. In the case of Garton et al. (2011), one of the authors, J. Connelly, was also one of the two editors of the monograph that Garton et al. (2011) was published in. (Both editors were authors on multiple papers in this monograph.)

There is the need for greater accountability and a more comprehensive review process for highly influential scientific papers used in ESA listing decisions (and of the listing decisions themselves). However, it is questionable whether an additional round of peer review or the convening of expert panels would be adequate. An extensive social psychology literature points to the reasons *why*: even with intentions of neutrality, traditional peer review and expert panels may be unable to uncover the whole truth because of inherent cognitive and motivational mechanisms that contribute unintentionally to bias (e.g. strategy-based errors, confirmation bias, or majority amplification; see MacCoun 1998 for an extensive review).

Better access to data

In an ideal world, all of the data used to develop a highly influential scientific paper would be publicly available to allow for independent replication and ensure the potential for falsifiability. Therefore, it is worth asking: why is this not the case with Garton et al. (2011) and many similar, highly influential papers, especially given that *"The* [sage grouse] *monograph is recognized by the USFWS and the Court as the primary source of science for the new review and listing determination."* (USGS 2009b).

Until such time that underlying data of highly influential studies used in ESA decisions are mandated to be publicly available, few options exist to gain access to these data. While the option to obtain data under FOIA from federal agencies is available, and has been used for replication and publication of analyses (e.g. Turner et al. 2004, 2006), federal agencies must possess the data if they are to be obtained under FOIA. However, the little known OMB Circular A-110 provides a second option for public access to data under FOIA when studies are federally grant-funded (OMB 1999):

> *"(d) (1) In addition, in response to a Freedom of Information Act (FOIA) request for research data relating to published research findings produced under an award that were used by the Federal Government in developing an agency action that has the force and effect of law, the Federal awarding agency shall request, and the recipient shall provide, within a reasonable time, the research data so that they can be made available to the public through the procedures established under the FOIA."*

Procedures are well established, as some agencies (such as the National Institute of Health) are familiar with the responsibilities of granting agencies and awardees. To our knowledge, no data requests under A-110 have yet been submitted to the USFWS.

A third potential remedy exists in the form of "requests for correction" under the IQA. This administrative procedure only allows for suggested corrections to the record and does not provide legal remedy should an agency fail to correct or provide information.

The remedy of last resort, costly and time consuming for all involved, but comprehensive in its potential depth, is the power of subpoena.

From our viewpoint, these remedies should not be necessary. It is in the best interests of biodiversity conservation, responsible agencies, and researchers, to provide ready access to data used in scientific papers and key decisions, either online or in publicly accessible archives.

CONCLUSIONS

It is our view that a scientifically critical review of the study by Garton et al. (2011) on greater sage grouse would have concluded that there was no scientific basis for a "warranted" decision (for a ESA threatened listing) because of fundamental problems with the available data as well as with the analyses. Instead, the decision should have called for development of better data collection, with the goal of revisiting the issue in 5 years, when the relationship of lek counts to actual population data might be better understood, or a probability-based census method implemented. This would have minimized Type I error without increasing Type II error.

We acknowledge that multiple studies have presented documentation of the loss of sagebrush in the western U.S. and Canada (i.e. Miller and Rose 1999; Schroeder et al. 2004), however, the extent to which this loss of habitat translates into loss of sage grouse, is not certain. Therefore, the policy-relevant questions about sage grouse should be: 1) are populations in decline; 2) if so, where; 3) why has it occurred; and 4) what can be done to insure the stability of these populations? In order to address these questions, reliable data on population numbers and trends are needed. Those data are currently lacking.

To their credit, Garton et al. (2011) called for establishment of range-wide, standardized methodologies based on probability sampling of leks, breeding males, and females, that would allow for more meaningful population analyses in the future (e.g. sentinel-lek and dual-frame sampling methods). Walsh et al. (2010) have recently proposed the application of mark–resight methods to estimate population size in sage grouse and other lekking species.

From our assessment, the data collected for more than 50 years by thirteen states and provinces are inadequate to answer the above questions regardless of the analysis applied. Repeated calls to reform this weak and outdated methodology, whose limitations have been clearly documented here and elsewhere, have not yet moved agencies into reforming their "business as usual" approach to counting male sage grouse on leks each spring. This puts the overall management of this species on a shaky database and will continue to hinder effective management until more biologically relevant and statistically defensible census methods are adopted.

The issues and potential solutions identified here also apply to the ESA listing of species outside of the U.S. (an increasing trend) and more broadly to endangered species laws of other nations (e.g. Australia's Environment Protection and Biodiversity Conservation Act of 1999, Canada's Species at Risk Act of 2002, and South Africa's National Environmental Management: Biodiversity Act of 2004), as well as international treaties (e.g. the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) of 1973). Failure to implement changes will result in falure to adequately protect species that are truly at risk of extinction.

References

Boyce, M. S. Population viability analysis: adaptive management for threatened and endangered species. Chapter 11. In: M.S. Boyce and A. Haney (eds.), Ecosystem Management: Applications for Sustainable Forest and Wildlife Resources. Yale University Press; 1997.

Casadevall, A., and Fang, G. C. Is peer review censorship? Infection and Immunity. 77: 1273–1274; 2009.

Connelly, J. W., Knick, S. T. Schroeder, M. A. and Stiver S. J.. Conservation Assessment of Greater Sage-grouse and Sagebrush Habitats. Unpublished report. Western Association of Fish and Wildlife Agencies, Cheyenne, Wyoming; 2004.

Connelly, J. W., and Schroeder, M. A. Historical and current approaches to monitoring greater sage-grouse. Pages 3-9. In: K. P. Reese and R. T. Bowyer (eds.), *Monitoring populations of sage-grouse*. Idaho State University College of Natural Resources Experiment Station Bulletin 88; 2007.

Conroy, M. J. Review of Chapter 16 "Greater sage-grouse population dynamics and probability of persistence" by Garton et al., Prepared at the request of the Colorado Division of Wildlife, Fort Collins. Available from the Research Library at Colorado Parks and Wildlife, Fort Collins, Colorado; 2009.

CDOW (Colorado Division of Wildlife). Colorado Division of Wildlife reviews of greater sage-grouse studies in avian biology chapters: Greater sage-grouse population trends and likelihood of persistence. June 30, 2009. Colorado Division of Wildlife, Fort Collins. Available from the Research Library at Colorado Parks and Wildlife, Fort Collins, Colorado; 2009.

COT (Conservation Objectives Team). Greater Sage-grouse (*Centrocercus urophasianus*) Conservation Objectives Team (March) 2013. Washington, D.C., U.S. Fish and Wildlife Service; 2013.

Emmons, S. R., Braun, C. E. Lek attendance of male sage-grouse. Journal of Wildlife Management. 48: 1023-1028; 1984.

Fang, F. C., Steen, R. G., Casadevall, A. Misconduct accounts for the majority of retracted scientific publications. Proceedings of the National Academy of Sciences of the United States of America. 109(42):17028-33. doi: 10.1073/pnas.1212247109. Epub 2012 Oct 1. 2012.

Ferraro, P. J., McIntoshb, C., Ospina, M. The effectiveness of the US endangered species act: An econometric analysis using matching methods. Journal of Environmental Economics and Management. 54: 245-261; 2007.

Fischman, R. L., Meretsky, V. J. Endangered Species Information: Access and Control. Washburn Law Journal. 41(1): 90-113; 2001.

Frankham, R. Stress and adaptation in conservation genetics. Journal of Evolutionary Biology. 18: 750-755; 2005.

Franklin, I. R. Evolutionary changes in small populations. pp. 135-149. in M. E. Soule and B. A. Wilcox (eds.) Conservation Biology: An Evolutionary-Ecological Perspective. Sinauer Associates, Sunderland, Massachusetts; 1980.

Garton, E. O., Musil, D. D. Reese, K. P. Connelly, J. W. Anderson, C. L. Sentinel lek-routes: an integrated sampling approach to estimate Greater Sage-Grouse population characteristics. Pages 31-41 In: K.P. Reese and R.T. Bowyer (eds.). *Monitoring populations of sage-grouse.* College of Natural Resources Experiment Station Bulletin 88. University of Idaho, Moscow, Idaho; 2007.

Garton, E. O., Connelly, J. W. Hagen, C. A. Horne, J. S., Moser, A., Schroeder, M. A. Greater Sage-Grouse population dynamics and probability of persistence. In: Knick, S. T. and J. W. Connelly (eds.) Ecology and Conservation of Greater Sage-Grouse: A Landscape Species and Its Habitats. Studies in Avian Biology. Originally available at: http://sagemap.wr.usgs.gov/monograph.aspx. (Archived copy available at: http://web.archive.org/web/20100527124712/http://sagemap.wr.usgs.gov/monograph.aspx Accessed October 1, 2011); 2009.

Garton, E. O., Connelly, J. W., Horne, J. S. Hagen, C. A., Moser, A., Schroeder, M. A. Greater Sage-Grouse population dynamics and probability of persistence. Pp. 293–381 in S. T. Knick and J. W. Connelly (editors). Greater Sage-Grouse: ecology and conservation of a landscape species and its habitats. Studies in Avian Biology (vol. 38), University of California Press, Berkeley, CA.; 2011.

GAO (Government Accountability Office). Endangered Species: Time and Costs Required to Recover Species Are Largely Unknown. US-GAO. Report to Congressional Requesters. GAO-06-463R; 2006.

Gibson, R. M., Bleich, V. C., McCarthy, C. W. Russi, T. L. Recreational hunting can lower population size in greater sage-grouse. Pp. 307-315 in B.K. Sandercock, K. Martin, and G. Segelbacher (eds.). Ecology, Conservation, and Management of Grouse. *Studies in Avian Biology* (vol. 39), University of California Press, Berkeley, CA.; 2011.

Goodson, N. J. Persistence and population size in mountain sheep: why different interpretations? Conservation Biology. 8: 617-618; 1994.

Hickey, V. Reauthorizing the Endangered Species Act in favor of wildlife and wild lands: an inevitable result of narrative changes in twenty-first century America? Journal of Environmental Law and Litigation. 24: 423-454; 2009.

Hilborn, R. Faith-based Fisheries. Fisheries. 31(11): 554-555; 2006.

Jenni, D. A., Hartzler, J. E. Attendance at a sage-grouse lek: implications for spring census. Journal of Wildlife Management. 42: 46-52; 1978.

Krausman, P. R., Etchberger, R. C., Lee, R. M. Persistence of mountain sheep. Conservation Biology. 7: 219-220; 1993.

Krausman, P. R., Etchberger, R. C., Lee, R. M. Persistence of mountain sheep populations in Arizona. Southwestern Naturalist. 41: 399-402; 1996.

Lande, R. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genetics Research. 26: 221-235; 1976.

Luikart, G., Cornuet, J. M. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology. 12: 228-237; 1997.

MacCoun, R. J. Biases in the interpretation and use of research results. Annual Review of Psychology. 49: 259-287; 1998.

Mahoney, M. J. Publication prejudices: An experimental study of confirmatory bias in the peer review system. Cognitive Therapy and Research. 1(2): 161-175; 1977.

McCook, A. Is Peer Review Broken? The Scientist. 20(2): 26-31; 2006.

Miller, R. F., Rose, J. A. Fire history and western juniper encroachment in sagebrush steppe. Journal of Range Management. 52: 550-559; 1999.

Nickerson, R. S. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. Review of General Psychology. 2(2): 175-220; 1998.

Noon, B. R. Review of Chapter 16 “Greater sage-grouse population dynamics and probability of persistence” by Garton et al. Prepared at the request of the Colorado Division of Wildlife, Fort Collins. Available from the Research Library at Colorado Parks and Wildlife, Fort Collins, Colorado; 2009.

NTT (BLM Sage Grouse National Technical Team). Report on National Greater Sage-Grouse Conservation Measures. Unpublished report. Sage Grouse National Technical Team, December 21, 2011.

OMB (Office of Management and Budget). Circular A-110 Revised 11/19/93 As Further amended 9/30/99. Available at http://www.whitehouse.gov/omb/circulars_a110#1. (Accessed 1 September 2010); 1999.

OMB (Office of Management and Budget). Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by federal agencies. Final Guidelines; 2002.

Pielke, Jr., R. A., Conant R. Y. Best practices in prediction for decision-making: lessons from the atmospheric and earth sciences. Ecology. 84(6): 1351-1358; 2003.

Ramey, R. R., Luikart, G., Singer, F. J. Genetic bottlenecks resulting from restoration efforts: the case of bighorn sheep in Badlands National Park. Journal of Restoration Ecology. 8(4S): 85-90; 2000.

Ramey, R.R. On the origin of *specious* species. In: Institutions and Incentives in Regulatory Science. J. Johnston (ed.), Lexington Books; 2012.

Reese, K. P., Connelly J. W.. Harvest management for Greater Sage-Grouse: a changing paradigm for game bird management. Greater Sage-Grouse population dynamics and probability of persistence. Pp. 101–111 in S. T. Knick and J. W. Connelly (editors). Greater Sage-Grouse: ecology and conservation of a landscape species and its habitats. Studies in Avian Biology (vol. 38), University of California Press, Berkeley, CA.; 2011.

Robertson, C. B. Judgment, identity, and independence. Connecticut Law Review. 42(1): 1-48; 2009.

Roy, R., Ashburn, J. R. The perils of peer review. Nature. 414: 393-394; 2001.

Runge, M.C. Review of Chapter 16 “Greater sage-grouse population dynamics and probability of persistence” by Garton et al., 18 December 2009. Prepared at the request of the Colorado Division of Wildlife, Fort Collins. Available from the Research Library at Colorado Parks and Wildlife, Fort Collins, Colorado. 2009.

Sandström, U., Hällsten, M. Persistent nepotism in peer-review. Scientometrics. 74(2):175-189. 2008.

Schroeder, M. A., Aldridge, C. L., Apa, A. D., Bohne, J. R., Braun, C. E., Bunnell, S. D., Connelly, J. W., Deibert, P. A., Gardner, S. C., Hilliard, M. A., Kobriger, G. D., McCarthy, C. W., Mitchell, D. L., Rickerson, E. V., Stiver., S. J. Distribution of sage-grouse in North America. Condor. 106: 363–373; 2004.

Sedinger, J. S., White, G. C., Espinosa, S., Partee, E. T., Braun, C. E. Assessing compensatory versus additive harvest mortality: an example using greater sage-grouse. Journal of Wildlife Management. 74(2): 326–332; 2010.

Soule, M. E. Thresholds for survival:maintaining fitness and evolutionary potential. Pages 151-169 In: Soule, M. E. and B. A. Wilcox, (eds.) Conservation Biology. Sinauer; 1980.

Soule, M. E. Introduction. In:Soule, M.E. (ed.) Viable Populations for Conservation. Cambridge University Press, Cambridge, pp. 1-10; 1987.

Suls, J., Martin, R. The Air We Breathe: A Critical Look at Practices and Alternatives in the Peer-Review Process. Perspectives on Psychological Science. 4: 40-50; 2009.

Turner, J. C., Douglas, C. L., Hallum, C. R., Krausman, P. R., Ramey, R. R. Determination of critical habitat for the endangered Nelson's bighorn sheep in southern California. Wildlife Society Bulletin. 32(2): 427-448; 2004.

Turner, J. C., Douglas, C. L., Hallum, C. R., Krausman, P. R., Ramey, R. R. Osterman's assumption of a flawed habitat model is premised on facts not in evidence: Turner et al. (2005) response to Osterman et al. (2005). Wildlife Society Bulletin. 33(4): 1465-1473; 2006.

US-DOI (U.S. Department of Interior). U.S. Department Of The Interior, Information Quality Guidelines Pursuant To Section 515 Of The Treasury And General Government Appropriations Act For Fiscal Year 2001. (Available at http://www.doi.gov/ocio/iq.html. Accessed 1 September 2010); 2002.

US-DOI (U.S. Department of Interior). Integrity of Scientific and Scholarly Activities. Department Manual. Part 305, DM, Chapter 3:1-38. Office of the Deputy Secretary, Department of Interior. (Available at http://www.doi.gov/scientificintegrity/index.cfm Accessed October 2, 2012.); 2011.

USFWS (U.S. Fish and Wildlife Service). U.S. Fish and Wildlife Service Information Quality Guidelines. (Available at http://www.fws.gov/informationquality/ Accessed 1 September 2010.); 2007.

USFWS (U.S. Fish and Wildlife Service). Endangered and Threatened Wildlife and Plants; 12-Month Findings for Petitions to List the Greater Sage- Grouse (*Centrocercus urophasianus*) as Threatened or Endangered. The Federal Register. 75(55): 13909-14014; 2010.

USGS (U.S. Geological Survey). Sage-Grouse Conservation: A Major Challenge in the West. News Release dated November 4, 2009. 1p.; 2009a.

USGS (U.S. Geological Survey). Scientific information for greater sage grouse and sagebrush habitats. Unpublished briefing Paper, Dated September 29, 2009. 6p.; 2009b.

WAFWA (Western Association of Fish and Wildlife Agencies). Greater sage-grouse population trends: an analysis of lek count databases 1965-2007. Unpublished report by the Sage- and Columbian Sharp-tailed Grouse Technical Committee, Western Association of Fish and Wildlife Agencies , Cheyenne, WY.; 2008.

Walsh, D. P., White, G. C., Remington, T. E., Bowden, D. C. Evaluation of the lek-count index for greater sage-grouse. Wildlife Society Bulletin. 32(1):56-68; 2004.

Walsh, D. P., Stiver, J. R., White, G. C., Remmington, T. E. Apa. Population estimation techniques for lekking species. Journal of Wildlife Management. 74(7): 1607–1613; 2010.

Wehausen, J. D. Rapid Extinction of Mountain Sheep Populations Revisited. Conservation Biology. 13(2): 378-384; 1999.

Weicher, M. Peer review and secrecy in the Information Age. Proceedings of the American Society for Information and Technology. 45(1): 1-12; 2008.

Wright, S. Size of population and breeding structure in relation to evolution. Science. 87: 430–431; 1938.

Wanger, O. W. Findings of fact and conclusions of law RE Plaintiffs' request for preliminary injunction against implementation of RPA Component 2 (a/k/a Action 3)(Doc. 433), The Consolidated Delta Smelt Cases, U.S. District Court for the Eastern District of California, May 27, 2010; 2010.