Peer review by A. Treves "of Removing the Gray Wolf (*Canis lupus*) From the List of Endangered and Threatened Wildlife" Federal Register **84**(51): 9648-9687 (2018). Final version 30 May 2019

I appreciate the opportunity given to me to serve as a peer reviewer of the science in the proposed rule and draft biological report. I want to recognize and applaud the Department of Interior and USFWS for enhancing peer review policies and respecting the public interest in its role as trustee of endangered and threatened species of the USA. I also am grateful for the opportunity to present new or overlooked information. This cover page summarizes my responses to the scope of work as a peer reviewer. Throughout, text sent to me by the U.S. Fish & Wildlife Service (USFWS or Service) or its contractor is italicized and my contributions are not italicized.

Response to questions posed to me

Questions posed for me about the Draft Biological Report and Proposed Rule

1*. Does the draft report provide an adequate and concise overview of gray wolf (Canis lupus) taxonomy, biology, and ecology as well as the changes in the biological status (range, distribution, abundance) of the gray wolf in the contiguous 48 United States over the last several decades?*

No, I find the draft biological report adequate on the taxonomy but not on the biology, ecology, or biological status.

2. Please identify any oversights or omissions of data or information, and their relevance to the report. Are there other sources of information or studies that were not included that are relevant to the biological report? What are they and how are they relevant?

In the review that follows, I detail the missing scientific information on all of the following topics: biology, ecology, or biological status. In summary, I am concerned by missing information on human-caused mortality, human attitudes leading people to kill wolves, and dispersal.

3. Does the proposed rule provide an adequate review and analysis of the factors relating to the persistence of the gray wolf population currently listed under the ESA in the contiguous 48-States (human-caused mortality, habitat and prey availability, disease and predation, and effects of climate change)?

No, the proposed rule does not address human-caused mortality or habitat suitability adequately.

4. Have we (the Service) adequately considered the impacts of range reduction (i.e., lost historical range) on the long-term viability of the gray wolf in its remaining range in the lower-48 states (outside of the northern Rocky Mountains) and, if not, what information is missing and how is it relevant?

Almost. I would have liked more information on dispersal and its related legal definition of discreteness and the causes and consequences of recolonization apparently having stopped.

5. Is it reasonable for the Service to conclude that the approach of Michigan, Wisconsin, and Minnesota to wolf management, as described in their Plans and the proposed rule and in the *context of wolf management in the Western Great Lakes area, are likely to maintain a viable wolf population in the Western Great Lakes area into the future?*

No. In the review that follows, I detail the many sources of evidence that are missing to draw such conclusions and the many contrary findings that seem to have been overlooked, which undermine the conclusions.

6. Please identify any oversights or omissions of data or information, and their relevance to the assessment. Are there other sources of information or studies that were not included that are relevant to the proposed rule and, if so, what are they and how are they relevant?

In the review that follows, I detail the missing scientific information on biology, ecology, or biological status. In summary, I am concerned by missing information on all of the following topics: human-caused mortality, human attitudes leading people to kill wolves, the effects of legal killing on illegal killing, and the uncertainties surrounding WGL wolf population census and monitoring.

7. Are there demonstrable errors of fact or interpretation? Have the authors of the proposed delisting rule provided reasonable and scientifically sound interpretations and syntheses from the scientific information presented in the draft biological report and the proposed rule? Are there instances in the proposed rule where a different but equally reasonable and sound interpretation might be reached that differs from that provided by the Service? If any instances are found where this is the case, please provide the specifics regarding those particular concerns.

There are demonstrable errors in the proposed rule and the draft biological report. Several of the Service's documents' interpretations and syntheses are neither reasonable nor scientifically sound. In several instances, a different and equally reasonable (or more) and sound (or more) interpretation has been reached in the scientific peer-reviewed literature. In several cases, results in the best journals (ranked independently on a worldwide scale of impact factors) were ignored or overlooked, in favor of non-peer-reviewed interpretations or results from lower-ranked journals. In a few cases, the stronger evidence was paid for by the USFWS or was co-authored by USFWS staff.

Summary of the scientific evidence in the order presented in my peer review

Following the purpose on p9649 of the proposed rule, I have evaluated "*data, assumptions, and analyses"* and the inferences and conclusions that followed, as scientific statements. For all of the data, assumptions, and analyses to be scientifically sound, I weighed whether I found them accurate, precise, and valid, both on their face, and in light of the literature with which I am familiar. Regarding assumptions specifically, I sought clear statements of assumptions and considered their validity once identified: were they reasonable and were assumptions more likely to be true than false. I flagged unstated assumptions because transparency about assumptions is a key principal of scientific integrity. For data and analyses, I attempted to weigh whether they met the highest standards of evidence, including whether they were generated, interpreted, and reported with high standards of integrity and scientific rigor— to assist the

USFWS in considering the "best available" science under the ESA. Overall, I focused on 4 related scientific topics in the proposed rule and biological report. I also focused on the following claim,

"The metapopulation in the Great Lakes area contains sufficient resiliency, redundancy, and representation to sustain populations within the gray wolf entity over time. Therefore, we conclude that the relatively few wolves that occur outside the Great Lakes area within the gray wolf entity, including those in the west coast States and lone dispersers in other States, are not necessary for the recovered status of the gray wolf entity." p9683.

1. The first topic was the definition of the entity as it relates to the scientific criteria for recovering or delisting under the ESA (Endangered Species Act of 1973 as amended). I see this as a scientific issue because one must agree on how to group or split subpopulations of wolves that exist on the ground before scientists can evaluate a claim about their status. I found the rationale for the listed entity coherent. By contrast, I found that the handling of evidence for dispersal, and treatment of vacant habitat and different subpopulations of the gray wolf entity were not consistent in a scientific sense. As a result, the conclusions drawn about current range, vacant habitats, and northeastern USA gray wolves were not well substantiated. My scientific judgment is that the gray wolf entity 's current range is not defined well by scientific standards.

2. The second topic was the definition of suitable habitat. I found the definition of suitable habitat did not conform to standard practice in ecology and conservation, and moreover it contained an unstated value judgment in place of scientific observation.

3. The third topic was gray wolf mortality. I found that the presentation of data, review of literature, analysis, and predictions about gray wolf mortality were sometimes incomplete, sometimes imprecise, sometimes inaccurate, or sometimes invalid. Taken altogether, the interpretation of the threat posed by human-caused mortality and its cumulative effects experienced by gray wolves was inaccurate and misleading.

4. The fourth topic was human attitudes to gray wolves. I see this as a scientific issue because the measurement and interpretation of human attitudes can help to predict future human action to conserve or eliminate gray wolves. I found the review of literature and analysis of attitudes was incomplete, inaccurate, and overall misleading.

In sum, I do not find the proposed rule and draft biological report present the best available science and I made numerous suggestions for improving the identification, presentation, and the analysis of evidence in the draft biological report and in the proposed rule.

Addressing process of the peer review as an issue of scientific integrity

I also address scientific process throughout my peer review, because the best available scientific and commercial data depends on scientific integrity, consistent standards of evidence, and strong inference. The USFWS purpose for per review guided me in this direction because it made clear a base of fact precedes the decision,

"The purpose of peer review is to ensure that our decisions are based on scientifically sound data, assumptions, and analyses." p9649

I found an overall problem with the process for coming to the legal determination in the proposed rule. This is a scientific problem. The problem is that the draft biological report was released simultaneous with the proposed rule for the peer review. Logically, the draft biological report which should stick to scientific evidence and scientific inference should be peer reviewed long before a political and legal decision about delisting is made. The nature of a base of evidence is that it is solid before one builds on it.

In other words, the evidence should inform the value judgments that underpin a political and legal decision. Without the sequence I recommend, the proposed rule looks like a predetermined conclusion. Moreover, the proposed rule is full of presentations of evidence (often flawed) rather than the proposed rule referring to inferences that were drawn (and passed peer review) in the draft biological report before being used in the proposed rule.

Science cannot tell us what we should do. I am identifying a scientific problem and recommending an improvement to the scientific integrity of the peer review. The current mix in the proposed rule confuses evidence with policies based on legal and ethical reasoning. In my opinion, a clearer separation between fact and value judgments would make the process more scientific. Indeed, separating the ethical review – along the lines of one conducted by the USFWS for spotted owls recently (Lynn 2018) – from the scientific review would do much to dispel the current muddle between fact and policy in the proposed rule. Facts and evidence established by scientific consensus should be regarded as distinct from the ethical reasoning about what we ought to do.

1. **Definition of the entity and whether that entity met the ESA range criterion for delisting**

2. **Definition of suitable habitat**

3. **Human-caused mortality**

4. **Human attitudes to gray wolves**

5. **Biological Report**

   Appendix

1. **Definition of the entity and whether that entity met the ESA range criterion for delisting**

I understand the legal definition of the gray wolf entity in the proposed rule.

"In this proposed rule, we consider the status of the gray wolf within the geographic boundaries of the two currently listed C. lupus entities... These two currently listed entities are: (1) C. lupus in Minnesota, and (2) C. lupus in the lower 48 United States and Mexico outside of Minnesota, the NRM DPS (Montana, Idaho, Wyoming, eastern third of Washington and Oregon, and north-central Utah), and the area covered by the experimental population area for C. l. baileyi…" p9653

I understand how the evidence of dispersal leads to a biological definition of discreteness, and I think I understand how that, in turn, applies to a legal definition for listable entities, starting with the treatment of the Pacific Northwest (PNW) wolves in the proposed rule:

*We determined that these wolves are not discrete, under our DPS policy, from wolves in the NRM DPS… and, therefore, are **not a valid listable entity under the Act**… Therefore, wolves in western Washington, western Oregon, and northern California are not a valid DPS because they are not discrete from the NRM DPS… " p9653 emphasis added.*

The above quotation clearly connects all wolves of Washington, Oregon, and California with the NRM wolves on biological grounds (dispersal in both directions). I believe the biological evidence is consistent with this claim. Also, the proposed rule goes on to link Wisconsin and Michigan wolves to Minnesota wolves through frequent dispersal (quite correctly, in my interpretation of several peer-reviewed papers). By this logic, the WI and MI wolves could not be a DPS either. So the PNW and NRM wolves proposed as one subpopulation, and we have the Minnesota, Wisconsin, and Michigan wolves as another subpopulation (WGL) of the gray wolf entity. So far so good. However, the criterion of discreteness was abandoned when it came to examining evidence of dispersal within the gray wolf entity.

If discreteness only applies to a DPS, then it would seem that the NRM could potentially be an invalid DPS. That might be a legal issue but scientifically, the handling of dispersal and discreteness was inconsistent in the proposed rule.

The facts of dispersal are not analyzed in the same manner to consider discreteness of the NRM DPS and Minnesota's wolves.

"*…a number of lone long- distance dispersing wolves have been documented outside core populations of the Great Lakes area and western United States since the early 2000s. Confirmed records of individual wolves have been reported from North Dakota, South Dakota, Utah, Colorado, Nevada, Missouri, Indiana, Illinois, Nebraska, and Kansas. The total number of confirmed records in each of these States, since the early 2000s, ranges from one in Nevada to* *at least 27 in North Dakota, with the latter also having an additional 45 probable but unverified reports " p9656*

Is Minnesota's entity non-discrete from the NRM DPS or the less-well-understood northeastern USA gray wolves? Are PNW wolves discrete from all the rest? Clearer terminology will strengthen the proposed rule and potentially reduce future disagreements.

Also, clarity about subpopulations within the gray wolf entity was hindered by a proliferation of terms in the proposed rule. Within p9655-9657, the proposed rule uses the terms metapopulation, extension of a metapopulation, subpopulation, "*core population*", and "*biologically part of (although outside the legal boundary of) an already recovered and delisted population*". The draft biological report does not present enough information on dispersal to clarify these terms.

I did not find an answer to my questions above or even a reasoned analysis about dispersers in the proposed rule or the biological report. Consider the scant information about northeastern USA wolves. The proposed rule pooled the northeastern canids with the gray wolf entity to delist, but did so without providing evidence such as data on dispersal, analysis of status, or an analysis of discreteness in either the proposed rule or the draft biological report.

Moreover, the proposed rule seemed to ignore vacant habitat while claiming to consider the status of gray wolves where they occur. It does not do so for the gray wolves in the Northeast. as the following quotations illustrate:

"Our analysis of threat factors below does not consider the potential for effects to C. lupus in areas where the species has been extirpated—rather, effects are considered in the context of the present population." p9659

Apparently, the proposed rule considers the status of wolves ONLY within the geographic boundaries of the two currently listed *C. lupus* entities.

"In this proposed rule, we consider the status of the gray wolf within the geographic boundaries of the two currently listed C. lupus entities to determine whether these wolves should remain on the List in their current status, be reclassified, or be removed from the List." p9653

The approach above seems to contradict the ruling by the 2017 decision of the D.C. Circuit Court of Appeals, which was cited in the proposed rule. While my comment might seem focused on policy, it actually reflects the need for additional evidence – on vacancy, occupied range, and dispersal – before delisting the gray wolf entity.

"[The U.S. Court of Appeals in 2017]… upheld the District Court's vacatur, concluding that the Service failed to reasonably analyze or consider two significant aspects of the rule: The impacts of partial delisting and historical range loss on the remainder of the listed entity." p9650

The D.C. Court of Appeals raises the critical issue of insufficient analysis. Although no specific percentage of range was mandated by the ESA, the claim that one can delist regardless of what portion of the listed entity's range is occupied seems an extreme position, a case of following the letter of the rule while ignoring its goal. Perhaps the USFWS is not obligated to recover a listed entity across all or a significant portion of historical range, but the range of the listed entity and a significant portion of that range both seem relevant for analysis. A facile assertion that "there are no wolves so there is no information" would be flawed. Scientists glean information from allied species when faced with an information gap about a particular place. For example, how are other large carnivores faring in habitat currently lacking gray wolves? More analysis of the vacant and the occupied range of the gray wolf entity is imperative. To improve clarity, the proposed rule should insert the plain terminology of the ESA, e.g. "the significant portion of range" clause.

Here the D.C. Court of Appeals raises the critical issue of thorough analysis. Although no specific percentage of range was mandated by the ESA, the claim that one can delist regardless of what portion of the listed entity's range is occupied seems an extreme position to take; a case of following the letter of the rule, but ignoring its goal. Perhaps the USFWS is not obligated to recover a listed entity across all or a significant portion of historical range, but the range of the listed entity and a significant portion of that range both seem relevant for analysis. The easy retort that 'there are no wolves so there is no information' would be flawed. For one, scientists glean information from allied species when faced with an information gap about a particular place. So for example, how are other large carnivores faring in the habitat currently vacant of gray wolves? Therefore I wanted more analysis of the vacant and the occupied range of the gray wolf entity. To improve clarity, the proposed rule should insert the plain terminology of the ESA, e.g., the significant portion of range' clause.

There is also a scientific question associated with what is a significant portion of range, given the definition of the listed entity's vast range (Figure 2). I would answer the question scientifically, as follows: Congress did not mean significant in its statistical sense (whatever that might mean) but in its common English usage. I also deduce they meant the geographic range of the listed entity as opposed to some other type of range. As a scientist, I doubt the definition of significant in common English would include less than half because we have a perfectly understandable phrase for that ("a minority of the range"). By the same logic, I doubt plain English understanding would mean 51% because they could have written "a majority of the range". I also doubt that plain English interpretation would include 100% because 'all' would be covered by the ESA phrase. So visual inspection of Figure 2 and the proposed rule both confirm that the gray wolf entity is currently not occupying a majority of the listed entity's range. Therefore, my scientific judgment is that the gray wolf entity has not recolonized enough of its range to meet the standard of a significant portion of range.

The scientific basis of the gray wolf entity and its range seems questionable on scientific grounds because I found neither consistent terminology for subpopulations of current wolves, nor consistent handling of data on dispersal, discreteness, range, or status across the entity.

2. Definition of suitable habitat

I agree with the proposed rule assertion, *"gray wolves are habitat generalists*" p.9654. But then to further claims,

"We consider suitable habitat as forested terrain containing adequate wild ungulate populations (elk, white-tailed deer, and mule deer) to support a wolf population. Suitable habitat has minimal roads and human development, as human access to areas inhabited by wolves can result in wolf mortality…" p9662 (citing Oakleaf, Mladenoff, Carroll)

First, habitat suitability is estimated at the scale of individual animals or breeding social units, not populations as in the latter quotation. In brief, for wolves, suitable habitat can be as small as the area needed for one breeding pair to raise pups that survive to adulthood. Second, it is standard practice in ecology to define suitability by observing where reproduction and survival occur to define suitability, not by imposing a human value judgment on it as hinted above and obvious in the following quotation.

"*Thus, the estimated 450 wolves in [Minnesota's] Zone B could be subject to substantial reduction in numbers. …* wolves *should be restored to the rest of Minnesota but not to Zone B (Federal Zone 5) because that area 'is not suitable for wolves' (USFWS 1992, p. 20)." p9669.*

The latter paragraph undermines the proposed rule's definition of suitable habitat because an area containing 450 wolves is suitable by its own definition.

The USFWS in 1992 confused a value judgment (e.g., we don't want wolves in Zone B) with a scientific evaluation of suitable habitat (e.g., wolves survive and reproduce here). The proposed rule endorses the 1992 recovery plan repeatedly, despite multiple unscientific statements about habitat suitability.

Third, defining a human behavior (wolf-killing) as a habitat feature is contrary to long-standing ecological practice. Not all humans kill gray wolves or even want to kill gray wolves (e.g., (Treves et al. 2013). Therefore, human density is a weak correlate of threat to wolves. Stronger correlates of inclination to kill wolves have been identified and they do not always occur where human population density is moderate or high (Smith et al. 2010). Therefore, any claim that the cause of suitability of habitat is the presence or density of humans would be erroneous.

I anticipate the rebuttal that habitat suitability indices often include mortality sinks or habitat patches that hinder reproduction or survival for wildlife. But ecologists do not define habitat as unsuitable because a predator **resides there**. Nor should the proposed rule define a habitat as unsuitable **because people live there**. Only when mortality or failed reproduction are recurrent phenomena in a restricted area might that area be classified as unsuitable. Therefore, the definition of suitable habitat does not seem to accord with standard practice in ecology or conservation science.

Clearly lacking from the proposed rule and draft biological report analysis of suitable habitat is a reasoned spatial model of where illegal killing and legal killing have been concentrated or are more likely in the future. All I can glean from the proposed rule about these topics is that after delisting, legal killing will increase as states and tribes make value judgments about where wolves will be allowed to recolonize. Such value judgments are not scientific evaluations of suitable habitat as explained above., Moreover the role of illegal killing deserves further analysis in light of the ESA admonition to treat overutilization (e.g., high human-caused mortality) as a threat to listed species. Defining human-caused mortality as a threat to be reduced seems to be the appropriate stance, rather than redefining spaces, in which human-caused mortality might occur, as unsuitable. The ESA permits predator control to protect listed species. The USFWS could treat illegal killing as predation and control it, rather than redefine it as an immutable factor in habitats.

Likewise, an unwillingness to curb human-caused mortality (by the agencies responsible) is a value judgment, not a scientific fact or prediction. Unwillingness to curb illegal killing does not make wolves less capable of using habitat. Similar points were made by (Bruskotter et al. 2013), in a document I am confident was shared with the USFWS. The latter seems to have been omitted from the draft biological report and proposed rule.

3. **Human-caused mortality**

    a. **Is human-caused mortality lower now?**

    b. **Illegal killing**

    c. **Wolf mortality in light of government measurements and regulatory mechanisms**

    d. **Legal killing to protect domestic animals**

    e. **Cumulative effects**

Starting on p. 9659, the proposed rule presented its summary of human-caused mortality. I do not find it to be a thorough and comprehensive review of the best available scientific and commercial data. Furthermore, even when the evidence summarized seems to be the best available, I find several key analyses and conclusions drawn from the review are unclear, illogical, or poorly reasoned. There is a serious gap in the draft biological report as pertains to wolf mortality also. Before recommending improvements to the proposed rule and draft biological report, I present findings from 2019 and I summarize a debate about wolf mortality dating from 2015–2018 in Box 1 below.

Box 1. Wisconsin and Michigan wolves appear to have been adversely affected by delisting and other policies that liberalized wolf-killing above and beyond the number of wolves legally killed.

Santiago-Ávila (in review) presented a table of the fates (time to event) for radio-collared wolves in Wisconsin 1980–2012 (Table A). In summary, half of all radio-collared wolves with a recorded fate had disappeared from 1980–2012 (n=243 of 485). Also, he found that periodic increases in the rates of disappearance were significantly associated with periods of permitted lethal control or delisting. Furthermore, the preliminary results about incidence (i.e., the proportion of individuals experiencing an event over time) of disappearances over time (or relative incidence) increased by 11–34% relative to periods with full protection.

The range of values depended on 3 scenarios for missing data and the resulting uncertainty left by gaps in state monitoring data. The state's 2012 population report and mortality data omit 26 radio-collared wolves entirely (5.1% of the sample). That was a majority of the 41 radio-collared wolves monitored from January 1–April 14, 2012. Months of requests for those data were rebuffed. Nevertheless, the observations, simulation and imputation scenarios converge in their results and together lead to a clear inference. The risk that a wolf was lost to monitoring rose 'considerably during periods with legalized wolf-killing. Table A presents the raw data.

It seems unlikely that radio-collars fail more often during such policy periods, or that emigration (but not immigration) changes when federal policy changes.

Table A. The distribution of disappearance events for 243 radio-collared wolves in Wisconsin 1980–2012 where 0 = periods without legal wolf-killing and 1= during briefer periods with legal wolf-killing. Results below correspond to a scenario of 12 out of 26 wolves imputed as Disappeared or lost to follow-up (45%), which is consistent with the overall proportion of wolves that disappeared (243/250 = 50%).

lib_kill	Events observed	Events expected
0	180	194.02
1	63	48.98
Total	243	243.00
	chi2(1) =	6.80
	Pr>chi2 =	0.0091

Note: I declined to present the full text of methods for the information above because I must protect the intellectual property of a junior colleague, however he informed me that he is willing to present the entire set of methods orally to the USFWS if needed. Therefore, this source of evidence is no different than the personal communications cited in the draft biological report and proposed rule, which do not contain full methods or manuscripts.

The above results derive from radio-collared wolves and corroborate initial suggestions from all wolves of Wisconsin (Treves et al.2017b) and population dynamics of wolves from Michigan and Wisconsin, which I summarize next.

In 2016, (Chapron & Treves 2016a, b) estimated the wolf population growth in Michigan and Wisconsin slowed by 5% with a year-long period of legalized wolf-killing. They demonstrated quantitatively that those periods of lethal management had slowed population growth in proportion to the length of time those periods lasted and regardless of the number of wolves killed. They could not find evidence nor a biological mechanism to explain why such a slow-down in growth would occur through reproduction or dispersal, so they inferred poaching had increased during periods of legalized killing. Such an explanation is consistent with Santiago-Ávila's results on radio-collared wolves being lost to follow up (off the air) at higher rates during periods of liberalized wolf-killing above.

The attempt by Pepin et al., 2017 and Stien, 2017 to rebut Chapron and Treves (2016a) failed, principally because neither provided biological evidence for the now dubious view that density-dependent growth had occurred (Chapron & Treves 2017a, b), and see Appendix here.

Also Olson et al. (2017) tried unsuccessfully to rebut the results of Chapron and Treves (2016a). Olson et al.'s prior analysis of poaching (Olson et al. 2015) had major shortcomings (Chapron & Treves 2017b) and Appendix here), their table of negative density-dependent population growth in wolves included studies that did not find such growth (Chapron & Treves 2017b), and Olson et al. again omitted essential information on Wisconsin wolf monitoring (Appendix here and (Chapron & Treves 2017b).

Although not directly addressing the preceding debate, work by Stenglein et al. (Stenglein 2014; Stenglein et al. 2015a; Stenglein et al. 2015b; Stenglein et al. 2015c; Stenglein & Van Deelen 2016; Stenglein et al. 2018) cannot support the rebuttals by Pepin et al. Stien, or Olson et al. because those papers do not accurately account for missing radio-collared wolves or biases in wolf census methods that might affect the detection of dead wolves (Appendix and Box 1). Nor does it contradict Santiago-Ávila (above). In sum, I find no credible rebuttal to Chapron & Treves (2016a,b) and it is now supported by Santiago-Ávila above, using different data and methods.

End of Box 1

a. <u>Is human-caused mortality lower now?</u>

The subtext or implicit assumption throughout the proposed rule are that since bounties on gray wolves no longer exist, all human-caused mortality now is regulated and therefore wolves will not be extirpated. This conclusion mistakenly assumes that bounties are the only cause of extirpation, and that other killings of gray wolves by humans do not match, even cumulatively, prior population losses. Extirpation depends on the relative mortality rate and reproductive rate, regardless of how mortality occurs. Several of my criticisms below*,* relate to the incomplete or missing quantification of mortality among gray wolves.

The assertion – ***"an active eradication program is the sole reason that wolves were extirpated from their historical range in the United States"*** (p9659 citing Weaver 1978, p.1.) – is both outdated and fails to define a key element, the "eradication program". Because the eradication program was not defined – just as the characteristics of a bounty are not specified in the proposed rule – scientists cannot evaluate the implicit assumption that today's wolf-killing programs plus illegal killing are different from past wolf-killing programs.

If the bounties of 50 years ago resemble the lethal control and public hunting, trapping, and hounding programs of today in rate and risk for gray wolves, then the burden of proof is to show these will be stopped rapidly and effectively before relisting targets are reached.

Furthermore, recall that the eradication of gray wolves in the WGL was due to a combination of factors (prey depletion, habitat loss, and human-caused mortality), not solely bounties (Thiel 1993). Therefore, the notion that gray wolf populations will persist despite the many increasing sources of human-caused mortality after delisting simply because states and tribes do not implement bounty systems seems unwarranted. At very least, we need a more careful analysis than the following effort.

The Stenlund 1955 comparison to Fuller 1989 suggests a possible decrease in mortality rate after bounties ended. The limitations of this one study are minimally acknowledged in the proposed rule. I don't agree with their conclusion that, *"Nonetheless, these figures provide clear support for the contention that human- caused mortality decreased significantly once the wolf became protected under the Act.",* because the studies even if comparable represent a sample size of 1 subpopulation before-and-after multiple changes in policy (bounty, several Acts of Congress, the ESA, and state policies between 1978 and Fuller's study years later). If such weak evidence were accepted in other topics (e.g., human attitudes, road density for suitable habitat), very different conclusions about delisting would have to be reached. A scientifically defensible inference from the Stenlund and Fuller studies if they are comparable would be, "gray wolf mortality declined between the 1950s and the 1980s, an interval in which several policies changed including the repeal of the bounty', habitat regenerated, and prey recolonized.

Likewise, the proposed rule misstated what we know about periods with and without legal wolf-killing (Box 1), when it asserted,

"Regulation of human-caused mortality has significantly reduced the number of wolf mortalities caused by humans, and although illegal and accidental killing of wolves is likely to continue with or without the protections of the Act, at current levels those mortalities have had little impact on wolf populations." p9661

Box 1 above contradicts the latter. Moreover, the proposed rule places the emphasis inappropriately on *"Regulation of human-caused mortality*' rather than on evidence that the WGL state regulations have reduced wolf mortality.

For example,

"*…the regulation of human-caused wolf mortality is the primary reason wolf numbers have significantly increased and their range has expanded since the mid-to-late 1970s.*" p9659.

I encourage the authors of the proposed rule to be more precise when making scientific assertions. Bounties represented up-regulation of human-caused mortality, and the ESA represented down-regulation of human-caused mortality, i.e., regulation up or down should be specified. This might seem trivial but for two implications of the latter quotation.

i. The quotation seems to imply that removal of bounties and replacement by regulated killing solved gray wolf extirpation. A naïve reader might interpret the quotation above to encourage more regulated wolf-killing. Imprecise language leading to inaccurate conservation prescriptions has been discussed recently in similar contexts (Treves et al. 2018).

ii. Regulation has to manifest in behavior to be effective. Regulation by itself is paper protection. That insight applies to the various assurances that currently state regulations are adequate to avoid a repeat of the history of extirpation.

iii. Was gray wolf range expansion facilitated more by recolonization of prey and regeneration of forests or by lifting of bounties? Alternate views include that societal values, habitat and prey restoration, or ESA listing of the gray wolf per se (without the prohibition on take being enforced) were sufficient to permit wolf recolonization (Mladenoff et al. 1997; Naughton-Treves et al. 2003; Schanning 2009; Smith et al. 2010; Treves & Bruskotter 2014). Even if we accept that past bounties caused more mortality than present private action, the proposed rule and biological report should show that current human-caused mortality is lower than past human-caused mortality as a hazard rate (% of wolves dying from that cause in a given time interval). It is not sufficient to point to population increase and range expansion since 1978 to argue *ipso facto* that the major cause of mortality has been removed or lessened. For greater scientific validity, we need efforts to understand mortality and wolf population estimates historically (Appendix and Box 1). That begins by rigorous measurement and interpretation of today's mortality rates and risks.

The proposed rule must present evidence why delisting will not repeat the past and jeopardize the WGL subpopulation. That explanation should include that much has changed in addition to bounties since 1950-1965, including increases in vehicle traffic, access to firearms, and wildlife-killing practices and technologies that did not exist in the 1960s. To my eye, the proposed rule does not adequately consider the modern risks and deregulations of the means to kill wolves.

b. Illegal killing

"Many wolf killings, however, are intentional, illegal, and never reported to authorities." p9660.

The proposed rule rightfully acknowledges this important point because the major cause of U.S. wolf mortality today is illegal killing (Treves et al. 2017a). However it was followed by many paragraphs that I find misleading, starting with a misstatement of the problem in measuring illegal killing.

"The number of illegal killings is… impossible to accurately determine because they generally occur with few witnesses." p9960.

The number of concealed wolf mortalities is very difficult to measure precisely, but it is not difficult to estimate it accurately with defined bounds of uncertainty, as the following two examples demonstrate. Several recent analyses have done so for radio-collared gray wolves in the USA (Schmidt et al. 2015; Treves et al. 2017a). Recall that the USFWS repeatedly miscalculated human-caused mortality for radio-collared wolves with a systematic bias (Treves et al. 2017a). Also, for unregulated mortality in Alaska, scientists estimated >80% of mortality went unreported to the state agency administering a program of legal wolf-killing (Adams et al. 2008). That troubling conclusion cast doubt on the confidence the proposed rule places in well-regulated, legal wolf-killing discussed next.

The best available evidence indicates that the majority of gray wolves died from illegal killing in the NRM and in Wisconsin **during the periods of ESA protection**, not to mention red wolves and Mexican wolves (Treves et al. 2017a; Treves et al. 2017c). Omission of those facts in the proposed rule and draft biological report is noteworthy for other reasons also. Treves et al. presented the raw data and based their conclusions on recalculation of mortalities published by the USFWS and allied scientists in state and federal agencies. In one case (NRM gray wolves), the agencies' miscalculation led to misidentifying the major cause of gray wolf mortality (Treves et al. 2017a). Because illegal killing of gray wolves was the major cause of their mortality despite ESA protection, the proposed rule should analyze the evidence that delisting gray wolves will not lead to dramatic increases in mortality when illegal killing is counted along with legal killing. The proposed rule seems to have ignored a valuable source of inference about the consequences of delisting in not reviewing the scientific debate over the effects of legaizing wolf-killing (also referred to as "tolerance hunting" (Epstein 2017; Epstein & Chapron 2018); Box 1). This is a key gap in both the proposed rule and the draft biological report.

***"**Liberg et al. (2011, pp. 3–5) suggest more than two- thirds of total poaching may go undetected, and that illegal killing may pose a threat to wolves; however, poaching has not prevented population resurgence in either the Great Lakes area or the northern Rocky Mountains, as evidenced by population growth in those areas... During the times that lethal control of depredating wolves was conducted in Wisconsin and Michigan, there was no evidence of resulting adverse impacts to the maintenance of a viable wolf population in those States."* p9660.

The above quotation misses the mark in several ways. First, it is not consistent with evidence in Box 1, and it is not apparent what the proposed rule means by "resurgence." Second, an estimate of unreported, concealed poaching was provided by Treves et al. 2017b (50%), but overlooked in the draft biological report and proposed rule. Third, where are data from 2012–2014, when wolves were delisted in the WGL? I am aware of only one peer-reviewed analysis of how the populations of the WGL were affected by public hunting and trapping seasons in 2013, and that analysis did not answer important questions about missing methods for wolf census or attendant biases associated with incomplete reporting of dead wolves (Appendix and Box 1).

c. Wolf mortality in light of government measurements and regulatory mechanisms

The adequacy of regulatory mechanisms to control human-caused mortality depends heavily on the adequacy of measuring mortality. If one cannot measure wolf mortality without systematic error, the regulation and enforcement of wolf-killing become systematically biased in the scientific meaning of that phrase. What follows is a summation of the inaccuracies, imprecision and misinterpretation of gray wolf mortality data contained in the peer-reviewed literature.

The discussion of wolf mortality in the proposed rule and the draft biological report largely misrepresents the evidence for super-additive mortality from Creel & Rotella (2010)) and (Vucetich 2012); the latter was published in a USFWS document as part of the public record in the Wyoming wolf peer review. I am aware that Creel & Rotella (2010) was disputed by (Gude et al. 2012) but Vucetich then resolved the discrepancies between the two and still concluded that high human-caused wolf mortality was likely to result in a decelerating wolf population growth and an accelerating decline in wolf population abundances (Vucetich 2012). See the section on the draft biological report below for why the proposed rule presents a misleading picture of sustainable rates of human-caused mortality and also (Creel et al. 2015) for NRM wolves.

In addition to a misleading summary of sustainability research, the proposed rule and draft biological report completely miss key evidence. Regarding the adequacy of mortality data, the proposed rule contains an unstated assumption that we have accurate and precise measurements of mortality from the states' monitoring programs. To the contrary, unchallenged scientific evidence indicates that rates of mortality (% of wolves dying in a given time period) have been under-estimated, and risk of mortality (the proportion of dead wolves dying of a given cause) has been miscalculated, particularly with regard to legal killing and illegal killing as explained below

(Treves et al. 2017a; Treves et al. 2017c) have demonstrated arithmetically that the state and federal wolf monitoring programs produced systematically biased estimates of the risk of mortality because they miscalculated the effect of disappearances in such a way that under-estimated illegal killing by large margins (also see Box 1). That analysis dealt with risk of mortality (the proportion of all dead wolves that died of a given cause), but analyses of mortality rates also raise concerns about the proposed rule.

In two studies, the rate of mortality was under-estimated by current methods of monitoring and analysis (Schmidt et al. 2015; Treves et al. 2017c). Furthermore, a majority of radio-collared wolves in at least one state (Wisconsin) went missing, fate unknown (Treves et al. 2017c) and

Box1. Contrary results by Olson et al. (2015) do not accurately account for missing radio-collared wolves or biases in wolf census methods that might affect the detection of dead wolves (Appendix and Box 1). The proposed rule and draft biological report fail to mention that inferences based on observed mortality of radio-collared wolves are imprecise, and the inaccuracies result in under-estimating illegal killing (Liberg et al. 2012; Treves et al. 2017a; Treves et al. 2017c).

I also found several problems with the proposed rule's handling of legal wolf-killing under state regulatory programs.

"We anticipate the level of mortality due to depredation control that would take place would be similar to what was observed during those times… [with permits or brief delisting 2003–2012]." p9660.

There is an error of extrapolation in this assertion. During recent short periods with permits or delisting (2003–2012), states received federal permits that either stipulated the number of wolves that could legally be killed, or else the period of delisting was brief, a very different scenario from the proposed rule whereby a state would not have to seek federal permission and the time frame for legal killing could be lengthy. For example, after the WGL delisting in January 2012, the state used lethal control to kill an unprecedented number of wolves, invoking catch-up' killing to compensate for not having had such authority the previous year (Treves et al. 2017b), supporting information) with public hunting and trapping allowed as well. Therefore, the expectation cited above seems fanciful and unlikely to come true. The reality is that gray wolf mortality is likely to be higher under the proposed rule than in the brief periods of delisting and issuance of permits for lethal control in the past.

Even if I set aside the scientific questions about whether non-target animals are killed and whether such killing protects human interests at all (Treves & Naughton-Treves 2005; Treves et al. 2016), I found two implicit assumption in the proposed rule's' treatment of wolf-killing to protect livestock.

The first assumption is that killing wolves is effective in protecting human interests. If it is not effective, the USFWS should discourage it, regardless if it has no effect on the viability of wolf populations. The USFWS should discourage it because it is expensive, wasteful, and adds incrementally to wolf mortality. A proposed rule for delisting wolves can and should discourage ineffective management interventions that add to cumulative threats. Indeed, a worldwide review of evidence published by 21 authors from 10 nations (van Eeden et al. 2018) observed the scant evidence and the generally poor quality of evidence used in USA and other government predator-killing programs. Using the most rigorous criteria for evaluating the effectiveness of lethal control, Treves et al. (2016) found more evidence of counter-productive results that would raise the risk for livestock. Research to date provides evidence both for and against the assumption that killing problem wolves protects domestic animals. A longitudinal analysis of wolf pack areas in the NRM suggested that killing entire wolf packs would reduce future livestock losses (Bradley et al. 2015). (A set of three weaker, correlational analyses –

starting with Wielgus et al. - have debated similar data from the NRM, but they merely involved correlations without spatial information.) Although Treves et al. (2016) accepted Bradley et al. (2015) as reliable, that no longer seems appropriate in light of the more recent study by Santiago-Ávila et al. (2018), which points out that Bradley et al. introduced a bias in favor of lethal control by following wolf pack territories rather than wolf packs, because that led them to count intervals in which an area contained no wolves as intervals of effectiveness of lethal control. That approach inflates the apparent effectiveness of whole-pack removal, because the authors counted the period of vacancy as if risk to livestock continued. The sometimes-lengthy period of vacancy would require recolonization and breeding by colonizers before risk resurfaced, which introduces a systematic bias favoring removal of entire wolf packs over other methods. Also Santiago-Ávila et al. (2018) pointed out methodological defects in Bradley et al. (2015) including incomplete data and methods; the latter authors could not provide data and a request for such material was not met. These are critical departures from the scientific requirement of replicability. Relying on Bradley et al. (2015) would be inappropriate unless the methodological problems are fixed, and the data published transparently.

The second implicit assumption was that wolf-killing is self-limited versus the possibility that it accelerates each time wolves are killed. By self-limiting, I mean that wolf-killing stops because the 'problem' wolves are removed and hence the rate of such killing tends to be low and diminish over time. The evidence for and against the assumption is not presented.

Santiago-Ávila et al. (2018) corrected the shortcomings in Bradley et al. (2015) and deployed better methods to the study of lethal control of Michigan wolves. They found corroborating evidence that lethal control did not prevent future livestock losses, and reported that any small reduction in risk for affected farms was outweighed by a subsequent greater risk for farms in neighboring townships. They interpreted patterns of wolf and livestock death to mean that lethal control might cause spill-over effects leading to more losses of domestic animals in subsequent periods. In a state like Michigan that did not monitor the effectiveness of killing wolves legally, and any state that contracts the same federal agency as Michigan did, might erroneously augment wolf-killing after spill-over effects spread to other farms.

Currently, the best available evidence suggests to me that lethal control is risky for domestic animals on farms and is not self-limiting, hence it leads to yet more wolf-killing. Incidentally, social scientific data on human inclinations to kill wolves are consistent with these biological results (See below).

The above concerns relate to the past, observed patterns of legal wolf-killing. Now consider Minnesota's policies and statements about wolf-killing, which are particularly important given Minnesota's essential role in preserving the WGL subpopulation for the entire gray wolf entity.

Without leaps of imagination on my part, I read Minnesota's management plans and the USFWS recovery plan of 1992 as predicting a high likelihood of unprecedented wolf-killing in Minnesota (MN). Consider for example, *"Thus, the estimated 450 wolves in [Minnesota's] Zone B could be subject to substantial reduction in numbers. … At the extreme, wolves could be eliminated from Zone B…(USFWS 1992, p. 20)… there is no need to maintain significant protection for wolves in Zone B in order to maintain a Minnesota wolf population that continues to satisfy the Federal recovery criteria after Federal delisting."* p9669.

Moreover, the proposed rule seems to anticipate changes in regulatory mechanisms, which seem to be in the direction of more wolf-killing, in Zone B, when they state,

"Significant changes in wolf depredation control under State management will primarily be restricted to Zone B, which is outside of the area necessary for wolf recovery (USFWS 1992, pp. 20, 28)." p9669.

Note the value judgment above about what is necessary for wolf recovery. Also, wide latitude has been given to private individuals to kill wolves in MN, as p9668-9669 plainly show. Assurances seem hollow without evidence, as in the following,

"We conclude that this action is not likely to result in the killing of many additional wolves, as opportunities to shoot wolves ''in the act'' would likely be few and difficult to successfully accomplish, a conclusion shared by a highly experienced wolf-depredation agent (Paul in litt. 2006, p. 5)." p9669

It would be more scientific to say that this is a prediction. Predictions entail assumptions and the assumptions here are notable. The proposed rule assumes that (a) owners do not bring their domestic animals to sites in which wolves congregate (e.g., dens or rendezvous sites or carcasses) to expose their animals to imminent risk as was apparently done in Washington State recently, and (b) owners do not bring with them means to kill multiple wolves in a short period (e.g., poison bait, automatic weapons). The possibility of a change in human behavior can reasonably be considered without great imagination. Imagine a Zone B wolf-killing contest which is not far-fetched because such contests have become common across the USA and routine in neighboring Wisconsin in the case of coyotes. I recommend the USFWS consider whether the MN DNR rules resemble a bounty and if not, what are the likely cumulative effects of each element of the MN rules, including whether they might represent a continuous drain on neighboring states or Zone A.

d. Cumulative effects

The proposed rule concluded that cumulative effects of delisting will not imperil WGL wolves. The MN wolves seem vital to that claim. On p9657, the proposed rule discusses Minnesota wolves at length and proposes a population goal of 1251-1440 wolves for long-term viability and genetic diversity. I suggest the draft biological report should fortify its evidence for this assertion by relating the effective population size to published recommendations on mammalian population sizes known to be viable and genetically diverse. For example, the proposed rule might use the best available science to present the probability that a population of 313-350 breeding females (assuming a pack size averaging four, which is typical in the WGL) will go extinct in the face of a demographic or environmental event causing massive mortality or cessation of reproduction for 1-2 years. Presenting reasoned evidence from population viability analyses would fortify overall scientific evidence for the MN population goal described on p9657. Without such reasoned evidence, the assurances in the proposed rule sound hollow.

Without such information, it is difficult to evaluate scientifically the USFWS claim that,

"this region contains sufficient wolf numbers and distribution to ensure the long-term survival of the gray wolf entity. " p9658.

To consider cumulative effects fully, the draft biological report should have presented state-by-state estimates of rates of mortality (observed and unobserved for radio-collared wolves, assuming mortality for those that went missing rather than simply omitting them). Regrettably, the proposed rule considers (see its Table 2) mortality only from lethal control in protecting domestic animal losses and from public hunting and trapping , without considering other causes (illegal, vehicular, nonhuman) and without proper attention to unexplained disappearances of radio-collared wolves (Box 1 for example).

Despite such defects, I pooled lethal control with the harvest data and found errors in the proposed rule. Combining the following scattered items – *"The number of wolves killed for depredation control while wolves were under State management for the second time (20121–2014) was slightly higher (203 wolves in 2011, 262 in 2012, 114 in 2013, and 197 in 2014) than during 2007 and 2008, but was still consistent with those killed under section 4(d) in the surrounding years (192 wolves in 2010 and 213 in 2015)."* p9669 with the USFWS estimates of the MN wolf population https://www.fws.gov/midwest/wolf/aboutwolves/mi_wi_nos.htm, and Table 2 – for four of the years covered in the above quotation, I offer Table B as a recalculation to help address the cumulative effects of delisting.

Table B. Legal wolf-killing in Minnesota before and during periods with state authority.

Winter	Population estimate	# killed in depredation control	# harvested	% killed legally
2015-2016	2,278	213 (9.3%)	0	9.3%
2014-2015	2,221	197 (8.8%)	272	21.1%

2013-2014	2,423	114 (4.7%)	238	14.5%
2012-2013	2,211	262 (11.8%)	413	30.5%

Footnote: I find the annual average for 'depredation control' to be 8.7% (4.7–11.8%), which is slightly higher than that presented in the proposed rule's Table 2.

Because the proposed rule itself predicted harvests will be repeated, the cumulative depredation control and harvest (legal kill) percentages should range from 14.5-30.5% (average 22.0%), without factoring in catch-up killing or higher quotas. The years 2012-2014 are most pertinent to that issue, because gray wolves were delisted and subject to hunting in some or all of those years in some or all of the states of the WGL. The range of 14.5-30.5% is not all the human-caused mortality yet it is already near or over the well-measured thresholds of 17% and 29% mortality rates before wolf populations are very likely to decline (Adams et al. 2008; Vucetich 2012).

The range of 14.5-30.5% does not account for vehicular, nonhuman, or the largest source of wolf mortality (poaching). Treves et al. (2017a) calculated that poaching exceeds legal wolf-killing when wolves are federally protected (and often 2 to 3 times higher). Recall also that delisting did not reduce poaching of radio-collared wolves (contra Olson et al. 2015 examining only observed wolf poaching), legal killing appeared to increase poaching (Box 1). illegal killing is virtually certain to add >9% to population declines, with lethal control, (Chapron & Treves 2016a,b) not even counting the possible rise in poaching with legal public hunting, trapping, or hounding. The weight of evidence suggests that poaching will increase and it is already the major source of mortality and is not well measured by state or federal agencies. Therefore, by past trends without any precaution against new ways of killing wolves at higher rates (see above), MN wolf delisting is almost certain to lead to wolf population declines. The size and speed of those declines depends on categories of mortality that the proposed rule and biological report do not carefully quantify or discuss at length. Therefore, a catastrophic decline in the MN wolf population post-delisting is entirely foreseeable.

In light of all WGL state policies legalizing wolf-killing after delisting and in light of under-estimating mortality in Table 2 of the proposed rule and in Box 1 that I present, I have to conclude that the proposed rule and draft biological report provide unwarranted assurances about the safety of wolves in the WGL after delisting.

The proposed rule initially does well to point out that the five-factor analysis mandated by the ESA and regulatory instructions that follow from it, require that all factors be examined individually and in conjunction for their cumulative effect. However by p. 9658 that lesson seems to have been forgotten:

"*However, the mere identification of factors that could affect a species negatively may not be sufficient to compel a finding that the species warrants listing. The information must include evidence sufficient to* ***suggest that the potential threat is likely to materialize and that it has the capacity (i.e., it should be of sufficient magnitude and extent)*** *to affect the species' status such that it meets the definition of an endangered species or threatened species under the Act.*" p9659 emphasis added.

This argument in the proposed rule is unscientific. While each separate threat might not pose an existential threat to a population, an accumulation of threats might well do so. Similarly, different populations within a listed entity might face different threats that pose a cumulative existential risk to each of those populations, such that looking at the entire listed entity may fail to identify a single existential threat even when the listed entity should be viewed as threatened or endangered by the cumulative effects from different threats. Because the proposed rule and draft biological report fail to adequately review the data for best available evidence, I cannot agree with the inference that delisting will not lead to excessive cumulative effects and to wolf population decline and possible collapse too quickly to be averted by relisting.

4. Human attitudes

"It took a considerable length of time for public attitudes and regulations to result in a social climate that promoted and allowed for wolf recovery within the gray wolf entity. The length of time over which this shift occurred, and the ensuing stability in those attitudes, gives us confidence that this social climate will persist." p9659 and

"*It is also possible that illegal killing of wolves in Minnesota will decrease, because the expanded options for legal control of problem wolves may lead to an increase in public tolerance for wolves (Paul in litt. 2006, p. 5).***"** p9669.

These assertions are not evidence-based. The citation to Paul is outdated and does not reflect any measurements of tolerance for predators or understanding of human behavior. The proposed rule and draft biological report neglected to cite research most relevant to the question of how attitudes to wolves change with policy changes (Browne-Nuñez et al. 2015; Hogberg et al. 2015) The former study used focus groups and data analysis (quantitative and qualitative) undertaken both before and after delisting and changes in policy that re-initiated lethal control of wolves. Browne-Nunez et al. (2015) found that attitudes to wolves did not change, and poaching plans appeared to stay the same or increase. Calls to kill more wolves through public hunting and trapping actually increased. Hogberg et al. (2015) used a quantitative mail-back survey to residents living in wolf range, comparing attitudinal measures from the same persons sampled in 2009 and resampled in 2013. Hogberg et al. found that a sample of residents of wolf range in Wisconsin had lower tolerance for wolves compared to their tolerance in 2009. The most significant change in policy was the inauguration of a wolf hunting and trapping season in 2012, after delisting and associated lethal control in two periods from 2009-2012. Therefore, the best available evidence shows that tolerance for wolves and inclination to poach wolves actually **increases** after delisting and during periods of liberalized wolf-killing .

Also, the USFWS misinterprets the review by Treves & Bruskotter (Treves & Bruskotter 2014), which inferred that policies and social norms can rapidly change attitudes to lower tolerance for predators and increase inclinations to poach predators. Missing the former articles and the point of the latter, the USFWS then asserts,

"Thus, it is unclear how delisting and the changes in wolf management subsequent to delisting, such as implementation of wolf harvests, may affect attitudes, human behavior and, ultimately, wolf mortality. " p.9662.

Given two independent empirical studies using different methods and different subjects at a different date, albeit in the same region, and a review cited above, the proposed rule errs in not summarizing these findings. The USFWS funded un-cited studies led by Browne-Nunez (a USFWS employee) and by Hogberg, so the agency cannot fairly claim ignorance about them. Also, Bruskotter et al. 2013 (entire) long ago addressed the weaknesses of the USFWS position on human attitudes to wolves.

The proposed rule's citation to Paul in litt. 2006, above, and the draft biological report's lack of scientific review of individual wolf-killing are troubling insofar as it seems to promote favorable, unqualified opinions over unfavorable evidence rather than addressing uncertainty, fundamental scientific debates or the weight of evidence. It is striking how diligently the proposed rule quantifies minor details of wolf biology (e.g., p9662 has dozens of citations and statistics on wolf use for various road densities) but entire sections on human attitudes and mortality risk and rate remain unquantified, despite the USFWS having access to studies that quantify these phenomena. No scientific justification supports that emphasis. Therefore, I find the proposed rule (and draft biological report next) are unscientific on their face.

Draft biological report

I generally found the biological report to be fairly strong within the limits of what it covers, except for the section on Wisconsin's wolf census and population model (see my Appendix) and on Minnesota's wolf census (see below). I write "within the limits" advisedly because the biological report overlooks essential information. Specifically, it substantially omits evidence relating to human-caused mortality and the cumulative effects of all causes of mortality or reproductive failure of wolf packs. I begin with the substantial gap and end this review of the biological report with suggestions for revision of several misleading paragraphs.

A major problem with the biological report is the entirely absent handling of the biological causes of wolf mortality. Namely, the biological causes of the vast majority of wolves are humans and the process and pattern of (and intervention against) human-caused mortality is completely ignored in the biological report. By analogy, if one avoided discussion of a nonhuman predator and its predation on another endangered species, one's findings would be seen as having a glaring gap. Likewise the ecology of the wolf as presented in the biological report fails to discuss the wolf's primary predator, humans. This gap leads to major errors in the proposed rule. The biological report should be revised substantially, to deal scientifically with the patterns and processes of human-caused mortality in wolves.

Understanding human-caused mortality of wolves requires an understanding of human attitudes, because much of the mortality of wolves is caused by human intentions rather than accidents. The biological report is wholly lacking in pertinent vital information on human attitudes and behavior as they relate to human-caused mortality in wolves. Therefore, the many claims about the survival of wolves after delisting, claims about suitable habitat and consideration of the risk of extirpation after delisting are not informed by a scientific analysis of the peer-reviewed literature on human-caused mortality.

For the biological report to address cumulative effects of mortality and reproductive failure in wolves of the WGL and PNW, it must be revised to meet the scientific mandates of the ESA five-factor analysis. It is impossible to scientifically evaluate the likelihood of a decline of gray wolf populations after delisting without a thorough and comprehensive look at all mortality causes within each subpopulation deemed essential to the gray wolf entity, followed by a thorough examination of cumulative effects across all subpopulations. Scientists need rigorous, peer-reviewed evidence concerning current causes of mortality plus causes anticipated following delisting, along with reasonable estimates of probable catastrophic declines. Absent such evidence, a reasonable scientist cannot assume that states can respond effectively and rapidly to prevent population declines below relisting levels, nor can we evaluate the likelihood of substantial changes in mortality rate for gray wolves in the foreseeable future. Assurances from one or even a handful of self-appointed experts do not constitute valid in scientific conclusions.

It is not adequate to assure us that Canadian wolves will repopulate WGL or PNW subpopulations without offering any supporting data on wolf immigration from Canada.

Likewise, the biological report also lacks any meaningful assessment of standards of evidence. It contains questionable conclusions that do not adequately consider disagreements within the scientific literature, basically treating them as if all sources of evidence were equal in strength of inference.

Instead, different sources should be evaluated in light of each other. Stronger inferences, hence stronger evidence, comes from sources that are more transparent, that consider contrary evidence and provide scientific reasons for why certain sets are weaker, and then present either a better method for drawing inferences or explain contrary results in a more robust manner. A paper that ignores contrary evidence fails on its face. These scientific principles are well illustrated by the recurring problem of citing Gude et al. (2012).

Vucetich (2012) explained why the claim that 48% anthropogenic mortality would be sustainable is in error and misleading. Until Gude et al. (2012) either correct their estimate, or else scientifically rebut the observations of error asserted by Vucetich (2012), Gude et al.

(2012)cannot fairly be listed as best available science. I view the following statements about Gude et al. (2012) as scientifically irresponsible:

"Some studies suggest that the sustainable mortality rate may be lower, and that harvest may have a partially additive or even super-additive effect (harvest increases total mortality beyond the effect of direct killing itself through social disruption or the loss of dependent offspring) on wolf mortality (Murray et al. 2010; Creel and Rotella 2010), but there is substantial debate on this issue (Gude et al. 2012)."

In my view this quotation is inaccurate to the point of being misleading. First, there is no credible evidence to contradict the finding that human-caused mortality has led to more wolf deaths than expected from permitted killing. Second, a plausible mechanism for the super-additive mortality has been proposed (Box 1 and citations to Brainerd et al. and Borg et al. above). Although, super-additive mortality resulting from instability or infanticide in wolf packs is hotly debated at present, the more likely mechanism proposed for super-additive mortality and the accelerating decline of wolf populations modeled by Vucetich (20125) is that legal killing prompted illegal unreported killing of wolves (Chapron & Treves 2016a,b, 2017a,b). And more recent results corroborate that idea with independent data (Box 1).

My last specific point relates to the Minnesota wolf census of 2012–2013 (Erb & Benson 2013). Given the ostensible importance of the Minnesota wolf population for the security of the WGL wolves, it would seem important to the USFWS to validate the Minnesota wolf census if one wishes to persuade scientists that this lynchpin population is secure. That case should be made in an appropriate scientific journal in front of truly independent reviewers if the claim of best available science is to be made. The USFWS has had years to encourage such a move towards the best available science. I am not aware if this census or a subsequent one has been subjected to rigorous, scientific peer review.

If I were asked to peer review Erb & Benson (2013) as it currently stands, I would point out several shortcomings in the methods that need revision. The major shortcomings are a lack of appropriate sensitivity analysis and handling of uncertainty, which together undermine my confidence in the accuracy of the 2013 Minnesota wolf population estimate and its putatively narrow confidence interval. With even a conservative effort at sensitivity analysis, the lower bound might cross the delisting threshold of 1251. Therefore, I describe those shortcomings below for the purpose of recommending that a revision to the draft biological report wait for a more scientific estimate of that population size as I attempt below.

Wolf biologists including myself have long acknowledged that it would be a daunting task to count every wolf pack in Minnesota. I accept that one must extrapolate from a sample. Such extrapolations are tricky and demand sensitivity analysis after transparent and rigorous handling if uncertainty in measurements, which further emphasizes the need for peer review of how extrapolation has been done. Extrapolation is tricky because of the risk of introducing sampling bias (how was the small sample of wolf packs chosen to treat it as representative of the state as a whole?) and measurement uncertainty (was the sample measured precisely and accurately to avoid propagating error by multiplying inaccurate or imprecise variables with each other?), among other issues.

Regarding the uncertainty due to sampling bias, the Minnesota estimate depends on "mean scaled territory size" estimated from a sample of 36 wolf packs (see their Figure 2 for the geographic distribution of those territories, two packs did not provide midwinter pack size counts). These 36 represent 8% of the total number of packs in the state, which on its face raises concerns about representativeness of even the best random sample of packs. The selection of those packs was not random but appears haphazard (without relation to the goal of extrapolating to the state), so we have no information on how representative the 36 wolf packs might have been. The analogy would be if the U.S., Census Bureau did not randomly sample for its long-form census but instead relied on a handful of other studies that had selected their samples by unknown and different criteria. Had the wolf packs been chosen randomly we might extrapolate with confidence. Haphazard sampling is not the same as random sampling because the more convenient sites or wolf packs tend to be sampled. For example, a common practice in carnivore biology is to study animals that are not likely to be killed during your study period, which would waste time and resources.

For example, wolf packs that are vulnerable to annual lethal control as conducted in Minnesota might be poorly represented in the sample of 36. Such packs subject to lethal control also tend to be smaller than average in Wisconsin (Wydeven et al. 2004a) for a variety of reasons. Therefore, extrapolating from larger packs selected for the convenience of other studies might tend to inflate the apparent size of the statewide population. I did not attempt to account for that difference between packs because I chose to be conservative in my sensitivity analysis. But someone else employing the precautionary principle, upheld by the U.S. Supreme Court ruling on the first ESA Case in TVA v Hill 1978, might argue that an endangered species decision demands precautions, so one should assume the majority of wolf packs in Minnesota average smaller size than the 36 wolf packs reported by Erb & Benson (2013).

Regarding measurement uncertainty, consider the following equation used to estimate N, the Minnesota wolf population size estimate: N = ((km2 of occupied range/mean scaled territory size)*mean pack size)/0.85." p.3, (Erb & Benson 2013). This pleasingly straightforward and simple equation raises a few questions about measurement uncertainty.

From left to right, the first variable is the area of occupied range. A precise estimate of the possible error in this value is impossible to gather from(2013), but they report ""total wolf range was estimated to be 95,098 km^2", p.4 and they provide a qualitative sense of uncertainty in the following statement, "Of the total estimated occupied range, 70% was confirmed to be occupied based on pack detection in the township and 30% was presumed to contain packs because of low human and road density..." p.4, (Erb & Benson 2013). Therefore, I treat the lower bound of their estimate of occupied area as 70% of the estimate they ended up using of 70,579 km^2 (i.e., eliminating the 30% of area for which no wolves had been observed). Note that a less conservative approach would require that evidence of breeding wolves be provided for every mapping unit (township) included in the range. The logic behind such a criterion would be further strengthened by the equation's enormous (and unjustified correction for lone wolves, see further below)..

The next variable was the average territory size. My efforts at estimating uncertainty could not unravel how the wolf pack territory sizes were estimated, given that radio-locations were removed from range estimates and two subsets of packs were handled differently – 19 estimates of size of well-studied packs' territories were used without correction but 17 estimates of size were scaled up by 37%, because too few telemetry locations were obtained, to correct for putative under-estimates (Erb & Benson 2013). These procedures might be valid but it is currently impossible to be certain. Nevertheless, we can estimate a minimum level of uncertainty. Had we been given the standard deviation of territory size (Erb & Benson 2013), which is an elementary statistic, we might have calculated the uncertainty around that average. Without that information, I will assume the standard deviation (SD) of pack territory size in Minnesota is the same as in Wisconsin from 2001–2006 (Wydeven et al. 2009), which was ±67 km^2. Hence, I estimate the standard error of the mean (SEM = SD / square root of n) for the 161 km^2 average territory size (Erb & Benson 2013) was 11 km^2. Again, a less conservative approach would have used the SD itself, not the narrower SEM.

Likewise, mean midwinter pack size estimated from 34 wolf packs was 4.3 wolves in Minnesota (Erb & Benson 2013) and probably an over-estimate statewide for the reasons I discussed above. By comparison; in 1997, all Wisconsin wolf packs averaged a size closest to the Minnesota pack size estimate, at 4.1 wolves in midwinter with SD ±2.1 (Wydeven et al. 2009). Therefore, I used that SD to estimate the SEM for the 36 Minnesota wolf packs at SEM 0.4. Again, a less conservative approach would use the SD or even the statewide pack size average from Wisconsin in 2013 (after a wolf hunt like in Minnesota) that was closer to 3 wolves per pack.

The last value in the above equation that has uncertainty is the apparent constant of 0.85, which was used as a correction factor for lone wolves, citing (Fuller et al. 1992). To estimate uncertainty about this value, I use both the estimate of the average and the SD from Wisconsin of 3.2% ± 0.7 SD during the period 1996–2007 (Wydeven et al. 2009). I bracket the 0.85 constant between 0.843 and 0..975 to capture the Minnesota estimate minus the SD from Wisconsin and the Wisconsin average plus its SD. Again, a less conservative approach would use 0.97 and its SEM not the 0.85 value that could substantially inflate the results of the equation and is the least well-substantiated parameter.

What does the above exercise in estimating uncertainty tell us? If we take the margins of error above and apply them to the equation on p.3 in (Erb & Benson 2013), we see the effect of propagating error in a far wider range of possible sizes of the Minnesota wolf population:

Maximum N = ((95098/150)*4.7)/0.843 = 3534 wolves

Minimum N = ((49405/172)*3.9)/0.975 = 1148 wolves

That range of estimates for the Minnesota wolf population size (1148–3534) is a conservative minimum estimate of the true uncertainty in my view, yet admits higher uncertainty than the official estimate of 1662–2640 (Erb & Benson 2013). I am not clear from their report how they claim such certainty. In particular, my lower bound of 1148 is substantially lower than the state's lower range. Either of us could be criticized for preferring our answer over the others answer, but science does not respect authority, only evidence. Given the short time for this peer review and the scarcity of raw data for Minnesota, I do not place great faith in my own calculations or those of Erb & Benson (Erb & Benson 2013). We need a fully transparent, comprehensive review to make a scientific judgment.

The precautionary approach upheld by the U.S. Supreme Court in TVA v Hill 1978 suggests a closer look at the Minnesota wolf population estimate and methods for extrapolating. Indeed, my estimate falls near the lower bound of the federal relisting threshold of 1251 wolves in MN. Given the risk and uncertainty, I recommend all the MN data be presented transparently with basic statistics about standard deviation and with sensitivity analyses applied at every step. There is a possibility that Minnesota contains fewer than 1251 wolves today.

In general, I find the draft biological report ignores a large number of relevant articles published in peer-reviewed journals of the highest rank. Being unaware of them does not seem plausible given the USFWS paid for some of the research in these articles and were sent many of them in previous rounds of delisting. If the draft biological report was written without the need to review past delisting efforts for gray wolves, these sorts of errors will persist, so I recommend a review of the policy for writing draft biological reports as a matter of scientific integrity. Moreover, the articles included appear haphazard. The USFWS shared with peer reviewers an array of non-peer-reviewed sources and lower ranked peer-reviewed journals. From this, I glean the following recommendation on future drafts of biological reports for wolves or other species.

The peer review process ideally provides independent validation of scientific findings. Personal communications, conference proceedings, reports from agencies, journals without editorial policies, etc. do not meet these standards and are therefore less reliable and provide weaker inference no matter how much one likes the conclusions. Of course, the ideal of peer review is sometimes not met even within the peer-reviewed literature. The literature falls short of the ideal when reviewers or editors share the same biases as the authors or when editors or peer reviewers take shortcuts in evaluating the evidence. Therefore, strong inference and progress in science depends on critical evaluation of the evidence itself using accepted standards of evidence and a lengthy process of scientific deliberation and consensus-building. The USFWS has had time for such deliberation and consideration given 20 years of contemplating delisting gray wolves (Refsnider 2009).

Even without consensus among scientists and extensive discussion, one can use long-established standards of evidence to rule out some of the evidence relied on by the draft biological report and proposed rule. For example, controlled comparisons (i.e., those in which a statistical control or experimental control is used) provide stronger inference than anecdote or correlation. Among controlled studies, random-assignment experiments are the gold standard in biomedical research (Ioannidis 2005) as in predator science alike (Treves et al. 2016). Furthermore, within controlled experiments, careful reading of methods can identify systematic bias in sampling, measurement, treatment, or reporting. The USFWS can set an important example for agencies in lower jurisdictions if it restricts itself to drawing conclusions from studies providing the strongest inference, settling for weaker inference when truly nothing else is available. Many examples of such careful discrimination are available in the scientific literature on wolves. Thanks for considering these recommendations for improvement.

End of main text of peer review

Appendix. Reanalysis of wolf population dynamics and monitoring in Wisconsin 1999–2013

Summary statement: The scientific debate about census, monitoring, and the effect of lethal management of wolves in Wisconsin is important to the proposed rule because a foundational claim of the proposed rule was that WGL wolves would be safe after delisting. Ostensibly, WGL wolves would be safe because wolves are being monitored scientifically (when alive and when dead), state management plans are science-based, and therefore that states will swiftly detect if their wolf subpopulations are in jeopardy and respond protectively. The information in this Appendix casts doubt on that assumption. The Appendix below and the following webinar (https://zoom.us/recording/play/unVXR_qjH5QJ0sp3rBbECrNTGR08x1yEmWEocSYSY6SZHihUrQlPQdc-QUk2NrJf?continueMode=true) detail why the state wolf population model and wolf census have been based on incomplete information that requires scientific revision and peer review. The appendix below begins the review and revision including the period 1994–2013 and covers material in the 1999 Wisconsin wolf management plan, its 2007 addendum, and at least half a dozen peer-reviewed papers involving the architects of the state wolf population model and wolf census. The information provided below undermines the scientific rationale for the current state's population goal, state delisting goal, and state harvest models. Although this appendix addresses only Wisconsin, a proper accounting by the USFWS would help to reveal if problems are more widespread.

A. The state wolf management plan of 1999

In 1999, the wildlife agency of the state of Wisconsin (WDNR), published its first wolf management plan (1999 Plan), as the USFWS prepared for delisting wolves in the WGL (Refsnider 2009). I use the federal government's population estimates (USFWS 2018) because they had primary jurisdiction over wolves for most years 1979–2019.The 1999 state wolf management plan included a projection of wolf population growth to 2020 (Figure A1), from the estimated 205 adult and yearling gray wolves counted in winter 1998–1999 (WDNR 1999).

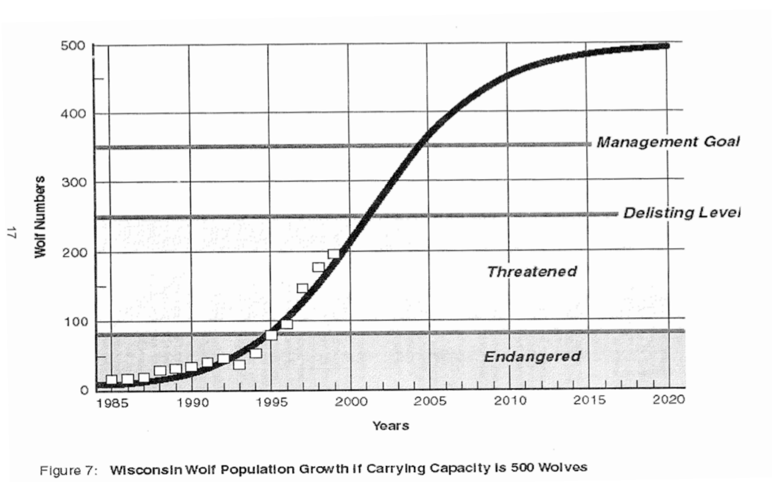

Figure A1. The 1999 Plan's forecasting model of wolf population growth (their Figure 7 (WDNR 1999). The 'Delisting Level' was set at 250, when the legal removal of wolves from the state's list of threatened and endangered species would begin. The 'Management Goal' codified a population target (N_{goal} = 350). N_{goal} is still the state population target today (Stepp 2013). Also note the estimate of carrying capacity (500 wolves) represented one of several estimates in 1999, all of which estimates later proved too low (Appendix).

The 1999 Plan codified two numerical values of population estimates that would alter policy. The first, Delisting Level, was set at 250 (Figure A1) and was apparently set at the midpoint of the range of outputs of a population viability analysis (PVA). The 1999 Plan stated that the PVA, "needs to be cautiously interpreted and should not be used by itself to set management goals… Based on [the Wisconsin PVA], a population between 200 to 300 seemed appropriate for delisting wolves in Wisconsin." p.16, (WDNR 1999). The second value was the Management Goal N_{goal} = 350 (Figure A1) also referred to as the 'population goal'. Ostensibly, N_{goal} was chosen to exceed the USFWS delisting criterion (N = 100), supportable by suitable habitat, compatible with the PVA mentioned above, and "socially tolerated" p.15, (WDNR 1999). Several scientific questions arise from Figure A1 that have not been answered adequately.

The first question relates to the dark logistic growth curve in Figure A1. It appears to be a model fit to the population estimates 1980–1999 and a projection of future growth. Importantly, the 1999 Plan made no mention of that logistic growth curve super-imposed on the population estimates (Figure 1A), or its relationship to an assumption of density-dependent population dynamics and harvest planning (described further below). The assumption of density-dependent population growth was common but far from universal in that a substantial number of wild animal populations did not show such dynamics (Fowler 1987). Apparently, the assumption was made without supporting evidence beyond fitting points to half of such a curve and an estimate of carrying capacity that proved inaccurate (see below). New evidence also suggests density-dependent dynamics are unlikely to characterize Wisconsin wolf population growth from 1980–2012. Subsequently, Brook & Bradshaw (2006) explained many of the reasons why populations like the Wisconsin wolf population might not show density-dependence, including (a) substantial errors or changes in sampling or measurement can mimic or obscure density-dependent dynamics (as we show below); (b) populations growing without spatial bounds and limited mainly by exogenous factors are not expected to show density-dependent dynamics; and (c) minimal changes in density over the sampling period might not produce density-dependent dynamics. Below I review how all of these factors played a role in Wisconsin's wolf population dynamics and the science used to model them.

The second question arose from introducing the criterion of human tolerance (a value judgment). Human tolerance was first introduced by an informal survey about N_{goal} after it had been selected, and the survey was reported without methods in the 1999 Plan. "During the review of the second draft of the [1999 Plan], of persons commenting on the population goal, 38% supported the goal, 38% felt it was too low, and 24% felt it was too high… a reasonable compromise between population capacity, minimum level of viability, and public acceptance." p.16, (WDNR 1999). That small-sample opinion poll probably consisted of individuals in a wolf advisory committee (1999 Plan), not by randomly-sampled, broad public opinion. Later scientific surveys did not support the result, starting in 2001 (Naughton-Treves et al. 2003; Treves et al. 2009a; Treves & Martin 2011). In 2004, the survey with the most representative sample of 1364 respondents sampled randomly from 6 postal areas stratified by region (wolf range or not) and human population density (urban or rural) found that 57% wanted a higher N_{goal} of 500, 1000, or "No cap". N_{goal} itself only garnered 15% support (Naughton-Treves et al. 2003; Treves et al. 2009a; Treves & Martin 2011). Claims to the contrary have been inaccurate or misleading (Groskopf 2014).

The third question that soon arose was about the estimate of carrying capacity in Figure A1. Carrying capacity (K) played an important role in codifying N_{goal}, (p. 15–16, WDNR 1999). Two estimates existed for K (both were exceeded by population growth by 2008–2009): a habitat-based estimate of K = 300–500 wolves or possibly up to 800 if marginal habitat were to be occupied, and a prey-based estimate of K = 262–662 (Mladenoff et al. 1995; Mladenoff et al. 1997). It does not seem a coincidence that the Delisting Level was half of K = 500, and the Management Goal was 53% of K = 662 (Figure A1). Setting population targets at half of K is characteristic of harvest models aiming for maximum sustainable yield (MSY). Such models are designed to kill as many organisms as possible annually, within presumed limits of sustainability. Yet, the 1999 Plan did not explain the model that had been used to set the dark logistic growth curve, the Delisting Level or the Management Goal in Figure A1.

By contrast, the value judgments underlying N_{goal} in the policy side of the 1999 Plan were clearer. The 1999 Plan read, "[N_{goal}] was intended to be the minimum level at which proactive control and public harvest would occur." (p. 16, WDNR 1999). Indeed, the first published draft of the 1999 Plan also included an appendix J, which qualitatively described a future framework for regulation of hunting and trapping wolves. Appendix J was eventually removed from the draft after public opposition (Treves 2008). Yet the science did not change.

Objectivity is a core principle of scientific integrity and therefore important to distinguishing the best available science from biased science (National Academy of Sciences et al. 1992). For the 1999 Plan to have been prepared in a scientifically objective way ("Of a person or his or her judgement: not influenced by personal feelings or opinions in considering and representing facts; impartial, detached." Oxford English Dictionary online 2019), one or more of the following scientific steps should have been taken: presenting alternative hypotheses about population growth (e.g., not density-dependent over the period in question); scrutiny of the underlying assumption of sustainability at N_{goal} or MSY; and exploring alternative scenarios if key parameters were found to differ substantially, e.g., K was exceeded. None of those steps were taken to my knowledge.

B. Changing wolf census methods from 1994–2004

Because wolves are cryptic, territorial, as well as ecologically and socially complex, interpreting wolf population estimates accurately requires a detailed understanding of the census methods used to arrive at population estimates. The source data from unpublished WDNR population reports are no longer easily available to the public, so I present them at http://faculty.nelson.wisc.edu/treves/data_archives/WDNR%20ER%20Bureau%20reports.zip. I use these reports to support the argument that there are four time-series of wolf population estimates in Wisconsin not one time series as has been erroneously depicted (Figure A1 for example). These reports support the following facts:

- From 1979–April 1994, the WDNR deployed a few citizen volunteers to help to census wolves, but in the winter of 1994–1995, the WDNR began a volunteer program for private citizens to track wolves in snow, by driving or on foot, independently of the WDNR (Troxell et al. 2009). In the winter of 1994–1995, the WDNR augmented the number of volunteers from 12 to 40 and those volunteers undertook independent tracking (Wydeven 1994; Wydeven & Megown 1995); see p. 12 and p.13 respectively. The following year, the WDNR deployed 83 volunteers, see p. 16, (Wydeven & Megown 1996). The start date of 1994–1995 when volunteers tracked wolves >526 km has been confirmed several times, most recently in (Wydeven et al. 2009); although a slow start compared with the thousands of km volunteer trackers would later drive (Table 6.1, (Wydeven et al. 2009). Every year thereafter, 50–100 volunteers tracked independently in the snowy months. Census methods changed again approximately 10 years later, but first I examine the effects of the 1994–1995 change.
- The first change in methods strongly affected the population estimate of April 1995, as evidenced by the following: "The 1994-1995 wolf population was 66% above the wolf population present in 1993-1994 (50-57 wolves). This increase probably represents more than just natural reproduction. Some wolves were probably missed in 1993-1994 surveys." p. 10, (Wydeven & Megown 1995).

The above spike and discontinuity in the population estimate can be seen more clearly in Figure A2 than in Figure A1. It represents a break in the first time-series 1980–1994 that created the appearance of a change in shape of the population growth curve. I will argue that this year produced the first (misleading) impression of density-dependent dynamics in this population. My inference is consistent with cautions published later by (Brook & Bradshaw 2006) that sampling effort or measurement methods account for the appearance of a change in slope.

Regrettably, in addition to an illusory pattern of density-dependence in my view, the timing of changes in wolf census methods became confused by an apparent error in the 1999 Plan.

The 1999 Plan stated, "A volunteer carnivore track survey was initiated by the WDNR in fall 1995 [sic] (Wydeven et al. 1996)." p. 20, (WDNR 1999). The reference to fall 1995 is an error that might have come about unintentionally. The 1999 Plan referenced Wydeven et al. 1996 that presented guidelines for those surveys, not the population report by Wydeven & Megown 1995 that presented the first evidence of such a volunteer tracker program (and confirmed in 2009 as summarized above). I presume the one-year discrepancy was an unintentional error in the 1999 Plan, possibly caused by the common confusion that a population report published in the summer of year t+1 was collated from records including the second half of year t, therefore the 1995 population report I cited above for a 66% increase in the number of wolves included late 1994 and is an official government population report in a format of a scanned image which is non-editable (Wydeven & Megown 1995); it precedes the 1996 guidelines on volunteer tracking cited in the 1999 Plan.

Confusion about timing was not restricted to the 1994–1995 change in census methods. For example, the following sentence appeared first in the winter 2003–2004 report, "All volunteers were required to attend weekend wolf ecology courses and day-long track training programs."

p. 5, (Wydeven et al. 2004b). Such training sessions had been implemented by the second half of 1996 and were a regular part of the program every year thereafter.

In all, I found four changes in census methods had been implemented piecemeal: addition of volunteers 1994–1995, training in the second half of 1996, some types of quality control on volunteer tracker data between summer 2000 and winter 2003–2004, and an overhaul of methods and volunteer turn-over in 2012–2013 described in section E below.

Changes in census methods affected the accuracy of wolf census. Importantly, wolf census by volunteer trackers improved the population estimate and routinely differed from those by WDNR staff members surveying the same census blocks at different times. The following quotations substantiate both assertions. The 2003 conference presentation entitled "Counting wolves--integrating data from volunteers", asserted that volunteer trackers improved the population census:

> "Wolves recolonized the state of Wisconsin in the mid-1970s after being extirpated for about 15 years. Between 1979 and 2002, the Wisconsin Department of Natural Resources (WDNR) maintained constant monitoring of the state wolf population through live-trapping and radio tracking, winter track surveys, summer wolf howls, and public reports. The wolf population in Wisconsin ranged from a low of 15 (1985) to 327+ in 2002 and ranged from 4 to 83 packs. Mean estimated pup survival from late gestation to about 9 to 11 months was 0.30 (range 0.16 to 0.57). During periods of disease outbreak of Parvovirus (mid 1980's), and initial outbreak of Sarcoptic mange (early 1990s), pup survival declined to < 3.0 during periods of severe disease outbreak. A mean of 36% (SD + 16.4%) of packs had no surviving pups in late winter. Survival of older wolves (1+yr) was 0.61 during the early 1980s when wolves declined but increased to 0.82 in the late 1980s and early 1990s and has remained high in recent years. Since 1985 the wolf population increased an average of 20% annually and experienced a slight decline during only one year. Areas occupied by wolf packs expanded from about 1500km 2 in the early 1980s to about 13,000km 2 in 2002. The wolf population continues to increase, but growth may decline or stabilize soon as most suitable habitat is occupied, and more liberal lethal controls are enacted." abstract, (Wiedenhoeft et al. 2003).

I do not have the full transcript of that presentation, so I do not know the period of censuses it covered. But the differences between volunteer trackers and WDNR census-takers began to be quantified in population reports by the winter of 2003–2004:

> "Forty-nine blocks were surveyed by both [W]DNR and volunteer trackers. [W]DNR detected 159–180 wolves, and volunteer trackers detected 130-143 wolves in these blocks. [W]DNR detected more wolves in 17 blocks, less [sic] wolves in 17 blocks, and the same in 15 blocks. In 19 of the 34 blocks where counts differed, the group with the higher count had tracked considerably more miles than the group with the lower count. In 5 of the 15 blocks where the counts were the same, both detected 0 wolves." p.9, (Wydeven et al. 2004b).

By contrast, in 2006, the state reported that volunteer trackers found more wolves in the same censes blocks counted by WDNR staff:

> "Both DNR and volunteer trackers surveyed wolves in 41 survey blocks. Overall rates of wolf detection were similar with 147 - 159 wolves detected by DNR trackers and 159- 173 wolves detected by volunteers. DNR detected more wolves in 13 blocks, volunteers detected more wolves in 19 blocks, counts were the same in 7 blocks, and no wolves were detected in 2 blocks. Overall rates of wolf detection indicate volunteers are providing suitable counts of wolves." (p.9, (Wydeven et al. 2006)

The differences between volunteer tracker and WDNR counts varied annually and the differences seemed to relate to search effort. For example, in winter 2003–2004, 55% of census blocks showed different wolf counts by volunteer trackers and WDNR staff, when calculated by the number of census blocks that differed divided by the total number of census blocks that both groups counted; or alternately, the wolf count summed across census blocks that both groups had counted differed -18% to -21% when calculated by the number of wolves counted by volunteer trackers divided by the number that WDNR staff counted in the same census blocks (Wydeven et al. 2004b). By contrast, the winter 2005–2006 report allowed calculations as above of 78% and +8% to +9% respectively (Wydeven et al. 2006). Although the reports from 2004 to 2012 suggest WDNR staff more often counted more wolves than volunteers counted in the same census blocks, the differences were clearly large and variable to the point where they sometimes exceeded the annual average for population growth (see below).

The unpublished population model and its revealing logistic growth curve (Figure A1) in the 1999 Plan might have been materially influenced by the addition of independent volunteer trackers in winter 1994–1995 and beyond. I expect the model would have changed because the average and variance of population estimates changed. However, the architects of the state wolf population model[1] did not transparently describe the above changes or its implications for the 1999 Plan or underlying, unpublished model.

I present a revised figure for Wisconsin's wolf population growth that reflects three time-series rather than one, as I can best recreate the appropriate start and end dates from the above information for 1980–2012. But I retain the official population estimates because providing alternative estimates is beyond my present scope (Figure A2).

[1] In the 2004 population report, the architect of the wolf census introduced the public to the architect of the wolf population model as WDNR staff, until October 2005–March 2006 when the University of Wisconsin–Madison employed the latter. "Between October 2003 and March 2004,… agency personnel were asked to report wolf observations to Tim Van Deelen with DNR Science Bureau." p.8, (Wydeven et al. 2004b) and then, "New research by graduate student Elizabeth (Lizzy) Berkley and Dr. Tim Van Deelen will be conducted, examining fatty acids to determine diet of wolves" p.11, (Wydeven et al. 2006).

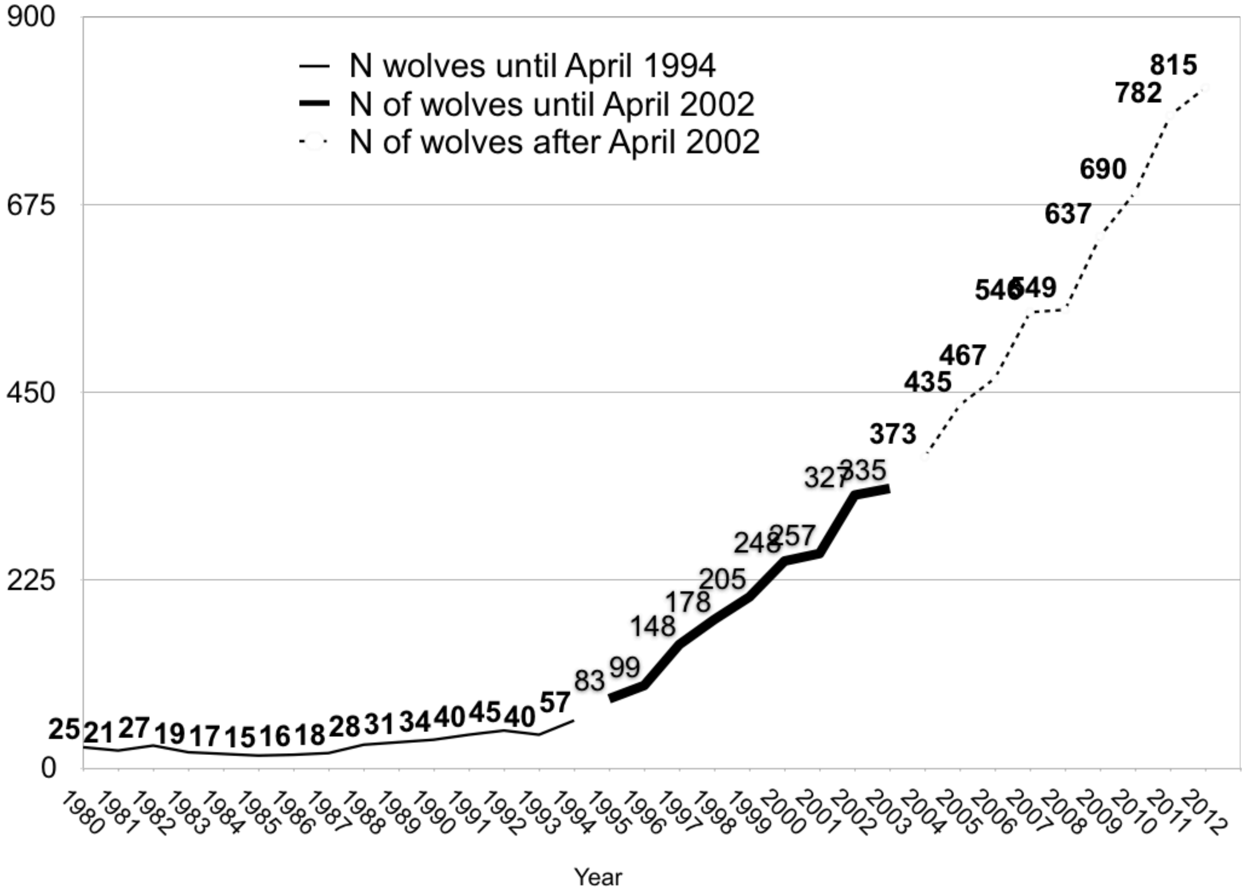

Figure A2. Wisconsin's wolf population estimates in three time-series, 1980–1994, 1995–2003, 2004–2012. The census population size of adults and yearlings in late winter depicted in three consecutive time series: the first (thin, solid line) before volunteer trackers counted wolves independently for the WDNR; the second (thick solid line) marks the major increase in deployment of independent volunteer trackers, until mid-2003 when the WDNR announced the quality controls over volunteer tracker data and search effort. The transition from the second period to the third period might have occurred as early as winter 2000–2001 or piecemeal during the interval 2000–2003. A fourth time-series also seems to have started after winter 2012–2013 (Appendix section E).

When I analyzed the population estimates in three consecutive time series (Figure A2), I found significant differences in annual growth rate averages or standard deviations between time series. Using $(N_{t+1} - N_t) / N_t)$, the annual average increase in the first period was 8% (sd 23%), then from April 1995 to April 2004, it was 22% (sd 15%, two-tailed test with unequal variances X^2=26.9, p=0.025). Instead, if one ends the second period after winter 1999–2000 (Wiedenhoeft et al. 2003), the average was even higher (29%, sd 15%). In the third period 2005-2012, growth averaged 10% (sd 6%, unequal variances X^2=62.9, p=0.001). Even if one demarcates the third period after April 2001, the average growth fell by >50% to 11%. The higher sd in period two than period three suggests that volunteers and WDNR staff might have compensated for undercounting in year *t* with greater effort or redirected effort in year *t+1*. Also the first two winters of volunteer tracking (1994–1996) seem to account for the high average growth rate in time series 2 and the appearance of a slow-down by time series three. The latter appearance of a slow-down might have created the illusion of density-dependent dynamics.

Also, one should consider that in April 2003 the WDNR began issuing permits for lethal control of wolves, adding an exogenous limiting factor, which is not evidence for an endogenous density-dependent factor limiting wolf population growth. Regardless of how one interprets the population estimates and their changes over time, scientific integrity demands disclosure of the changes in methods (National Academy of Sciences et al. 1992) associated with changes in population estimates and average and standard deviations of growth rates.

I conclude that a significant methodological artifact permeates the population estimates for Wisconsin's wolves from 1980–2012. Moreover, the initiation of lethal management of wolves on 1 April 2003 – after quality control was imposed on volunteer trackers' data between 2000 and 2003 – and its cessation or re-initiation in five more periods until 15 April 2012 (Chapron & Treves 2016a), would complicate any simple inference that population growth showed density-dependence.

Because the effect of the volunteers described above was only detected after the 1999 Plan and perhaps only fully grasped in mid-2004, the ensuing years would have been critical for remedying the problems with the unpublished population model, improving the descriptions of methods, and revealing the potential effects of those changes to the public. After 2004, an opportunity arose to improve transparency of wolf population science; explicitly and objectively discuss assumptions, scenarios, and alternative hypotheses; subject census methods and population models to independent peer review; and encourage replication. Little of this was attempted to my knowledge. In sum, the science in the 1999 Plan had gaps in both transparency and objectivity. Those gaps could have been easily remedied scientifically between 1999–2006, but they were not when the WDNR published an addendum (WDNR 2007).

C. Addendum to the 1999 Plan

Collaborations between the architects of the wolf census and population model soon led to the second wolf population model. However, again the WDNR and the architects themselves did not publish their work in a way that could be reviewed by peers or the public. In the 2007 addendum to the 1999 Plan, the architects of the state wolf census and population model wrote,

> "Van Deelen (unpublished) fit simple growth models to a XX [sic] year time series of wolf population estimates. Models fit were the discrete logistic model (CITATION) [sic] and the discrete Ricker model (1975) of the general form Nt+1 = f(Nt) where N = population size. Model fitting was based on a least squares algorithm and jackknife procedures were used to generate variance estimates because of the inherent temporal autocorrelation (Dennis and Taper 1994). The best fit logistic model estimated an equilibrium (or carrying capacity) of 505 (95% C.I. = 501 - 518, P <0.0001, R^2 = 0.99) whereas the best fit Ricker model estimated an equilibrium of **522 (95% C.I. = 295 - 635, P <0.0001 0.** [sic] R^2 = 0.99). Model selection criteria (Burnham and Anderson 1998) suggested that

> these 2 models were nearly equivalent given the data. Nonetheless, a Ricker model is probably more useful because of less restrictive assumptions about the shape of the growth curve. Despite wide use to characterize the growth in a time series of population growth estimates (Lotts et al. 2004) this model fitting approach has recently been criticized in favor of a risk analysis (Population Viability Analysis) that can be generated from the same data (Lotts et al. 2004). Still this exercise demonstrates that the original estimates of 300-800 wolves (depending on the extent to which marginal habitat was used) were reasonable and probably quite accurate." emphasis added, p.7, (WDNR 2007).

The long quotation above serves several purposes in this context. First, the 2007 addendum never mentioned the substantial differences between volunteer trackers and WDNR staff, nor the significant differences in averages or standard deviations of wolf counts in the different census periods I identified in Figure A2. Second, the time series used was omitted as was the citation to the method of fitting a discrete logistic model (see [sic] above), which perpetuated the 1999 Plan's lack of transparency.

I surmise the writing was rushed to meet the publication date: "The addendum to the wolf plan was presented and approved by the Natural Resources Board at their meeting on June 28, 2006 and updated on August 15, 2007." p.1, (WDNR 2007). The date is important because the architects reaffirmed their model outputs estimating K at 522 with confidence intervals of 295–635 (boldface phrase above) when the population had been estimated at 546 wolves in April 2007 (Figure A2) – yet also reaffirming estimates of K at 500–800 wolves. It is unclear why the models described in the quotation above were trusted when the predicted mean values were already surpassed. The architects seemed themselves to grant credence to the habitat suitability models that provided an upper bound of 800 wolves originally credited in the 1999 Plan.

Further details of wolf census and population model only surfaced in an incomplete fashion in 2009. In 2009, the architects co-edited a book on wolf populations in Wisconsin and surrounding states. The relevant chapters were entitled, "History, population growth, and management of wolves in Wisconsin" (Wydeven et al. 2009), and "Growth Rate and Equilibrium Size of a Recolonizing Wolf Population in the Southern Lake Superior Region" (Van Deelen 2009). Neither chapter cites their own findings of differences between volunteers and WDNR staff – p. 91-94 and Table 6.1 in (Wydeven et al. 2009).

The chapter on population growth and modeling presented only one scenario for killing wolves: "For example, given a growth rate of 1.31… an additive maximum sustained yield of 92 wolves (7%) would maintain the [Southern Lake Superior] population at 60% of estimated carrying capacity (770 wolves)." p. 150, (Van Deelen 2009). That author did not specify the model fully nor mention the changes in census methods, the apparent spike in population estimate in April 1995, why summing Michigan and Wisconsin's wolf populations tends to obscure the effects of volunteer trackers, or even the existence of volunteer trackers as shown by the following quotation:

> "These counts were supplemented by howling surveys and winter track surveys." p. 139, (Van Deelen 2009). "Biologists began radio-collaring adult wolves in

> Wisconsin in 1979 (Wydeven et al. 1995) and in Michigan in 1992 (Potvin et al. 2005) to facilitate the identity [sic] and location of wolf packs for aerial counts. **These counts were supplemented by howling surveys and winter track surveys**, but radio tracking has remained a centerpiece of population monitoring (Wydeven et al. 1995). Taken together, these efforts have provided rigorous annual counts of wolves." (p. 139, (Van Deelen 2009).

D. Independent science and an unresolved debate about Wisconsin's wolf population dynamics

Chapron and Treves independently published an alternative model starting with wolf population estimates from 1995 to account for the changes produced by volunteer trackers (Chapron & Treves 2016a, b, 2017b). They could not replicate the findings of the architects and their junior colleagues about population dynamics, but instead concluded that negative density-dependence on population growth was not apparent.

A lack of density-dependence in the period 1980–2012 need not be surprising. the fundamental assumption of density-dependence of Wisconsin's wolf population growth 1980–2013 has not yet been supported transparently or reproducibly. No density-dependence on adult mortality has been reported in any peer-reviewed report. Density-dependence on juvenile recruitment to adulthood is currently contested and vulnerable to problems of statistical dependence. Lack of a density-dependent change in life history parameters need not be surprising for wolves. For example, wolves in Western Poland grew substantially with no sign of density dependence (Nowak & Mysłajek 2016). Wolves may benefit from helpers (hence a higher local density) to raise a litter of pups, and for defending a territory against neighbors. Moreover, the very slight increase in density over time – 0.9% annually from 2000–2011 (Chapron & Treves 2017b), or even the 3% annually from 1995–2007 claimed by (Wydeven et al. 2009) – might be artefacts of measurement uncertainties about the geographic area occupied. Reproduction or mortality need not necessarily change with increasing population size or density. After all, population size may rise by geographic expansion without changing densities in the core areas. Or increasing densities within a pack territory might lead to more successful reproduction (positive density-dependence) if supernumerary adults help breeders reproduce and resources are not scarce at critical periods (Moehlman 1989).

Chapron & Treves (2016a, b) analyzed the time series of wolf population dynamics from 1995–2012 as a single time series, aware of the first change in census methods but unaware of the second (Figure A2). The architects and their junior colleagues responded, "[Chapron & Treves] selectively chose to analyse a subset of wolf population and life-history data (1995 –2012), yet these datasets extend to 1980 and 1989 for Wisconsin and Michigan, respectively (Beyer et al. 2009; Wydeven et al. 2009). Inclusion of the full range of wolf population and life-history data would probably have produced contrary results…" p.1–2, (Olson et al. 2017). Chapron & Treves rebutted, "[Olson et al.] insinuate that we chose to start our analysis in 1995 because it somehow supported our hypothesis. Our choice is justified by two of [Olson et al.] co-authors writing how monitoring substantially improved after 1995 [Wydeven et al. 2009]. The papers they cite [Stenglein et al. 2015, Stenglein and van Deelen 2016, van Deelen 2009] that begin analyses earlier do not seem to account for that change in census methods, which may affect their results." p. 2, (Chapron & Treves 2017b).

The preceding debate in 2017 reflected controversy about a series of papers published by the architects and their junior colleagues, which omitted the changes in census methods (sections B and C above). The first peer-reviewed population model introduced misleading information and apparent errors, when it stated,

> "The population grew slowly from 1980 to 1995 at which point the winter count surpassed **the endangered status of 80 wolves [sic, a]** (Wydeven et al. 2009). Since 1995, the wolf population increased dramatically, and management policy **changed with respect to the degree to which managers may kill wolves to address depredation problems [sic, b]**. Hence, policy changes and population growth interacted to define three recovery periods… **During 1996–2002, wolves were listed as endangered under the US Endangered Species Act [sic, c]** and protected from all hunting and trapping. In 2003, wolves were downgraded to threatened status and lethal control actions [followed]…. The period 2003–2012 was dominated by this on-again and off-again lethal control management..." (internal citations relating to lethal control omitted, emphasis added, p. 371, (Stenglein et al. 2015c).

I perceive errors or misleading text where I inserted **[sic a–c]** above: **(sic a)** Reclassification is a legal designation not a biological one and no change in state or federal policy was made before 2003 as summarized in this Appendix (nor was the claim of policy change substantiated (Stenglein et al. 2015c). Moreover, the 1995 spike in the population estimate that was associated with a change in census methods was detected in the winter of 1994-1995 before the observation that wolves had exceeded 80 individuals in April 1995. Therefore, any ostensible change in state or federal management policies (for which there is no record) follows the change in methods, not the other way around; **(sic b)** The history of lethal management policy is not accurately presented in the quotation above. Authority for killing wolves was not granted to the state of Wisconsin until 1 April 2003 (Chapron & Treves 2016a); **(sic c)** Similarly, wolves were federally listed as endangered since the late 1970s, so identifying a break in policy relating to lethal control in 1995 or 1996 appears misleading.

The errors or inaccuracies noted above were not insignificant given the population modeling used the three recovery periods as parameters, "…we fit a model with three correction factors that were constant within each recovery period (1980–1995, 1996–2002, and 2003–2011)." p.372, (Stenglein et al. 2015c). The periods they chose replicate the first errors in timing, which I detailed above. The first period should end April 1994 before volunteer trackers data led to a methodological artifact. The second period should end sometime between summer 2000 and winter 2003–2004 as I detailed above, when the WDNR first analyzed the substantial differences between volunteer trackers and WDNR staff and undertook changes to control the quality of data. Moreover, lethal methods were only permitted on 1 April 2003, which means 95.9% of that wolf-year (15 April 2002–14 April 2003) should be assigned to the prior policy period without lethal management (Chapron & Treves 2016a). Furthermore, the latter authors showed that lethal management changed 5 more times in the period under question, so I cannot understand why Stenglein et al. (2015c) did not define more policy periods. In sum, I perceive their policy periods obscured the real changes in methods and policy for killing wolves.

The problems with Stenglein et al. (2015c) did not end with their designation of illusory policy periods. Although I agree that adult survival did not seem to vary with density from 1995–2012 (Stenglein et al. 2015c; Chapron & Treves 2016a), I do not agree with their claim of negative density-dependence on recruitment (Stenglein et al. 2015c). Chapron & Treves (2017b) found problems with that claim, which seem to require reanalysis. Chapron and Treves (2017b) could not find Stenglein et al.'s (2015c) quantitative estimates for how recruitment changed from 1995–2012. Instead, the conclusion seems to be based on low-resolution, line graphics in Figures. Moreover, Stenglein et al. (2015c) found weak effects as shown here, "The evidence for a negative slope of the line for t > 18 was 69.0% (proportion of posterior that was <0)" (p. 372, Stenglein et al. 2015c). (Note that t > 18 refers to their policy periods found to be erroneous above.) and here, *"*48.4% of the time, the estimated population sizes in Wisconsin from 1981 to 2011 were within the 95% posterior intervals of μ_t *"* (p. 372, Stenglein et al. 2015c), implying that more than "half the time the estimates failed this relatively undemanding test" (p. 1, Chapron & Treves 2017b).

The new problem I report here is that the raw data on reproduction were not based on the few, rare, direct observation of pups around dens or pup survival to late fall, as done in other studies (Fuller 1989; Fuller et al. 2003). Rather, the estimates of recruitment relied on (a) indirect estimates for a minority of packs via summer howl surveys, which were recently shown to be widely variable between observers (Palacios et al. 2017); and (b) retrospective inference about pup production in a majority of packs, based on censuses taken 6–9 months later (Wydeven et al. 2004a; Wydeven et al. 2009). In the former 2004 article, we wrote, "The pup count… [is] based on a combination of direct and indirect evidence collected in both the summer and winter. As a result. pup count is statistically related to total pack size because DNR biologists estimated past pup production from current- and previous-year counts of adults and yearlings." p.35, (Wydeven et al. 2004a). In short, population growth necessarily reflected the observation of more wolves, which led to a circular inference that reproduction was responsible for packs that had grown in size. Migrants, and individuals undertaking long-distance movements, were not uncommon in this population (Treves et al. 2009b), which can lead to changes in pack size when census is based on brief encounters every few weeks (e.g., by volunteer trackers conducting snow surveys or pilots counting packs from the air). By contrast, Chapron & Treves (Chapron & Treves 2016a) treated reproduction as a binary variable (reproduction or no reproduction), which is more robust to uncertainty about the number of pups (Palacios et al. 2017), compared to a continuous measure of pup survival. Therefore, the evidence for density-dependence on reproduction seems so weak as to require new data not skimpy reanalysis.

In 2016, Stenglein and Van Deelen attributed population dynamics to biological mechanisms explicitly, although they addressed (and dismissed) methodological artifacts,

> "We did not find reduced fecundity in pups per pack or in the proportion of breeding females in the population pre-1995 compared to 1995–2007 (Stenglein unpublished). However, the proportion of lone wolves prior to 1995 (roughly 10% of the population) was higher compared to 1995–2007 when only 4% were

> lone wolves (Wydeven et al. 2009). **The difference in proportion of lone wolves could be due to sampling and detection issues;** however a real difference provides support for a mate-finding component Allee effect in early recovery because it suggests that wolves had difficulty finding mates at low densities, resulting in more lone wolves… When the wolf population was small, it may have been more difficult to count wolves, packs and occupied territory. A**t small population sizes, failing to count just one pack and then finding and counting it in the next year could lead to the appearance of substantial population growth which would be due to observer error rather than real growth**." (emphasis added, p. 11-12, (Stenglein & Van Deelen 2016).

All the most recent peer-reviewed articles of which I am aware ignored the debate about census methods entirely and continued to attribute biological causes to the methodological artifact that Treves & Chapron (2016a, 2017b) had made them aware.

The most recent by Stenglein, Wydeven and Van Deelen (2018) described their methods as follows, "The dataset consisted of > 42,000 weekly locations of 501 wolves captured and tracked (November 1979–December 2013)… [documenting] (2) known collar failure or lost to follow-up during the study and then found dead sometime after its endpoint (known censoring)." p.102, (Stenglein et al. 2018). Finding dead wolves whose radio-collars had stopped transmitting or were otherwise 'lost to follow up' requires observers on the ground, yet the methods and changes in those methods over time for volunteer tracking were not presented. See Box 1 main text for the large number of wolves lost to follow up. The same criticisms apply to (Olson et al. 2015; Olson et al. 2017). And because the latter article followed the debate with the architects and close collaborators (Olson et al. 2015; Olson et al. 2017) and similar debates about detecting dead wolves (Treves et al. 2017a; Treves et al. 2017c), the latest publication in 2018 seems to perpetuate misleading methods and analyses (Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018)(Stenglein et al. 2018).

In sum, articles from 2015–2018 modeled wolf population dynamics, birth, or deaths with essential methods having been omitted (Stenglein 2014; Olson et al. 2015; Stenglein et al. 2015a; Stenglein et al. 2015b; Stenglein & Van Deelen 2016; Olson et al. 2017; Stenglein et al. 2018). When specifying methods, they cited only these articles (Wydeven et al. 1995; Van Deelen 2009; Wydeven et al. 2009), which treated the wolf population estimates as a single time series (erroneously in my view), treated the time series as an outcome of density-dependent dynamics, contrary to warnings from scientific reviews of the topic (Brook & Bradshaw 2006), and treated the time series as the statistically independent from estimates of reproduction and mortality despite their own acknowledgments of the shortcomings of monitoring methods. Therefore, I do not consider the work of the architects of the state wolf census and population model to be accurate reflections of the Wisconsin wolf population status or dynamics, nor of the effects of lethal management on illegal killing or population dynamics (Box 1).

E. After delisting in 2012

WDNR plans to allow extraction of wolves in a regulated season with hunting, trapping, and hounding went into effect on 15 October 2012 (Natural Resources Board 2012; Treves et al. 2017b). The next year, the Secretary of the WDNR reaffirmed N_{goal} and the pivotal role of one of the state architects and their junior colleagues,

> "…several scientists, including Timothy Van Deelen and his research associate Jennifer Stenglein, have presented to the [WDNR] wolf committee on multiple occasions.[The newspaper article] suggests that because Van Deelen isn't a member of the committee that his inputs are not considered. To the contrary, their model has been utilized for two consecutive years to help model and project how the state's wolf population will respond to harvest.… Next the assertion that we will get to our goal of 350 wolves in one year ignores the science. The UW's own population model indicates this year's harvest could result in a 13 percent reduction in the state's wolf population.… we have been clear that we will honor the established population goal…." no pagination in original, (Stepp 2013).

The above quote supports my contention that the 1999 Plan and 2007 addendum continued to model wolf populations with an implicit assumption about public hunting of wolves. However, state implementation and regulations for killing wolves did not match the model assumptions as far as I can tell, for the following reasons.

Harvest models based on maximum sustainable yields assume that 'surplus' individuals are 'harvested' or killed such that future reproduction is unimpeded. Wisconsin's eventual wolf hunting regulations made no distinction between killing breeders, supernumerary adult helpers, or young (2012 Wisconsin Act 169) and (Natural Resources Board 2012, 2013, 2014). Evidence suggests those regulations would not protect the reproductive capacity of the wolf population and left it to stochasticity. The total number of breeding females legally killed in Wisconsin in 2012 has not been published. But neighboring Minnesota held a similar public hunting season in 2012 and presented the following estimates: 51% of wolves killed were females, of which 22% appeared to have once been breeders (Figure 9b, Stark & Erb 2013). Without considering the effect of deaths of breeding males on pack reproduction (Brainerd et al. 2008; Borg et al. 2015), and assuming one breeding female per pack, I use the Minnesota estimate to predict wolf reproductive potential diminished by 22% in the late winter after Wisconsin legalized public hunting and trapping for wolves. Diminished reproductive potential following loss of breeders can persist for more than one breeding season (Brainerd et al. 2008; Borg et al. 2015). Also, the 22% estimate from Minnesota did not consider hounding that was not allowed in Minnesota. One might also expect the 22% might be additive to other causes of breeder loss. Without better data on reproduction, the effects of Wisconsin's wolf-hunting seasons on population size and growth are uncertain.

At no time did the scientists assisting the state explain to the public whether or how the risk of population crash had changed (Natural Resources Board 2012, 2013, 2014).

I anticipate one rebuttal that Wisconsin's wolf population is thriving at an estimated census of 866 in April 2016 (USFWS 2018). But questions remain about the censuses from 2013–2016, especially after the change in census methods after winter 2012–2013. In July 2012, 28 veteran, volunteer census-takers opposed wolf policy publicly and many later resigned (Ericksen-Pilch et al. 2012); they represented approximately 42% of the previous year's number of volunteers (Wydeven et al. 2012). In response to a need for wolf monitors, the state seems to have relaxed the criteria in period 3 (dashed line in Figure A2) for effort, quality, and training. Then, the state reduced the former transparency of the tally of the wolf census, "That the DNR sort of has to come up with a count **in less open system** because giving the exact location of every wolf and every pack is occurring in the state is no longer appropriate when they're a hunted species." (emphasis added, transcribed from radio interview at 38:08 min:sec, Wydeven 2016). I do not understand the logic behind that statement given that poaching has long been the major cause of wolf mortality when the more transparent census was conducted (Treves et al. 2017a; Treves et al. 2017c) and the state revealed all wolf pack locations after 2012, with high spatial resolution (Wisconsin Department of Natural Resources 2018). The census policy and reporting would seem to obscure the wolf population estimate while making wolf packs easier not harder to locate.

F. Conclusion of Appendix

The original scientific justifications for Wisconsin's wolf population model (K, density-dependence, a single time series for census estimates) have been discredited or questioned with equivalent or stronger evidence. To uphold scientific integrity and identify the best available scientific and commercial data, we need an open and thorough peer-reviewed test between alternative hypotheses for the pattern of wolf population growth in Wisconsin. Without that, the state's wolf policy should not be considered the best available science.

Gaps in scientific integrity relating to objectivity (section A), transparency (Sections A-C), and replicability of results (Sections C and D) are all considered problematic by the National Academies of Science (National Academy of Sciences et al. 1992). Currently, the primary defense – against gaps in scientific integrity and wore transgressions of scientific ethics by a government agency – is academic freedom fortified by tenure or other protections for the independence of scientists. The U.S. Fish & Wildlife Service can and should contribute to finding and promoting the best available science by holding lower jurisdictions to higher standards of scientific integrity and academic freedom.

End of Appendix

References Cited

Adams LG, Stephenson RO, Dale BW, Ahgook RT, Demma DJ. 2008. Population dynamics and harvest characteristics of wolves in the Central Brooks Range, Alaska Wildlife Monographs **170**:1-25.

Beyer DE, R.O. Peterson, J.A. Vucetich, Hammill. JH. 2009. Wolf population changes in Michigan. Pages 65-85 in Wydeven AP, Van Deelen TR, and Heske EJ, editors. Recovery of Gray Wolves in the Great Lakes Region of the United States: an Endangered Species Success Story. Springer, New York.

Borg BL, Brainerd SM, Meier TJ, Prugh LR. 2015. Impacts of breeder loss on social structure, reproduction and population growth in a social canid. Journal of Animal Ecology **84**:177-187.

Brainerd SM, et al. 2008. The effects of breeder loss on wolves. Journal of Wildlife Management **72**:89-98.

Brook B, Bradshaw C. 2006. Strength of evidence for density dependence in abundance time series of 1198 species. Ecology **87**:1445-1451.

Browne-Nuñez C, Treves A, Macfarland D, Voyles Z, Turng C. 2015. Tolerance of wolves in Wisconsin: A mixed-methods examination of policy effects on attitudes and behavioral inclinations. Biological Conservation **189**:59–71.

Bruskotter JT, Vucetich JA, Enzler S, Treves A, Nelson MP. 2013. Removing protections for wolves and the future of the U.S. Endangered Species Act (1973) Conservation Letters **7**:401-407.

Chapron G, Treves A. 2016a. Blood does not buy goodwill: allowing culling increases poaching of a large carnivore. Proceedings of the Royal Society B **283**:20152939.

Chapron G, Treves A. 2016b. Correction to 'Blood does not buy goodwill: allowing culling increases poaching of a large carnivore'. Proceedings of the Royal Society B **Volume 283**:20162577.

Chapron G, Treves A. 2017a. Reply to comment by Pepin et al. 2017. Proceedings of the Royal Society B **2016257**:20162571.

Chapron G, Treves A. 2017b. Reply to comments by Olson et al. 2017 and Stien 2017. Proceedings of the Royal Society B **284**:20171743.

Creel S, et al. 2015. Questionable policy for large carnivore hunting. Science **350**:1473-1475.

Epstein Y. 2017. Killing Wolves to Save Them? Legal Responses to 'Tolerance Hunting' in the European Union and United States. RECIEL **26**:19-29.

Epstein Y, Chapron G. 2018. The Hunting of Strictly Protected Species: The Tapiola Case and the Limits of Derogation under Article 16 of the Habitats Directive. European Energy and Environmental Law Review **June**:78-87.

Erb J, Benson S. 2013. Distribution and abundance of wolves in Minnesota, 2012–2013. Grand Rapids, MN.

Ericksen-Pilch M, et al. 2012. The Coalition fo Wisconsin Wolf Trackers (CWWT): Testimony by 28 volunteer wolf census-takers declaring opposition to state wolf policy. Pages 1-9.

Fowler CW. 1987. A Review of Density Dependence in Populations of Large Mammals. Pages 401-441 in H.H. G, editor. Current Mammalogy. Springer, Boston.

Fuller TK. 1989. Population dynamics of wolves in north central Minnesota. Wildlife Monographs **105**:3-41.

Fuller TK, Mech LD, Cochrane JF. 2003. Wolf population dynamics. Pages 161-191 in Mech LD, and Boitani L, editors. Wolves: Behavior, ecology, and conservation. University of Chicago Press, Chicago.

Fuller TK, W. E. Berg, G. L. Radde, M. S. Lenarz, Joselyn GB. 1992. A history and current estimate of wolf distribution and numbers in Minnesota. Wildlife Society Bulletin **20**:42-55.

Groskopf L. 2014. Prototype wolf control resolution for county boards or other decision making bodies or clubs. 7-29-14. Pages pp 39-41 in the pdf but labeled P33-P35 in Adams

County Board of Supervisors Meeting Report (Amended) November 18, editor. Adams County Board, WI, Adams County Board Room.

Gude JA, Mitchell MS, Russell RE, Sime CA, Bangs EE, Mech LD, Ream RR. 2012. Wolf population dynamics in the U.S. Northern Rocky Mountains are affected by recruitment and human-caused mortality. Journal of Wildlife Management **76**:108-118.

Hogberg J, Treves A, Shaw B, Naughton-Treves L. 2015. Changes in attitudes toward wolves before and after an inaugural public hunting and trapping season: early evidence from Wisconsin's wolf range. Environmental Conservation **43**:45-55.

Ioannidis JP. 2005. Why Most Published Research Findings Are False. PLOS Medicine **2**:e124.

Liberg O, Chapron G, Wabakken P, Pedersen HC, Hobbs NT, Sand Hk. 2012. Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe. Proceedings of the Royal Society of London Series B **270**:91-98.

Lynn WS. 2018. Bringing Ethics to Wild Lives: Shaping Public Policy for Barred and Northern Spotted Owls. society & animals **26**:217-238.

Mladenoff DJ, Haight RG, Sickley TA, Wydeven AP. 1997. Causes and implications of species restoration in altered ecosystems. BioScience **47**:21-31.

Mladenoff DJ, Sickley TA, Haight RG, Wydeven AP. 1995. A regional landscape analysis and prediction of favorable gray wolf habitat in the northern Great Lakes region. Conservation Biology **9**:279-294.

Moehlman P. 1989. Intraspecific variation in canid social systems. Pages 143-163 in Gittleman JL, editor. Carnivore Behavior, Ecology and Evolution. Comstock Associates, Ithaca.

National Academy of Sciences, National Academy of Engineering, Institute of Medicine. 1992. Responsible Science: Ensuring the Integrity of the Research Process. National Academy Press, Washington, DC.

Natural Resources Board. 2012. Adoption of Board Order WM-09012(E) relating to wolf hunting and trapping regulations, establishment of a depredation program, and approval of a harvest quota and permit level. Madison, WI.

Natural Resources Board. 2013. Request approval of a wolf harvest quota and number of licenses to issue for the 2013-2014 wolf hunting and trapping season. Madison, Wisconsin

Natural Resources Board. 2014. Request approval of a wolf harvest quota and number of licenses to issue for the 2014-2015 wolf hunting and trapping season. Madison, Wisconsin

Naughton-Treves L, Grossberg R, Treves A. 2003. Paying for tolerance: The impact of livestock depredation and compensation payments on rural citizens' attitudes toward wolves. Conservation Biology **17**:1500-1511.

Nowak S, Mysłajek RW. 2016. Wolf recovery and population dynamics in Western Poland, 2001–2012. Mammal Research **61**:83-98.

Olson ER, Crimmins S, Beyer DE, MacNulty D, Patterson B, Rudolph B, Wydeven A, Van Deelen TR. 2017. Flawed analysis and unconvincing interpretation: a comment on Chapron and Treves 2016. Proceedings of the Royal Society of London B **284**:20170273.

Olson ER, Stenglein JL, Victoria Shelley, Adena R. Rissman, Christine Browne-Nuñez, Zachary Voyles, Wydeven AP, Deelen TV. 2015. Pendulum swings in wolf management led to conflict, illegal kills, and a legislated wolf hunt. Conservation Letters **8**:351-360.

Palacios V, Font E, García EJ, Svensson L, Llaneza L, Frank J, López-Bao JV. 2017. Reliability of human estimates of the presence of pups and the number of wolves vocalizing in chorus howls: implications for decision-making processes. European Journal of Wildlife Research **63**:59-66.

Refsnider R. 2009. The Role of the Endangered Species Act in Midwest Wolf Recovery. Pages 311-330 in Wydeven AP, Van Deelen TR, and Heske EJ, editors. Recovery of gray wolves in the Great Lakes region of the United States: An endangered species success story. Springer, New York.

Schanning K. 2009. Human dimensions: public opinion research concerning wolves in the Great Lakes states of Michigan, Minnesota, and Wisconsin. Pages 251-266 in Wydeven AP, Van Deelen TR, and Haske EJ, editors. Recovery of Gray Wolves in the Great Lakes Region of the United States: An Endangered Species Success Story. Springer, New York.

Schmidt JH, Johnson DS, Lindberg MS, Adams LG. 2015. Estimating demographic parameters using a combination of known-fate and open N-mixture models. Ecology **56**:2583–2589.

Smith DW, et al. 2010. Survival of colonizing wolves in the Northern Rocky Mountains of the United States, 1982 – 2004. Journal of Wildlife Management **74**:620 – 634.

Stark D, Erb J. 2013. 2012 Minnesota wolf season report. Grand Rapids, MN.

Stenglein J. 2014. Survival of Wisconsin's gray wolves from endangered to harvested, 1980 – 2013. University of Wisconsin-Madison.

Stenglein JL, Gilbert JH, Wydeven AP, Van Deelen TR. 2015a. An individual-based model for southern Lake Superior wolves: A tool to explore the effect of human-caused mortality on a landscape of risk. Ecological Modelling **302**:13-24.

Stenglein JL, Van Deelen TR. 2016. Demographic and Component Allee Effects in Southern Lake Superior Gray Wolves. PLOS ONE **11**:10.1371/journal.pone.0150535

Stenglein JL, Van Deelen TR, Wydeven AP, Mladenoff DJ, Wiedenhoeft J, Langenberg JA, Thomas NJ. 2015b. Mortality patterns and detection bias from carcass data: An example from wolf recovery in Wisconsin. Journal of Wildlife Management **7**:1173-1184.

Stenglein JL, Wydeven AP, Van Deelen TR. 2018. Compensatory mortality in a recovering top carnivore: wolves in Wisconsin, USA (1979–2013). Oecologia **187**:99–111.

Stenglein JL, Zhu J, Clayton MK, Van Deelen TR. 2015c. Are the numbers adding up? Exploiting discrepancies among complementary population models. Ecology and Evolution **5**:368-376.

Stepp C. 2013. DNR's Cathy Stepp: Wolf management is an art as well as a science. The Cap Times, Madison, WI.

Thiel RP 1993. The Timber Wolf in Wisconsin: The Death and Life of a Majestic Predator. University of Wisconsin Press, Madison, WI.

Treves A. 2008. Beyond recovery: Wisconsin's wolf policy 1980-2008. Human Dimensions of Wildlife **13**:329-338.

Treves A, Artelle KA, Darimont CT, Parsons DR. 2017a. Mismeasured mortality: correcting estimates of wolf poaching in the United States. Journal of Mammalogy **98**:1256–1264.

Treves A, Artelle KA, Paquet PC. 2018. Differentiating between regulation and hunting as conservation interventions. Conservation Biology **33**:472–475.

Treves A, Bruskotter JT. 2014. Tolerance for predatory wildlife. Science **344**:476-477.

Treves A, Chapron G, López-Bao JV, Shoemaker C, Goeckner A, Bruskotter JT. 2017b. Predators and the public trust. Biological Reviews **92**:248-270.

Treves A, Jurewicz RL, Naughton-Treves L, Wilcove D. 2009a. The price of tolerance: Wolf damage payments after recovery. Biodiversity and Conservation **18**:4003–4021.

Treves A, Krofel M, McManus Jec-a. 2016. Predator control should not be a shot in the dark. Frontiers in Ecology and the Environment **14**:380-388.

Treves A, Langenberg JA, López-Bao JV, Rabenhorst MF. 2017c. Gray wolf mortality patterns in Wisconsin from 1979 to 2012. Journal of Mammalogy **98**:17-32.

Treves A, Martin KA. 2011. Hunters as stewards of wolves in Wisconsin and the Northern Rocky Mountains, USA. Society and Natural Resources **24**:984-994.

Treves A, Martin KA, Wiedenhoeft JE, Wydeven AP. 2009b. Dispersal of gray wolves in the Great Lakes region. Pages 191-204 in Wydeven AP, Van Deelen TR, and Heske EJ, editors. Recovery of Gray Wolves in the Great Lakes Region of the United States: An Endangered Species Success Story. Springer, New York.

Treves A, Naughton-Treves L. 2005. Evaluating lethal control in the management of human-wildlife conflict. Pages 86-106 in Woodroffe R, Thirgood S, and Rabinowitz A, editors. People and Wildlife, Conflict or Coexistence? Cambridge University Press, Cambridge, UK.

Treves A, Naughton-Treves L, Shelley VS. 2013. Longitudinal analysis of attitudes toward wolves. Conservation Biology **27**:315–323.

Troxell PS, Berg KA, Jaycox H, Strauss AL, Struhsacker P, Callahan P. 2009. Education and Outreach Efforts in Support of Wolf Conservation in the Great Lakes Region. Pages 297-309 in Wydeven AP, Van Deelen TR, and Heske EJ, editors. Recovery of Gray Wolves in the Great Lakes Region of the United States: an Endangered Species Success Story. Springer, New York.

USFWS USFWS. 2018. Wolf Numbers in Minnesota, Wisconsin and Michigan (excluding Isle Royale) - 1976 to 2016, Available from http://www.fws.gov/midwest/wolf/aboutwolves/mi_wi_nos.htm (accessed 15 November 2018.

Van Deelen TR. 2009. Growth Rate and Equilibrium Size of a Recolonizing Wolf Population in the Southern Lake Superior Region. Pages 139-154 in Wydeven AP, Van Deelen TR, and Heske EJ, editors. Recovery of Gray Wolves in the Great Lakes Region of the United States: an Endangered Species Success Story. Springer, New York.

van Eeden LM, et al. 2018. Carnivore conservation needs evidence-based livestock protection. PLOS Biology **https://doi.org/10.1371/journal.pbio.2005577**.

Vucetich JA. 2012. Appendix: The influence of anthropogenic mortality on wolf population dynamics with special reference to Creel and Rotella (2010) and Gude et al. (2011) in the Final peer review of four documents amending and clarifying the Wyoming gray wolf management plan. Congressional Federal Register **50**:78-95.

WDNR 1999. Wisconsin Wolf Management Plan. Wisconsin Department of Natural Resources, Madison, WI.

WDNR. 2007. Wisconsin Wolf Management Plan Addendum 2006 and 2007. Madison.

Wiedenhoeft JE, Boles SR, Wydeven AP. 2003. Counting wolves--integrating data from volunteers. Page abstract only. World Wolf Congress 2003: Bridging Science and Community, Banff, Alberta, Canada.

Wisconsin Department of Natural Resources. 2018. Wisconsin gray wolf monitoring report 15 April 2015 through 14 April 2016 in Resources DoN, editor, Madison, WI.

Wydeven AP. 1994. Wisconsin Endangered Resources Report #102: Status of the timber wolf in Wisconsin, performance report 1 July 1993 through 30 June 1994 in Resources BoE, editor. Wisconisin Department of Natural Resources, Madison, WI.

Wydeven AP. 2016. Joy Cardin show in Cardin J, editor, Wisconsin Puboic Radio.

Wydeven AP, Megown RA. 1995. Wisconsin Endangered Resources Report #104: Status of the timer wolf in Wisconsin, performance report 1 July 1994 through 30 June 1995 in Resources BoE, editor. Wisocnisin Department of Natural Resources, Madison, WI.

Wydeven AP, Megown RA. 1996. Wisconsin Endangered Resources Report #111: Status of the timer wolf in Wisconsin, performance report 1 July 1995 through 30 June 1996 in Resources BoE, editor. Wisocnisin Department of Natural Resources, Madison, WI.

Wydeven AP, Schultz RN, Thiel RP. 1995. Monitoring a recovering gray wolf population in Wisconsin, 1979-1995. Pages 147-156 in Carbyn LN, Fritts SH, and Seip DR, editors. Ecology and Conservation of Wolves in a Changing World. Canadian Circumpolar Institute, Edmonton, Alberta, Canada.

Wydeven AP, Treves A, Brost B, Wiedenhoeft JE. 2004a. Characteristics of wolf packs in Wisconsin: Identification of traits influencing depredation. Pages 28-50 in Fascione N, Delach A, and Smith ME, editors. People and Predators: From Conflict to Coexistence. Island Press, Washington, D. C.

Wydeven AP, Wiedenhoeft J, Schultz RN, Thiel RP, Jurewicz RR, Kohn B, Van Deelen TR. 2009. History, population growth and management of wolves in Wisconsin. Pages 87-106 in Wydeven AP, Van Deelen TR, and Heske EJ, editors. Recovery of Gray Wolves in the Great Lakes Region of the United States: an Endangered Species Success Story. Springer, New York.

Wydeven AP, Wiedenhoeft JE, Schultz RL, Thiel RP, Boles SH, Heilhecker E. 2006. Progress Report of Wolf Population Monitoring in Wisconsin for the Period October 2005 - March 2006 in Resources WDoN, editor. Wisconsin Department of Natural Resources, Park Falls, Wisconsin PUB-ER- 2006.

Wydeven AP, Wiedenhoeft JE, Schultz RL, Thiel RP, Boles SH, Heilhecker E, Hall WHJ. 2004b. Progress Report of Wolf Population Monitoring in Wisconsin for the Period October - March 2004 in Resources WDoN, editor. Wisconsin Department of Natural Resources, Park Falls, Wisconsin PUB-ER- 2004.

Wydeven AP, Wiedenhoeft JE, Schultz RN, Bruner JE, Boles SR. 2012. Status of the timber wolf in Wisconsin: Performance report 1 July 2011 through 30 June 2012 (also PROGRESS REPORTS FOR 15 ARPIL 2011-14 APRIL 2012, and 2011 summaries). Resources WDoN, Madison, WI.

Text of the peer review ends here