Guidance for scientific assessment

We (WDFW) will evaluate whether our agency's recent vessel regulations benefit Southern Resident killer whales (SRKW). This question is challenging to answer, because effects of the regulations on SRKW well-being can be extremely difficult to isolate. Therefore, we look to associated – or proxy – effects that can be isolated and measured, acknowledging simplifying assumptions. In evaluating recent small vessel rules implemented by WDFW to benefit SRKW, we seek to answer questions such as 1) Does the 2019 reduced speed regulation benefit orcas?; 2) Does the 2019 increased buffer distance regulation benefit orcas?; 3) Do the 2021 commercial whale-watching licensing program rules benefit orcas? Recognizing the difficulty of addressing these directly, we instead might seek to answer more measurable questions such as:

- Has disturbance of SRKW (e.g., noise) decreased with implementation of the rules?
- Has the rate of vessel speed infractions decreased with implementation of the rules?
- Has the distance of regulated and recreational vessels from SRKW increased with implementation of the rules?
- Has the number of vessels within a half nautical mile of SRKWs decreased with implementation of the rules?
- Are there unintended consequences of the rules?

The strongest evidence is derived from studies designed, analyzed and interpreted to answer questions directly addressing questions about the effectiveness of the rules. However, these studies are difficult to design and conduct. In the interest of evaluating and incorporating all useful information, we have developed the following guidance for considering pieces of evidence proposed to measure the effectiveness of the commercial whale-watching (CWW) rules, based on the Weight of Evidence (WoE) approach. WoE is a process by which evidence is evaluated and synthesized to determine the relative support for possible answers to a question.

Characteristics of useful scientific information:

Signal strength – the magnitude of the measured effect relative to the uncertainty.

- Low = effect is small relative to the uncertainty associated with the measurement
- Medium = effect is moderate, uncertainty is moderate
- High = effect is large, uncertainty is relatively small

Methodology – the format, design, and process of data collection.

- Low = anecdotal evidence and/or line of inquiry has only single potential outcome
- Medium = data collected without design, i.e., opportunistically
- High = data collection designed to isolate the variable(s) of interest for appropriate analysis; line of inquiry has multiple potential outcomes

Analysis – level of statistical evaluation of data.

- Low = product is raw data, no analysis or measure of uncertainty; inference exceeds the limits of the analysis

- Medium = descriptive statistics (e.g., mean with variability); post-hoc analysis of data collected with limited/no study design
- High = inference is limited to what is supported by and consistent with the analysis; product is estimate of effect with uncertainty

Relevance – measure of association between effect of interest/concern and line of evidence provided.

- Low = variable measured is many/uncertain steps from, or has undefined connection to, the effect of interest
- Medium = variable can be connected to effect of interest through a multi-step chain; logical connection is clear but contains uncertainty
- High = variable of interest is measured

Repeatability/replicability of study

- Low = methods and data collection are described/conducted in such a way that they could not be reproduced; consistency among data collectors is not tested or is very low
- Medium = methods and data collection are described/conducted such that they could be partially reproduced
- High = methods and data collection are described/conducted in such a way that they would produce the same results in consecutive trials; consistency among data collectors is assessed and found to be high

Scientific context

- Low = lacks grounding in contemporary scientific context, particularly with respect to underlying mechanisms and relationships; no external scientific review
- Medium = some consideration of scientific context
- High = grounded in contemporary scientific context; peer review by independent scientists

Bias:

Bias can occur in the planning, data collection, analysis and interpretation phases. Bias is often present in degrees (not present or absent), so a) data collectors must attempt to eliminate bias, and b) evaluators must attempt to identify the types of bias and whether/how that bias may influence the data or conclusions. Controls to minimize bias increase the utility of the findings.

Potential types of bias:

Selection bias – occurs when the data are selected subjectively, i.e., non-random and non-standardized sampling. This renders the data unrepresentative of the system.

- Low bias- sampling designed using a random process and strictly followed
- High bias- Sampling is determined opportunistically and/or by interpersonal connections

Confirmation bias – occurs when the data collector or analyzer expects a certain result and intentionally or unintentionally seeks to generate it. This includes scenarios where the collector or supporter may benefit financially or otherwise from supporting a certain conclusion.

- Low bias- hypothesis has a chance of being unsupported

- High bias- the study design has very low/no potential to disagree with the hypothesis, contradictory data are not collected

Observer bias – occurs when the data collector or data collection process influences the variable to be measured. This can be conscious or unconscious.

- Low bias- data collection process randomized and/or standardized to minimize effect on product; data collection is unknown to subject
- High bias- product clearly impacted by data collection process; protocol for data collection is not documented ahead of time; is only "known" by the researcher

Confounding factors – occurs when a variable of interest is not adequately isolated and data/outcomes are co-influenced by multiple factors.

- Low bias- relationship of interest has been successfully isolated through study design; assumptions addressed
- High bias- relationship of interest is obscured or inflated due to other factors; assumptions not addressed

Assumptions

Assumptions are a critical part of scientific inquiry. Identifying the information that is assumed to be true simplifies analysis and allows conclusions to be drawn and understood. It is important to think critically about what has been assumed in the generation and interpretation of data. Assumptions should be explicitly stated and justified, so that the audience can understand how conclusions are drawn.

Relevant pieces of evidence must address the following to be considered:

- Clearly define the question intended to be answered by the evidence; be as specific as possible.
- Please describe data collection, as well as any study design and/or analysis, in detail.
- How relevant is the metric to one of the effectiveness questions (see above)? Describe the logical steps and established relationships associating the evidence to the question(s) of effectiveness.
- Please acknowledge/address any assumptions.
- What are potential sources of bias in the data collection? How have these been mitigated/addressed?
- Is there any other information WDFW should consider regarding the usefulness of this evidence?

What to expect:

WDFW will evaluate submitted evidence according to the above considerations. We may follow up with additional questions/clarifications in order to fully evaluate each piece of evidence. The evidence will be weighed and considered in the context of the remaining evidence. Information that does not satisfy the above characteristics of rigorous scientific inquiry or does not address the above points may not be included as scientific evidence.

We acknowledge that not all of the variables of interest can be isolated, and where necessary we will fall back on the summary of science produced by the Washington State Academy of Sciences during the development of the rules. Additionally, we will continue to be guided by the precautionary principle in instances of substantive uncertainty.