SC/68C/SM/21

Sub-committees/working group name: SM

Report on using expert elicitation to estimate total unique vaquitas and calves in the Zero Tolerance Area with recommendations for future research efforts

Lorenzo Rojas-Bracho, Cormac Booth, Barbara Taylor, Len Thomas, Tim Gerrodette, Annette Henry

Report on using expert elicitation to estimate total unique vaquitas and calves in the Zero Tolerance Area with recommendations for future research efforts.

Lorenzo Rojas-Bracho[1], Cormac Booth[2], Barbara Taylor[3], Len Thomas[4], Tim Gerrodette[3], Annette Henry[3]
1: Comisión Natural de Áreas Naturales Protegidas; Ensenada, B.C. México; 2: SMRU Consulting, St Andrews, Scotland; 3: Southwest Fisheries Science Center, La Jolla, CA USA; 4. Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, Scotland

1. Executive Summary

Between September 2 and October 27, 2019 surveys were conducted from two ships (the *Narval* and either the *Farley Mowat* or the *Sharpie* provided by Operation Milagro of the Sea Shepherd Conservation Society). Surveys stayed within the Zero Tolerance Area (ZTA), a rectangle of about 288 km^2 or roughly 12 by 24 km (which is a bit smaller than the urban area of Tijuana). Passive acoustic data continue to indicate that the ZTA is the area used as primary habitat by vaquitas. Experienced observers were used to find and track vaquitas with methods developed for the earlier effort to capture vaquitas in 2017. 7 sightings were made on 4 days when winds were low enough to sight and track vaquitas, and only 2 photographs were matched to previous efforts. There were no photographic matches within the 3 weeks. All but one sighting contained a smaller individual that could have been a calf.

There were insufficient photographic matches to make a mark-recapture estimate, so expert elicitation was used to estimate the number of unique calves and unique vaquitas (including calves). This report summarizes results from 3 different expert elicitation methods: Structured Expert Decision Making (SEDM), Expert Elicitation using the Rational Impartial Observer method (EE-RIO) and Expert Elicitation using the Roulette method (EE-Roulette). The second two methods were used in a separate exercise conducted Aug 31 through Sept 3, 2020. The exercise was designed to not only potentially improve 2019 estimates but also train observers for future efforts where the expectation is that expert elicitation will again be needed to interpret sightings made over several weeks.

For SEDM, observers were independently sent by email a number of questions with the goal of estimating both the range and most likely values for the unique number of calves and the number of total unique individuals. Unlike the accepted method of SEDM, there was no discussion of the results followed by an opportunity for experts to change their estimates. Experts were given an evidence dossier summarizing the sightings that was later used for the EE exercise. The mean estimate for number of calves seen was 3 with a 63% belief that there were at least 3 calves. The mean estimate for the number of unique vaquitas seen in all 7 sightings was 9.4 with 40% belief that there were at least 10.

The EE-RIO exercise required all experts to take 3 hours of online training and read the evidence dossier covering the 7 vaquita sightings. Elicitation was facilitated over a 3-day period using video-conferencing for a total of 7 hours. Experts independently input their lowest and highest plausible values for number of unique calves and unique total vaquitas together with their 25^{th}, 50^{th} and 75^{th} percentiles. These values were used to generate a distribution for each expert and each expert was asked to verbally give their rationales. Discussion revealed that there was additional photographic evidence that could influence estimates and that evidence was added to the dossier. Thus, the EE exercise utilized more evidence than the SEDM exercise. Facilitators then led a discussion towards a consensus distribution that the group believed could be convincingly argued to a Rational Impartial Observer. Experts agreed that a RIO would estimate that 3 was the most likely integer value for the true number of calves, with

approximately equal belief that the true value lied between 1 and 3 or between 3 and 5 calves (Figure 4.2.1). There was a 97.4% chance of 2 or more calves and a 71.4% belief that the true number of calves sighted was 3 or more. Experts agreed that a RIO would judge, based on the discussions, that it was implausible that the true number of vaquita sighted was less than 7 or more than 15, and that the most likely value was 11 with values close to 11 being almost as likely. The RIO distribution indicates a 71.8% belief that the true number of unique vaquita sighted was 10 or more.

Following the video conference, an email was sent to experts with a spreadsheet. Experts were requested to ignore the RIO results and cast 'votes' for different discrete numbers of potential unique calves and unique total vaquitas. Each expert was given the same weight and values were summed to give results for EE-roulette. This is similar to the SEDM, but benefited from having the same evidence as the EE-RIO and allowing experts to participate in and hear the discussion about the evidence. The mean estimated number of calves seen was 3.1 with a 73% belief that there were at least 3 calves. The mean estimated number of unique vaquitas seen in all 7 sightings was 10.4 with 66% belief that there were at least 10.

2. Introduction

2.1. Summary of rationale for developing visual/acoustic survey method.

The decline in vaquita numbers has been well documented. The first effort to cover the full vaquita distribution used visual line-transect methods (Jaramillo-Legorreta et al. 1999). This effort noted the difficulty in sighting this species because of small group size, inconspicuous surfacings and avoidance of the survey vessels. Imprecise abundance estimates raised concerns about timely detection of potential declines in abundance (Taylor and Gerrodette, 1997). Acoustic monitoring methods were developed to increase precision of estimating both abundance and trends in abundance (Jaramillo-Legorreta et al. 2017), and a combination of visual and acoustic methods were used to estimate vaquita abundance in 2008 (Gerrodette et al., 2011) and 2015 (Taylor et al., 2016). Acoustic monitoring indicated that the vaquita population continued to decline rapidly, about 50%/year, through 2018 (Thomas et al. 2019, Jaramillo-Legorreta, et al. 2019).

Recent developments, however, have made both acoustic monitoring and visual line-transect methods difficult. Illegal fishermen have begun removing the acoustic devices (CPODs) used to record vaquita clicks. The data recorded on each device is lost, and it is expensive to replace the stolen CPODs. Unless enforcement of the fishing ban is effective and the theft of equipment is stopped, acoustic monitoring of the vaquita population using past methods is not feasible.

Visual line-transect methods face a different problem. The number of vaquitas is now so low that the number of sightings is not sufficient to estimate the parameters necessary to estimate abundance. If a line-transect survey was carried out utilizing the same ship as in past surveys (the *David Starr Jordan/Ocean Starr*), an estimate of abundance would be possible with relatively few sightings, because the probability of detection has been estimated for this ship. However, chartering this vessel and hiring experienced observers for the necessary time would be expensive, at least US$3,000,000 for a survey. Unless such funds are available, the size of the 2020 vaquita population cannot be estimated using line-transect methods. Therefore, CIRVA, the international vaquita recovery team, recommended using photo-identification methods to produce an estimate of minimum abundance, and to refine our understanding of aspects of vaquita life history, potentially including birth and death rates.

Faced with these difficulties, vaquita researchers have turned to photographic identification, which requires high quality photographs to identify individual vaquitas. Photographic identification of vaquitas began in 2008 (Jefferson et al. 2009). Opportunistic efforts resumed in 2017 during the VaquitaCPR effort.

2.2. Summary of 2018 and 2019 research

In September 2018 a 7-day photo-identification effort produced the first evidence that vaquitas could calve annually (Taylor et al. 2019), and showed that a minimum of 6 healthy animals remained in a small area near San Felipe, Mexico. This minimum abundance estimate was the number of animals seen simultaneously and was influential in the abundance estimate for that year (Jaramillo-Legorreta et al. 2019).

Two short efforts focused on photographic identification were conducted in 2019. The first effort had sightings during the set up period in late August and a single sighting (001) during the survey

period from September 2 to 6. The second, and larger, effort from October 15-27 had sightings 002-007, which are detailed in Appendices 1-3. Full details of the 2019 field effort are available in the survey report (Rojas-Bracho *et al.* 2019 available at iucn-csg.org).

2.3. Description of Expert Elicitation

Expert elicitation is a means by which available data can be combined with expert judgements, to temporarily fill a knowledge gap, via the development of probabilistic distributions. Specifically, expert elicitation is a formal technique, first developed in the 1950s and 60s (Brown 1968, O'Hagan et al. 2006), now widely applied in ecology and conservation where there is a relative lack of data but an urgent need for conservation or management decisions (Runge et al. 2011, Martin et al. 2012). Specifically, Morgan (2014) indicates: *"Expert elicitation should build on and use the best available research and analysis and be undertaken only when the state of knowledge will remain insufficient to support timely informed assessment and decision making".* Martin et al. (2012) describe how this technique can be used to access substantive knowledge on particular topics held by experts and such techniques have been discussed and used widely recently (e.g. MacMillan and Marshall 2006, Aspinall 2010, Knol et al. 2010, European Food Safety Authority 2014, Sivle et al. 2015), including in the assessment of risks from climate change (Lenton et al. 2008) and future sea level rise (Bamber and Aspinall 2013). The technique can also be used to translate and combine information obtained from multiple experts into quantitative statements, while minimizing bias in the elicited information, and ensuring that uncertainty is accurately captured. The formal process of expert elicitation aims to address many of the well documented problems, heuristics and biases that arise when the judgements of only a few experts are canvassed or where expert knowledge is sought in an unstructured matter (Kynn 2008, Kahneman 2011, Morgan 2014). In the field of marine mammals, a number of elicitations have been conducted in recent years involving the project team and seeking to improve the methods for marine mammal issues (Booth et al. 2016, Tollit et al. 2016, Booth and Heinis 2018, Booth and Thomas 2021). Several different approaches have been used to elicit experts' beliefs to quantify what was seen in the 2019 survey and are described below in 2.3.1 through 2.3.3. Experts and the exercises in which they participated are given in Appendix 1.

2.3.1. Description of Structured Expert Decision Making (SEDM)

Structured Expert Decision Making (SEDM) is a process that first went by the name of the Delphi method and is also called Expert Elicitation. It was developed by the RAND corporation in 1948 to overcome groupthink[1] and the influence of dominant individuals. Burgman (2015) modified this method and called it IDEA (for Investigate, Discuss, Estimate, Aggregate). SEDM follows these steps: 1) provide experts with relevant evidence; 2) elicit expert beliefs; 3) examine and discuss initial results to clarify whether differences arise from true differences in beliefs, or from evidence that was available to some but not all experts, or from issues arising from different understandings of the question being posed; 4) clarify question and extend evidence if necessary; 5) revise expert beliefs independently if necessary; 6) express results as a distribution of averaged expert beliefs, and sometimes also as independent distributions anonymously in an appendix. This method has been used in many status reviews for species at risk, including many for marine mammals within the United States (Krahn et al., 2002, Oleson et al., 2010, Rosel et al., 2016).

[1] A process in which arbitrary starting positions and hidden agendas of a few lead a group to a decision that does not reflect the individual participants' private opinions.

2.3.2. Description of Expert Elicitation using Rational Independent Observer (RIO) process (EE-RIO)

The objective of an expert elicitation is to construct a probability distribution to accurately represent the knowledge and beliefs of an expert or group of experts regarding a specific Quantity of Interest (QoI). For each QoI, which has a true but unknown value, 'X', each expert is asked to provide their individual judgements of X using a number of parameters: the plausible limits, median, lower and upper quartiles. The plausible limit was defined such that it may be theoretically possible for the true value of X to lie outside these limits, but that the expert would regard it as extremely unlikely that X was outside this range. Plausible limits are elicited first to try to avoid well-known biases of overconfidence (where experts do not consider extreme cases for X) and anchoring (where experts start with a value of X in mind). Following that, experts are asked to consider their median value for X (such that there is equal probability that the true value of X lies above or below the median (but within the plausible limits). Experts also provide lower and upper quartile values for X.

Experts are then asked to input their personal judgements into a web-interface form (Figure 2.3.1.1) and to send the data to the facilitator (via the form). The judgements are then input into SHELF and distributions fitted to each individual expert judgement with the best statistical fit (determined in SHELF as the distribution with the lowest sum of squared deviation between the specified quantiles and those fitted using a least-squares algorithm; candidate distributions were normal, t, shifted gamma, lognormal, log-t and shifted scaled beta). The facilitator then presents the anonymized individual judgements of all experts to the group (see an indicative example in Figure 2.3.1.2). During the process, the mechanisms that experts had considered in making their individual judgements are discussed among the group.

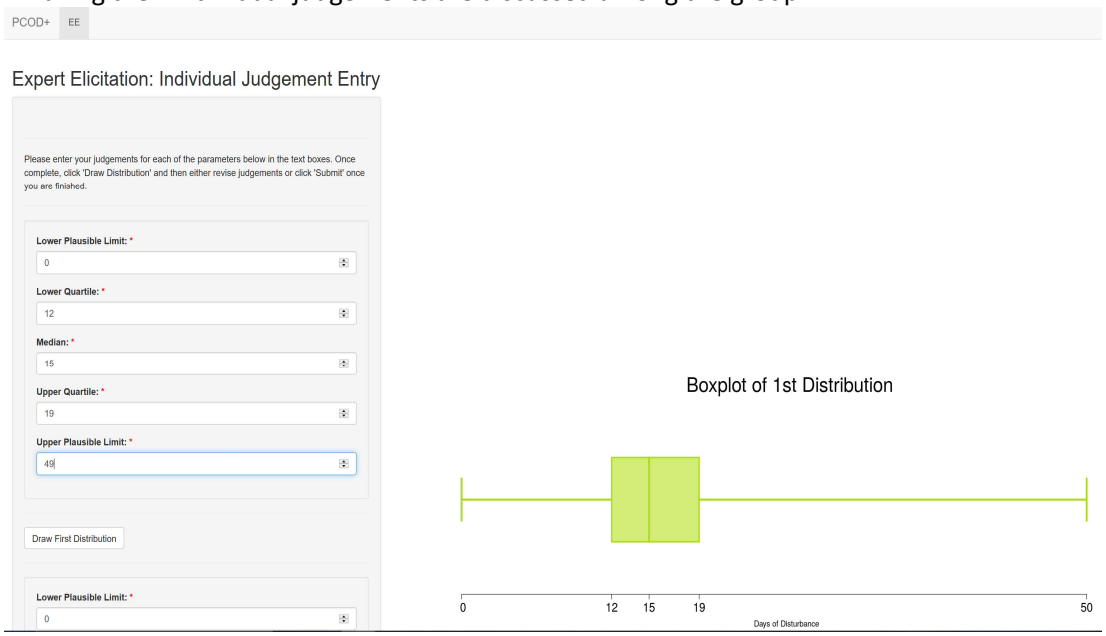


Figure 2.3.1.1- Example screengrab of online tool used by experts to provide their personal judgements, allowing them to be easily collated by the EE facilitator.

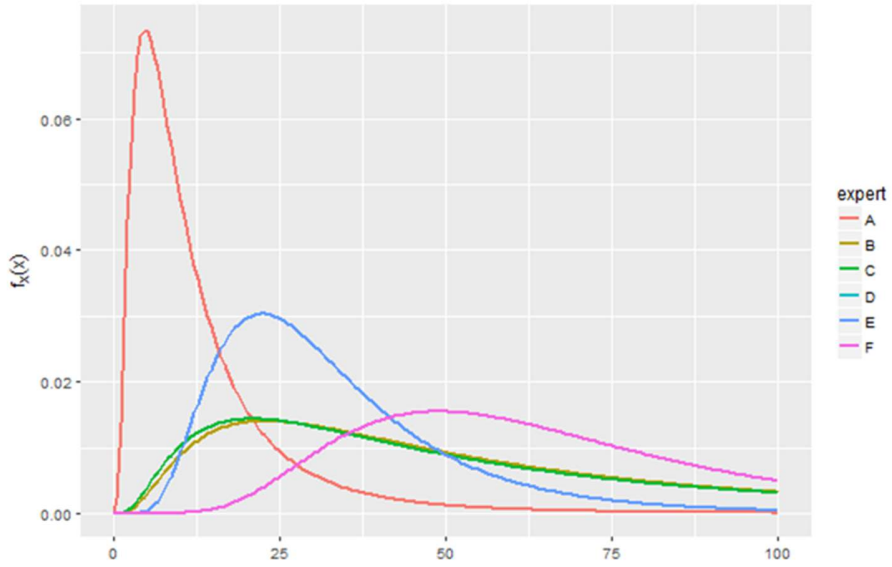


Figure 2.3.1.2 - Example fitting of example individual judgements using SHELF v3.0 software (not from this elicitation).

Experts are invited to justify their judgements, particularly those that were divergent, to ensure that the range of judgements had been discussed openly. Following this, the group was asked to reach a 'group consensus' judgement (in the form of a probability distribution). It is important to note here (and stated clearly to experts), that there was no expectation that the experts would reach complete agreement on a probability distribution for our QoI. That is because it is unlikely that there is one single distribution that would be accepted as perfectly representing the opinion of all experts. Instead, experts are asked to discuss and agree upon a distribution representing the reasoned opinions of a theoretical external observer, called a Rational Impartial Observer (or RIO), who was party to all of the information and discussions that had taken place. The RIO would not have identical views to any one of the experts but would instead find some merit in all the differing arguments or justifications – and give some weight to each.

2.3.3 Description of the Roulette method

The Roulette method elicits expert's beliefs for a set of integer values and can be used for either the SEDM or EE RIO elicitation types. Experts are given a certain number of 'probs' (for example 10) and asked to distribute their probs among a specified set of values determined that contains all values the experts might consider plausible. If using the SEDM process the following steps would be taken:
1) take training course; 2) read evidence dossier; 3) conduct a virtual meeting with introduction to process the evidence; 4) independently submit beliefs in integer values for questions; 5) discuss rationales for choices; 6) independently resubmit belief distributions. The final distribution for SEDM simply weights each expert the same and sums all beliefs.

3. Methods
   3.1. Description of process used for SEDM for 2019 vaquita data

Experts consisted of participants in the field efforts including both observers and data recorders. All experts had extensive experience with vaquitas and often harbor porpoise. Such experience contributed to their understanding of both behavior and life history that could bear on the estimation questions at hand. Experts were provided by email the document in Appendix 2 and asked to provide their expert opinion on a number of questions (Appendix 5). All 13 experts who participated in the field effort were asked to participate. Nine experts responded and participated in the elicitation effort (i.e., only a subset of those who participated in EE-RIO and EE-roulette). Some logical inconsistencies in the responses were identified, prompting a second elicitation. A second elicitation is common practice, and was deemed important in this case both because the probability problem presented to experts was complex and the initial elicitation was given hurriedly to observers at the end of the survey. This second elicitation used responses from the first to give the experts some feedback that would aid in their estimation of the total number of calves seen and the total number of vaquitas seen. For example, sightings 002 and 003 both had animals that could have been calves and could have been resightings of the same pair at different times during the same day. Therefore, depending on the probabilities that each small individual was a calf and that the sightings were the same or different, there could be zero, one or two calves. The experts were given these probabilities from their original responses (Appendix 6 shows the probabilities for an anonymous expert, but each expert was only provided their own probabilities).

Unlike the process described for SEDM above, there was no discussion session as the effort was done after the survey by email. Virtual meetings were not commonly used at this point in time and many experts were in the field and unavailable for such a virtual meeting.

3.2. Description of process used for EE-RIO for 2019 vaquita data

In September 2020 an expert elicitation was conducted and facilitated by Cormac Booth and Len Thomas from SMRU Consulting and Centre for Research into Ecological and Environmental Modelling (CREEM, University of St Andrews). This pair has developed expert elicitation approaches for marine mammal issues and facilitated many similar elicitations in recent years (Booth & Thomas, 2021). All 13 observers/recorders participated in this exercise (Appendix 5). Observers were given the same vaquita sighting summary that was now called an 'evidence dossier' (Appendix 2). All experts took the expert elicitation training.

The expert group discussed and agreed to provide their judgements for the following QoI:

1. How many unique individual vaquita were sighted during the 2019 survey (including adults, juveniles and potential calves)?
2. How many unique individual calves (only) were sighted during the 2019 survey?

To avoid linguistic uncertainty, the definition of a calf for the purposes of the elicitation was:

"*While calves are biologically defined as less than one year old, age cannot be determined in the field. Therefore, inferences about whether an animal is a calf is determined through a combination of size (as compared to the other member of the pair) and behavior. Vaquitas in October are roughly 6 months old and therefore the size can be both small and only slightly smaller than the adult. The dorsal fin may appear only slightly smaller. Because of this, the behavior is also key. Dependent calf behavior is to surface within one body length and slightly behind the mother. With more time observed with a pair in this conformation (larger individual in the lead followed by a smaller individual within a body length), confidence that the pair is a cow/calf increases. However, because October is roughly the time of weaning, the calf can be*

expected to be more independent and may not always be in this position. The evidence dossier uses both the term 'calf' and the term 'juvenile'. It is assumed that observers using the term 'calf' had confidence that they were observing the behavior above. Use of the term 'juvenile' indicates less confidence but definitely does not preclude that the smaller individual could have been a calf."

The exercise took place on 3 different days with 6 hours of 'face-to-face' meetings on Zoom. Booth served as the primary facilitator with Thomas assisting with statistical input and advice. The experts were asked to input their personal judgements into a web-interface form (see Figure 2.3.1.1) and to send the data to the facilitator (via the form). The judgements were then input into SHELF and distributions fitted to each individual expert judgement with the best statistical fit (determined in SHELF as the distribution with the lowest sum of squared deviation between the specified quantiles and those fitted using a least-squares algorithm; candidate distributions were normal, t, shifted gamma, lognormal, log-t and shifted scaled beta). Each observer received their results and had an opportunity to consult with Booth and Thomas to adjust their input to better match their desired distribution.

The facilitator then presented the anonymized individual judgements of all experts to the group (Appendix 6). The Zoom meeting then 'went around the room' so that each observer could give their rationale for each question (question 2 on calves considered on the first day of discussion and question one on the second and final day). To obtain the RIO distribution, an initial distribution was drawn summarizing the views across the panel of experts. From there the distribution of amalgamated results was critically assessed, beginning with a re-examination of the lower and then upper plausible values. It was clear among the experts that the primary drivers in judgements were the lower and upper estimates of the total number of animals sighted by the observers during the 2019 survey, the photos of some of the sightings, the descriptors of animals sighted – which was critical in assessing the potential for resighted animals. Via these discussions, it became clear that the evidence dossier was incomplete with respect to photographs of both vaquitas that could be individually identified and series of photographs that showed behavior that could improve judgement on whether a calf was present. Observers were given evidence dossier addendums (Appendices 3 and 4). Experts were allowed to resubmit individual judgements after examining the addendums.

3.3. Description of process used for EE-roulette for 2019 vaquita data

Because the primary EE method for this exercise was EE-RIO, the EE-roulette was done immediately after the virtual meeting, so for this exercise not all experts did step 4 described in the methods above. For EE-roulette, step 5 in 2.3.3 above was done in the EE-RIO and experts were asked immediately following those discussions to independently resubmit belief distributions (step 6).

4. Results

4.1. SEDM

The mean estimated number of calves seen was 3 with a 63% belief that there were at least 3 calves (Figure 4.1.1). The mean estimated number of unique vaquitas seen in all 7 sightings was 9.4 with 40% belief that there were at least 10 (Figure 4.1.2).

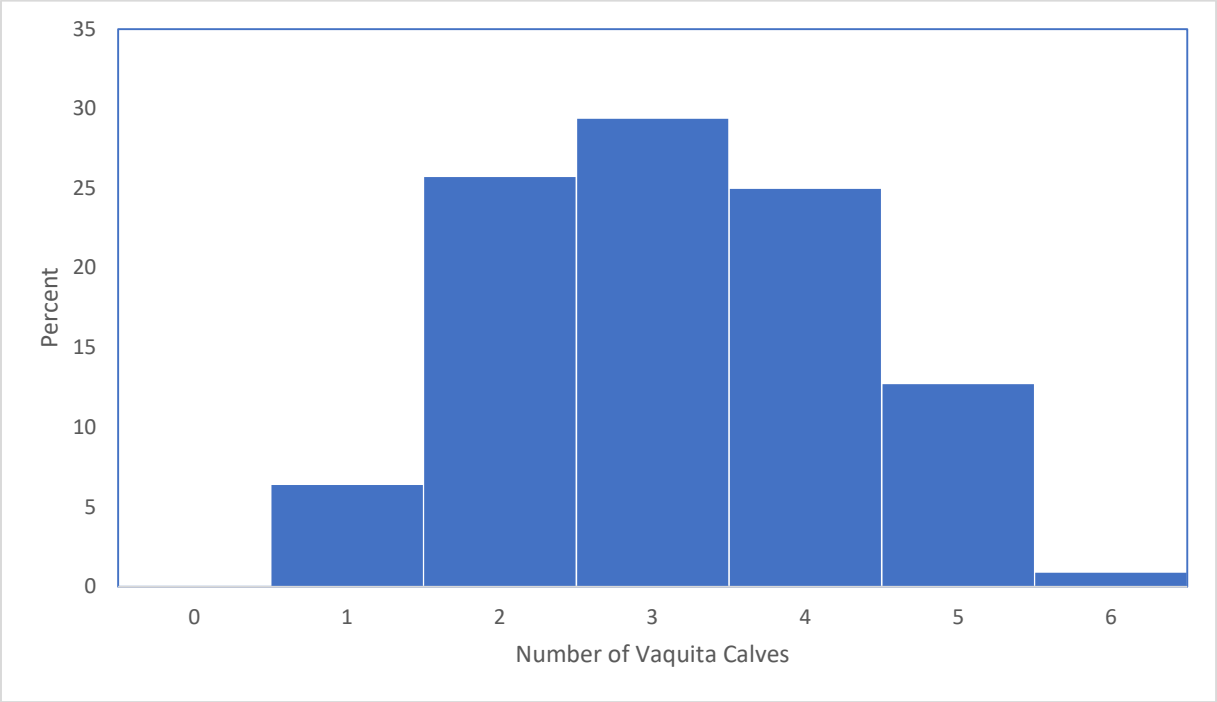


Figure 4.1.1 The distribution for the number of unique calves.

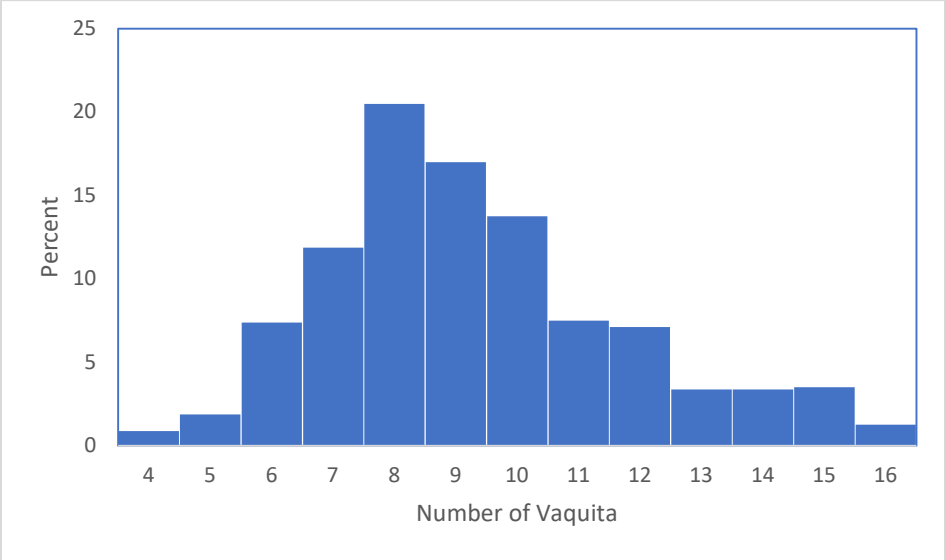


Figure 4.1.2 The distribution for the number of unique vaquitas including calves. Recall that the evidence available to experts did not include the full set of photographs.

4.2. EE-RIO

Experts discussed the main possible mechanisms or drivers for their differing views of the number of calves sighted (initial distributions in Appendix 6). The main drivers were regarding the true number of animals in a group (which experts agreed was affected by the length of the sighting) and the potential for resighted animals (particularly given the intense spatial coverage of the survey). Critically, experts were comfortable eliminating 0 as a potential lower plausible value (this discounts the possibility that no calves were seen during the survey). Experts had varying judgements regarding the upper plausible limit, but agreed that a RIO would agree that 5 calves was a plausible value – but with very low belief that this was the true number of calves sighted. This was due to a widespread belief that there were a significant number of resights during the surveys (in part due to the intensive spatial coverage of a small area during the surveys). Experts agreed that a RIO would estimate that 3 was the most likely integer value for the true number of calves, with approximately equal belief that the true value lied between 1 and 3 or between 3 and 5 calves (Figure 4.2.1). There was a 97.4% chance of 2 or more calves and a 71.4% belief that the true number of calves sighted was 3 or more.

Similar rationale existed in the expert's individual judgements of the true number of unique vaquita sighted during the survey (Figure 4.2.2). Experts agreed that a RIO would judge, based on the discussions, that it was implausible that the true number of vaquita sighted was less than 7 or more than 15, and that the most likely value was 11 with values close to 11 being almost as likely. The RIO distribution indicates a 71.8% belief that the true number of unique vaquita sighted was 10 or more.

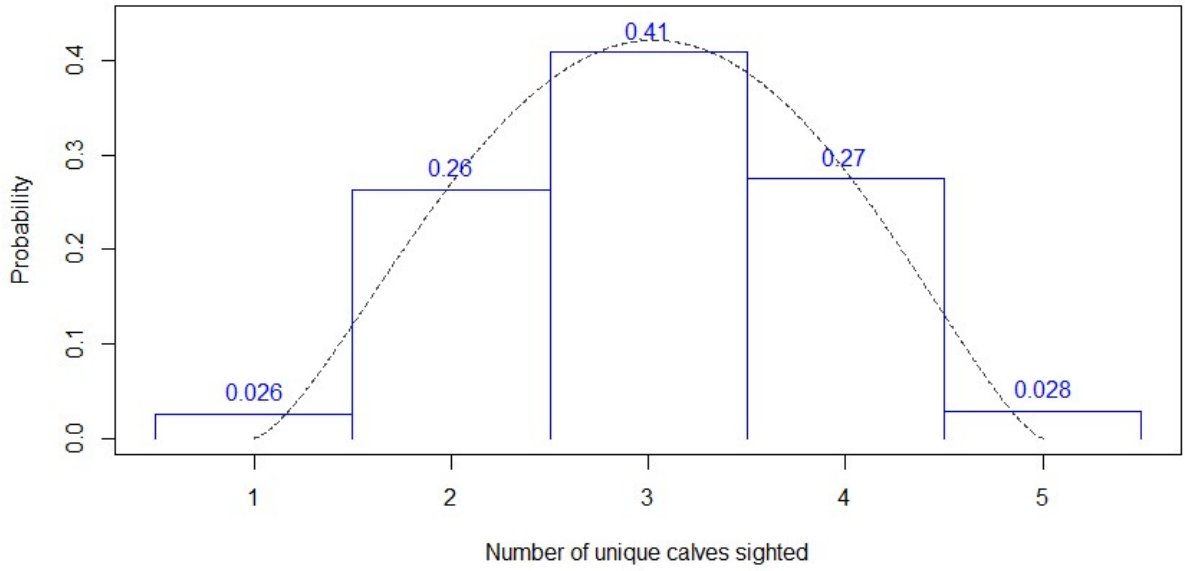


Figure 4.2.1 - The elicited probability distribution for the number of unique vaquita calves sighted during the 2019 survey. The bins show the probability for each integer value and the dotted line shows the resulting continuous distribution.

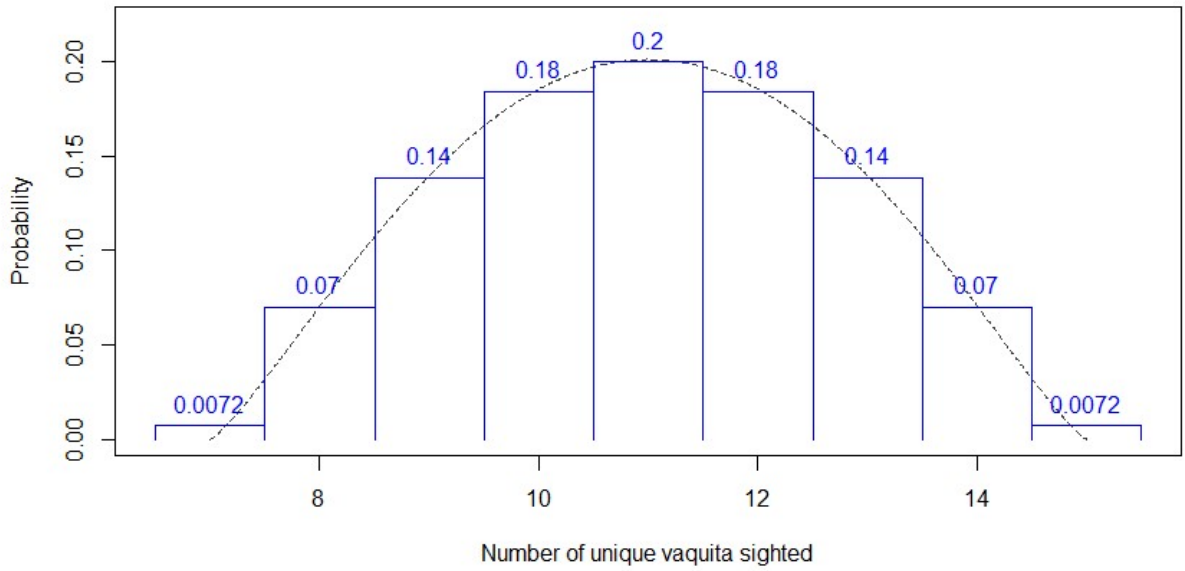


Figure 4.2.2 - The elicited probability distribution for the number of unique vaquita sighted during the 2019 survey. The bins show the probability for each integer value and the dotted line shows the resulting continuous distribution.

4.3. EE-roulette

The mean estimated number of calves seen was 3.1 with a 73% belief that there were at least 3 calves (Figure 4.3.1). The mean estimate for the number of unique vaquitas seen in all 7 sightings was 10.4 with 66% belief that there were at least 10.

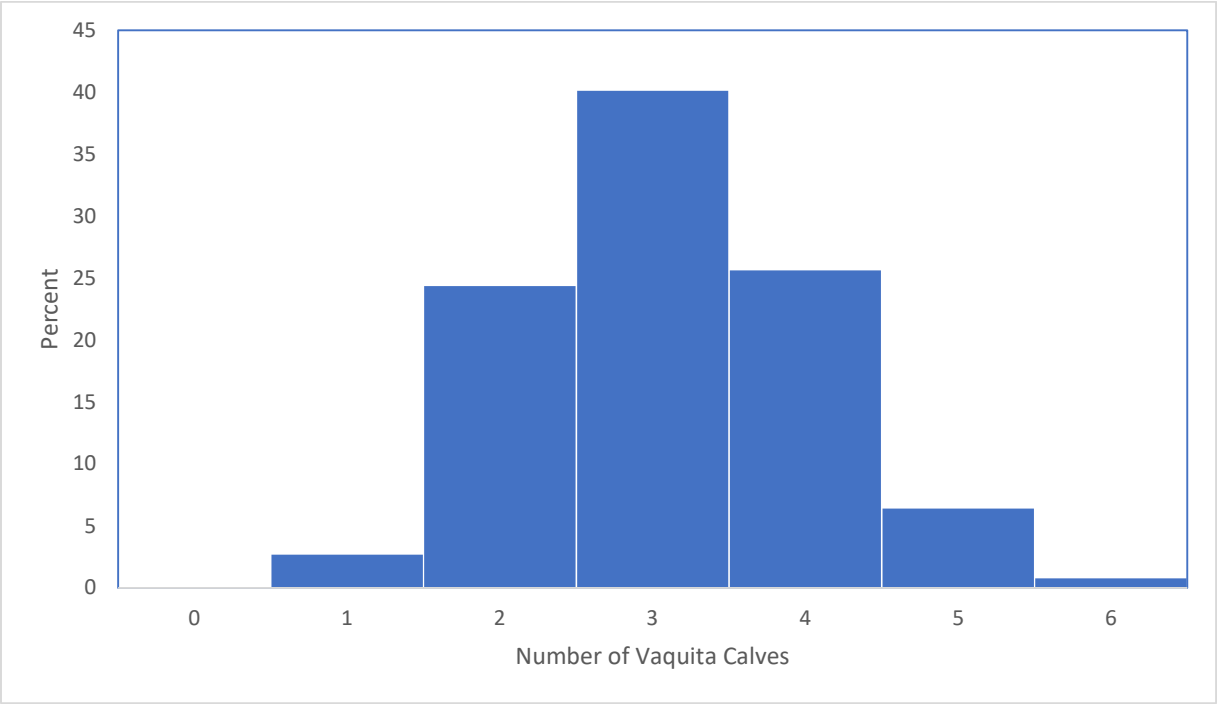


Figure 4.3.1 shows the distribution for the number of unique calves.

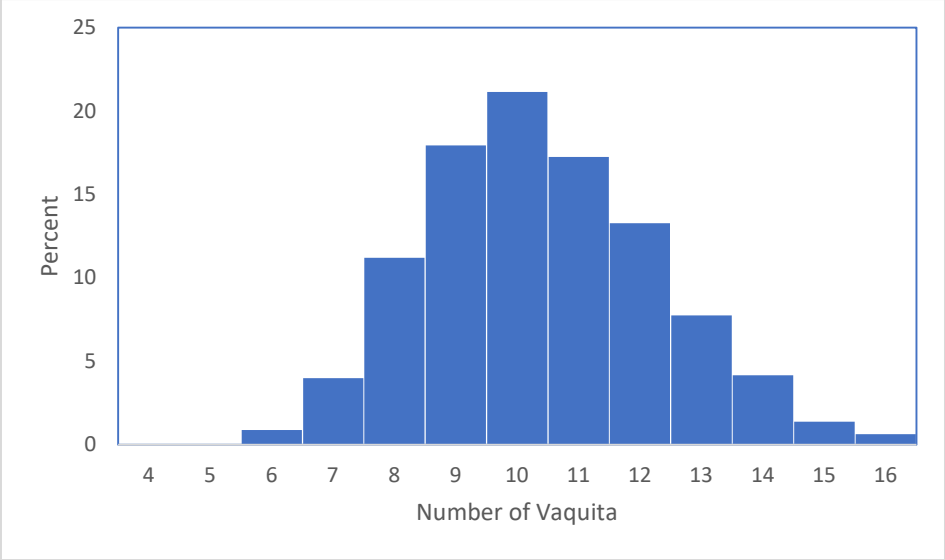


Figure 4.3.2 shows the distribution for the number of unique vaquitas including calves.

4.4. Comparison of different approaches (with caveats)

Because the SEDM was not done with discussion (as it should be) and did not have the same evidence presented to the experts (i.e. the SEDM exercise was done without the photographic evidence in Appendices 3 and 4), it is not considered in this comparison.

The distributions for the number of unique calves was very similar between the EE-RIO and the EE-Roulette. The difference in skew is more pronounced when comparing EE-RIO for total unique vaquitas (which is symmetrical) to EE-Roulette (which has a right skew). The differences are small, but may be worth investigating if these distributions are used in the future in models to estimate rates of decline or abundance.

5. Conclusions

5.1. SEDM post-survey exercise positives and negatives

Following the October 2019 vaquita field effort, a summary of the sightings was written up and distributed to observers who were asked to respond to a number of questions that included:

1) what were the total number of unique vaquitas sighted, and
2) what were the total number of unique calves sighted.

As noted earlier, this post-survey email exercise did not follow standard SEDM steps of providing evidence, obtaining initial responses, discussion, potential modification of responses and summary of responses. Omitting the discussion was a major flaw that can now be rectified by virtual meetings.

After reviewing the initial responses, a second round of elicitation was completed that gave observers some feedback on their responses in the first iteration. This was done because it was clear that observers were giving non-trivial probabilities to two different sightings within one day as both having calves and being separate sightings and yet were giving high probabilities when considering all 7 sightings to there being only a single calf. This second iteration did indeed have higher estimates for the number of calves, but the total unique was basically unchanged.

Positives:

- The elicitation was done immediately after the survey when expert's memories were best
- The framing of the exercise was in integer terms that were easily understood by experts

Negatives:

- Not all experts participated
- The question method used was apparently too detailed for many observers
- The email method did not allow for discussion

5.2. EE-RIO positives and negatives

Positives:

- Initial anonymous entry of judgements allows for all to contribute their personal views – ensuring no one is overlooked.
- Standardized framework for reporting and tracking expert judgement.
- Full discussions after initial judgements allows clarification of any linguistic uncertainty and addition of evidence missed in the initial dossier
- Final product is a single distribution that is designed to reflect scientific consensus (via the RIO process), and is insensitive to any potential extreme views and to the exact make-up of

the expert panel (since experts are encouraged to think about the views of those who are not attending, but would be given the same evidence).

Negatives:

- Such processes are mentally taxing, for even the most quantitative of scientists, and it therefore requires careful planning to avoid fatigue.
- The estimation of how many animals were resighted was a key component of this elicitation. Unfortunately, not all information experts wanted to have regarding sightings was always available (e.g. details of the sighting, photo records to compare between sightings). This is something to be addressed in future EEs. For example, in designing the survey methodology, specific notes or data (i.e. those that informed estimation of resights the 2019 survey EE) could be prioritized during the sightings of animals (see 6.1.3 below). This would ensure it can be available for discussions/expert panels.
- Use of continuous distributions was potentially problematic when only a small number of integer answers were possible. The lowest and highest plausible values were straightforward because these were likely given as integers, but the quartiles were not intuitive values. The RIO method could be done allowing experts to quantify their beliefs as integers. That process would also require less 'help' from outside experts to fit distributions, work with experts to help them obtain distributions that fit their expectations and then to produce the summary distribution for the group.
- The process of the group considering the plausibility of the highest and lowest plausible values has the potential of some experts who hold the most extreme values as plausible being subject to peer-pressure, i.e. there is potential for 'group think' if some experts have a strong ability to persuade others. The facilitator role is essential in mitigating this.
- A possible negative for vaquita surveys is that an experienced facilitator is needed to guide the RIO discussion. The exercise would be best done immediately after the survey ends when memories are fresh and it may be difficult to get an experienced facilitator on that date.

5.3. EE-roulette positives and negatives

Positives:

- Initial anonymous entry of judgements allows for all to contribute their personal views – ensuring no one is overlooked.
- Distribution belief 'points' to integer values is intuitive for experts.
- Full discussions after initial judgements allows clarification of any linguistic uncertainty and addition of evidence missed in the initial dossier
- Although some facilitation is needed to ensure all observer's opinions are heard and respected, the need for fitting distributions and bringing the group to a RIO consensus is avoided.
- The SEDM process is well-known and has been established as a valid method to avoid 'group-think' because after discussion, experts again express their judgement independently.

Negatives:

- The final distribution will contain integer values that may be more difficult to communicate statistically.

- The final distribution is likely to represent more 'extreme' views that are excluded in the RIO method (i.e. will give more weight to the tails). This result was observed in this exercise, however, the greater weight in the tails of the distribution was small.
- The final distribution is more dependent on which experts take part than the RIO method.

6. Recommendations
   6.1. Changes to field practices (observers)
      6.1.1. It would be useful to have a printer or a general use laptop, so that all observers could contribute to and read the evidence dossier at it was assembled.
      6.1.2. Agree a working definition of 'calf' prior to research starting (Barb/Lorenzo/Tim)
      During the exercise, it was apparent that different experts had different standards for the amount of evidence required to believe a small individual was a calf of the year. Assessing whether small individuals are calves in the fall is difficult because calves are about 6 months old and are spending more time separated from the mothers as they wean. Recording the following standard data for pairs would be helpful: 1) were dorsal fins seen such that relative size could be assessed well, 2) if there was a smaller individuals, was it observed in the 'calf position', 3) was the pair disturbed such that a calf might be expected to move into 'calf position', 4) what was the total time observed, 5) what was the closest distance observed.
      6.1.3. Agree elements that should be considered in descriptions of each sighting. In addition to the factors 1-5 above, the following should be recorded: 1) general behavior (milling, traveling in a consistent direction, reacting to vessels or nets), 2) were photographs taken and if so by whom, 3) if photographs were not taken, can verbal descriptions of dorsal fins be given (and were the fins seen in a perpendicular position).
      6.1.4. Develop evidence dossier as sightings occur and discuss with experts after each event. This would mean compiling photographs from events and making them available to all observers as sightings occurred. If possible, all participants in the sighting (which may be on multiple vessels) should meet to discuss the sighting events in the evening of the event. The write-up of the sightings should be made available to all experts as the survey progresses. Experts would be asked to update their responses to the 2 main questions (number of unique calves and number of unique vaquitas) after each day that had a sighting.
      6.1.5. Schedule an extra day for final elicitation
      The ideal practice would be to have experts make initial judgements on the previous evening to the final elicitation day. The judgements could then be complied for use in the discussion the following day. Discussion should happen in the morning so that the evidence dossier can be adjusted to include any new information or to clarify any language. Details will depend on whether EE-RIO or EE-Roulette is used. If a skilled facilitator like Booth or Cormac is unavailable, then the likely best option if to use EE-Roulette.
      6.1.6. Start the discussion with calf estimate, followed by total. This order simplified the estimation of the total because experts knew that their total unique individuals had to include the estimates of unique calves multiplied by two as a minimum.
   6.2. Recommended training
      6.2.1. All observers should take the free EE training before going into the field

6.2.2. Those who have taken the training, should take again but concentrate on the 'probs' section (note that we need to discuss this…would be better to decide whether to use discrete or continuous distributions beforehand)

6.2.3. If possible, new training should be developed to help observers with a closer example to what they are actually doing.

6.3. Changes to analysis process

6.3.1. Be prepared with a plan A and a plan B if EE-RIO is selected as the 'best' method, but an experience facilitator may not be available on the needed day (Barb/Lorenzo/Tim and consult with Cormac and Len)

6.3.2. Agree how the data will be used in analyses that project the vaquita population forward (Len's proposal to choose minimum from distribution first, then project and reject if trajectory is less that minimum value) (writing for this report by Barb/Lorenzo/Tim, but action item for analysis team)

6.3.3. Agree how to evaluate whether the trend in abundance has changed (writing for this report by Barb/Lorenzo/Tim, but action item for analysis team)

References

Aspinall, W. 2010. A route to more tractable expert advice. Nature **463**:294.

Bamber, J. L., and W. Aspinall. 2013. An expert judgement assessment of future sea level rise from the ice sheets. Nature Climate Change **3**:424.

Booth, C., C. Donovan, R. Plunkett, and J. Harwood. 2016. Using an interim PCoD protocol to assess the effects of disturbance associated with US Navy exercises on marine mammal populations Final Report.

Booth, C., and F. Heinis. 2018. Updating the Interim PCoD Model: Workshop Report - New transfer functions for the effects of permanent threshold shifts on vital rates in marine mammal species. Report Code SMRUC-UOA-2018-006, submitted to the University of Aberdeen and Department for Business, Energy and Industrial Strategy (BEIS), June 2018 (unpublished).

Booth, C.; Thomas, L. An Expert Elicitation of the Effects of Low Salinity Water Exposure on Bottlenose Dolphins. Oceans 2021, 1, 179–192. https://doi.org/10.3390/oceans2010011

Brown, B. B. 1968. Delphi process: A methodology used for the elicitation of opinions of experts. Rand Corp Santa Monica CA.

Burgman, M. A. 2015. Trusting judgements: how to get the best out of experts. Cambridge University Press, Cambridge, UK.

European Food Safety Authority. 2014. Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment.

Kahneman, D. 2011. Thinking, fast and slow. Macmillan.

Knol, A. B., P. Slottje, J. P. van der Sluijs, and E. Lebret. 2010. The use of expert elicitation in environmental health impact assessment: a seven step procedure. Environmental Health **9**:19.

Krahn, M. M., P. Wade, S. Kalinowski,M. Dahlheim, B. Taylor, B. Hanson, G. Ylitalo, R. Angliss, J. Stein and R. Waples. Status Review under the Endangered Species Act: Southern Resident Killer Whales (*Orcinus orca*) 2002. NOAA Tech. Memo NMFS-NWFSC-54, 135p.

Kynn, M. 2008. The 'heuristics and biases' bias in expert elicitation. Journal of the Royal Statistical Society: Series A (Statistics in Society) **171**:239-264.

Lenton, T. M., H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber. 2008. Tipping elements in the Earth's climate system. Proceedings of the National Academy of Sciences **105**:1786-1793.

MacMillan, D. C., and K. Marshall. 2006. The Delphi process–an expert-based approach to ecological modelling in data-poor environments. Animal Conservation **9**:11-19.

Martin, T. G., M. A. Burgman, F. Fidler, P. M. Kuhnert, S. Low-Choy, M. McBride, and K. Mengersen. 2012. Eliciting expert knowledge in conservation science. Conservation Biology **26**:29-38.

Morgan, M. G. 2014. Use (and abuse) of expert elicitation in support of decision making for public policy. Proceedings of the National Academy of Sciences **111**:7176-7184.

O'Hagan, A., C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley, and T. Rakow. 2006. Uncertain judgements: eliciting experts' probabilities. John Wiley & Sons.

Oakley, J., and A. M. O'Hagan. 2016. SHELF: The Sheffield Elicitation Framework (version 3.0). School of Mathematics and Statistics, University of Sheffield, UK, Available from: http://tonyohagan.co.uk/shelf., Sheffield, UK:.

Oleson, E.M., C.H. Boggs, K.A. Forney, M.B. Hanson, D.R. Kobayashi, B.L. Taylor, P.R. Wade, and G.M. Ylitalo. 2010. Status review of Hawaiian insular false killer whales (*Pseudorca crassidens*) under the Endangered Species Act. NOAA Technical Memorandum NMFS-PIFSC-22.

Rosel, P.E., P. Corkeron, L. Engleby, D. Epperson, K.D. Mullin, M.S. Soldevilla, B.L. Taylor. 2016 Status review of Bryde's whales (Balaenoptera edeni) in the Gulf of Mexico under the Endangered Species Act. NOAA Technical Memorandum NMFS-SEFSC-692.

Runge, M. C., S. J. Converse, and J. E. Lyons. 2011. Which uncertainty? Using expert elicitation and expected value of information to design an adaptive program. Biological Conservation **144**:1214-1223.

Sivle, L. D., P. H. Kvadsheim, C. Curé, S. Isojunno, P. J. Wensveen, F.-P. A. Lam, F. Visser, L. Kleivane, P. L. Tyack, and C. M. Harris. 2015. Severity of expert-identified behavioural responses of humpback whale, minke whale, and northern bottlenose whale to naval sonar. Aquatic Mammals **41**:469.

Tollit, D., J. Harwood, C. Booth, L. Thomas, L. F. New, and J. Wood. 2016. Cook Inlet Beluga Whale PCoD Expert Elicitation Workshop Report. Prepared by SMRU Consulting North America for NOAA Fisheries.

ACKNOWLEDGEMENTS
We would like to thank an anonymous funder for supporting this workshop. Also, to all the observers for their hard work during our surveys and for the workshop. We thank St. Andrews for supporting Len Thomas' time and Len's dedication to vaquita work. Our appreciation to Diego Ruiz Sabio and Museo de la Ballena y Ciencias del Mar for supporting the survey, the captain and crew of the Narval. To Sea Shepherd Conservation Society and Eva Hidalgo for their diligent work and use of their vessels the Farley Mowat and Sharpie. We especially thank the crews and volunteers of these vessels and the fishermen who drove the small boats. We also appreciate the support of the Mexican Navy, the staff of the Sector Naval de San Felipe and the personnel of PROFEPA during the surveys. A very special thanks for Paco Valverde for permitting this research and his continual support of vaquita science and conservation. His death from COVID is a huge personal loss to us all and to vaquitas as a species.

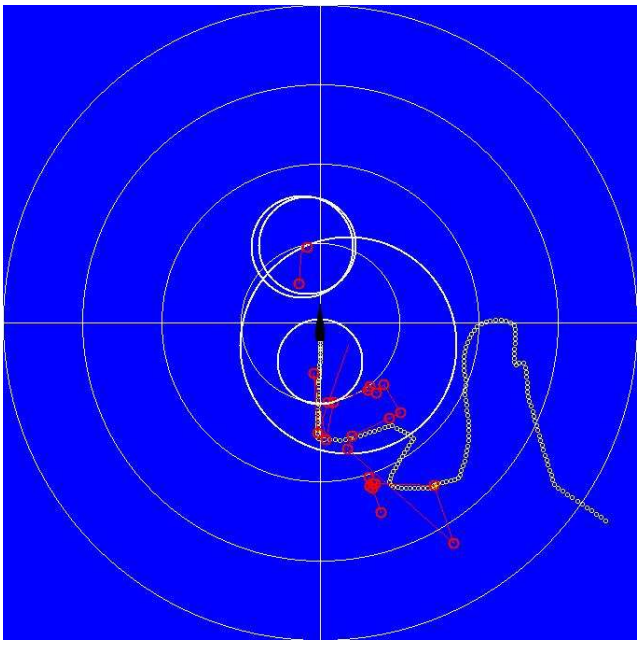
Appendix 1. List of experts

All experts participated in one or more of the field efforts as either an observer, recorder or survey leader. Those participating in the original SEDM exercise are noted with an asterisk (*) and those participating in all the other expert elicitations are noted with a hash (#).

Jay Barlow *#
Tim Gerrodette *#
Chris Hoeffer #
Tom Jefferson *#
Sergio Martinez #
Sarah Mesnick #
Paula Olson *#
Robert Pitman *#
Todd Pusser #
Lorenzo Rojas-Bracho *
Juan Carlos Salinas #
Barbara Taylor *#
Adam U *#
Ernesto Vazquez #
Suzanne Yin *#

Appendix 2. Evidence Dossier
OBSERVATIONS ON VAQUITA SIGHTINGS 3 SEPTEMBER, 17, 19, & 27 OCTOBER 2019

General information on vaquitas and porpoise
The average group size of vaquitas in all surveys (1997, 2008, 2015) was 2. Although the winter period remains mostly unobserved, neonate vaquitas were recovered from early totoaba fishing in February. Thus, calves in October would be 6-8 months old, which is thought to be roughly the age of weaning. The individual marked below as C18 was considered very likely to be a calf.



The red symbols are that same pair from 2018 recorded in a WinCruz screen shot. The pair had a meandering pattern. Each of the regularly spaced, concentric circles is 1 nmi. (Ignore the bright yellow circles of differing sizes.) The total time of the sighting was 1 hour. The most distant points of this pair within the hour covered 4 nmi. Conceivably if vaquitas travelled in a straight line, they could cover 8 miles in an hour, but more likely you wouldn't expect them to move more than what is seen here (about 4 nm).

Notes for 2019
Even though there were only 7 sightings, it may be helpful to write each down and draw lines connecting various sightings that could be duplicates to help with your final estimate of the plausible numbers seen in 2019.

Summary of Observations of the 7 sightings in 2019

3 September 2019
Sighting number 001
Sighting length: 59 minutes
Observation conditions: Beaufort 1
Observed by: Pitman, seen by all observers and recorders on both vessels
Time of day: 9:22
Behavior summary: The pair remained closely associated throughout the period but did not surface with one routinely in the 'calf' position.

Narrative: 2 individuals were sighted by Robert Pitman at 09:22. 3 small boats were launched but the vaquitas did not surface as closely to those vessels as to the 2 ships (Narval and Sharpie) where the best photographs were taken. Tracking was halted at 10:21 when it was determined that photographs of ID quality had been taken and the search could resume for different vaquitas. This was the only sighting of the September 2-6 period. This pair were adults and one was matched to a photograph from 2018. The pair remained closely associated throughout the period but did not surface with one routinely in the 'calf' position.

17 October 2019
Two sightings of vaquita were detected on 17 October 2019, occurring approximately *five and a half hours* and 10 nautical miles apart (sightings no. 002 and no. 003). Both groups consisted of a pair of animals, one larger and one smaller.

Sighting 002
Sighting length: 16 minutes
Observation conditions: Beaufort 2
Observed by: Olson,
Time of day: 7:23
Behavior summary: The pair were always within one body length of one another and the smaller animal was always observed to be trailing the larger.

Observer narratives:
Paula Olson (Farley Mowat):
In sighting no. 002 the larger animal was adult size and the smaller was juvenile size. Five surfacings were observed over a 16 minute period. The pair were always within one body length of one another and the smaller animal was always observed to be trailing the larger. Swim directions at the surface were South, East, South, South, and South. On the fourth surfacing, the pair emerged less than 30 meters from a panga. Distant photos were obtained from this surfacing.

Adam Ü (Narval):
Sighting 002 was initially seen from the Farley Mowat. Once the Narval was alerted to the sighting we were able to see the pair of animals for 4-5 surfacings. The animals were seen swimming along a "slick". A gillnet panga rapidly transiting the area stopped ~30m from the animals and appeared to be preparing to set their net. The animals surfaced again ~30m from the other side of the panga, at which point photos and videos were obtained. There were two animals visible on each surfacing of this sighting, with the 2nd animal being smaller and positioned slightly behind the first in the "calf position".

Sighting 003
Sighting length: 5 minutes? 3 surfacings and then lost
Observation conditions: Beaufort 2
Observed by: Ü
Time of day: 13:15
Behavior summary: 2nd animal swimming in the "calf position" alongside and slightly behind the first animal

Observer narrative:

Adam Ü was on the Narval and observed both pairs (002 and 003). Based on his description of the dorsal fin of the adult in the second pair (sighting no. 003), I think the dorsal fins of the adults may have been different shapes. He described the second adult has having a conical, erect dorsal fin. I had good, multiple looks at the dorsal fin of the first adult; this adult appeared to have a more backward-oriented dorsal fin and faintly falcate.

Given the different shapes of the dorsal fins, and the distance between the sightings (although certainly vaquita could travel ten nmiles in four hours) these may have been different pairs.

Sighting 003 was only seen for three surfacings. The first surfacing appeared to be a single animal and it was so quick that I was unable to confirm that it was in fact a vaquita. I gave the bearing and reticle (3.1 or 3.2) to Sarah (data recorder) to input as an "object" and alerted Bob on the 2nd set of bigeyes that I had a potential vaquita sighting. We decided to slow down and turn towards the potential sighting to confirm. Maybe 15 seconds later I saw the animal a second time (3.4-5 reticles) and was able to confirm it was indeed a vaquita. ~15 seconds after that I saw the animal a third and final time (3.8 reticles) and noticed there was a 2nd animal swimming in the "calf position" alongside and slightly behind the first animal. The 2nd animal's fin looked significantly smaller but since it was on the opposite side of the larger animal I did not get a good view of the entire body/surfacing. All three surfacings of the larger animal gave clear perpendicular views (heading W) and the fin of the larger animal was noticeably not falcate; the trailing edge seemed to have a bit of a convex flare to it instead of a concave/falcate curve.

We called the Farley Mowat over to assist with tracking and launched our two small boats but unfortunately we were unable to locate the animals after the third surfacing. I was the only person to see this sighting. I did not get a good enough look at the animals in sighting 002 to be able to confirm/deny that they were the same individuals in sighting 003 but PAO's description of the first sighting suggests the adult animals could have been different.

19 October 2019

Paula Olson (Farley Mowat):
Sighting no. 004 was initially detected over 2 nmiles from the ship; this detection consisted of a single animal with a tall, conical dorsal fin. The animal was seen twice and was swimming south while at the surface. The ship turned toward the sighting location and after moving some distance a resight was made, approximately 0.25 nmile from the original location. This time the group consisted of a mother/calf pair. Both dorsal fins were slightly falcate. The calf was smaller than the juvenile in sighting no. 002 from 17 October. Distant photos were obtained of the mother calf pair (Adam Ü). A third resight was made, again in almost the same location; this time the mother/calf pair was seen as well as a second pair, two adults, that surfaced several body lengths from the mother/calf pair. The two pairs did

not surface at the same time. (See Ernesto's comments, below.) The final resighting from the Farley Mowat was that of the mother/calf pair. The mother/calf were always observed swimming within one body length of one another. The initial single animal may have been one of the pair of adults. A single animal was seen later by the Narval (see below).

Ernesto Vazquez #125 (Farley Mowat):
Resight of sighting no. 004. As the cow /calf was resighted for a couple surfacings, while this 2 went down another couple vaquitas surfaced some body lengths to their right, seen just one time at the surface, they did not surface at the same time, never had a view with the 4 vaquitas at the same time at surface. The pair that was leading the 4 seemed different in size, one larger than the other, but cannot say that was another cow/calf pair, although they were close to each other by a body length.

R. Pitman #005 (Narval) seen about 30 min after the last resight of sighting 004 within about 1-2 miles. A single individual seen twice as it rolled over perpendicular to the vessel. It appeared small and I suspect that an adult was nearby but unseen

October 27 Sightings 006 and 007
Sighting #006
This sighting is seen for 48 minutes. It takes 44 minutes to confirm that there are 4 animals and they are described as not surfacing super close together with one a 'tiny bit smaller' than the others. The two photographed were by then somewhat separate and appeared to be similar in size (see below and compare to photo of cow/calf above). This photo taken from a panga estimated to be about 75-100 m away.

From Tom Jefferson: TAJ- I was on the port bigeyes. I have little to add, except the following. On multiple surfacings of the group, I saw 4 animals, all of adult/juvenile size, but none looked to be a calf of the year. On one surfacing, I saw the 4 animals surface in quick succession, and then another surfacing about 4-5 sec later. That final surfacing could possibly have been a fifth animal, but I think it much more likely that it was one of original 4 surfacing a second time. On later resights of the group, I never saw any indication of a fifth animal.

Adam's photo from this group appears to be a different animal to the pair above. He feels this photo matches to one he saw in sighting 003 above (17 October, which had an adult and a smaller individual).

Sighting 007 seen 51 minutes after the last resight of 006 about 3.2 miles away

From Adam: ACÜ –A mom/calf pair was tracked for a handful of surfacings, during which time Tom and I had multiple good broadside looks at the pair slow-rolling, with the calf being noticeably smaller than mom (Tom says "calf of the year" size") and surfacing in the calf position every time. We launched the small boats but were not able to track the pair for close approaches.

Throughout the sighting Tom and I felt confident that the calf in 007 was smaller than the juvenile we had seen in 006. My photos from sighting 006 showed the aforementioned closely associated pair, but if the pair in 007 were part of 006 it would mean they split from the other pair and traveled in a straight line away from the area they had been consistently occupying when we were following them and away from the direction of travel they were swimming when we first saw them.

From Todd: -- I was able to see several surfacings of both animals, often times nice broad side views in good light as they milled around a small area near two pangas that had gillnets in the water. It was quite obvious that it was a cow/calf pair and the calf was small, I would estimate around half the body length of the adult. The dorsal fin was considerably smaller and I realized instantly that it was a different animal from the group of 3 to 4 (sighting 006) I had seen earlier in the day.

After the sighting, it was asked if it was possible that the cow/calf pair was part of the group of 4 seen earlier in the morning. What I can say for sure is that the cow calf pair was not the pair of the animals we photographed from the panga in the earlier sighting. My estimate from that group in the bigeyes before I joined the panga was 3/4/3. I did not get a good look at the 4th animal in the group and cannot offer any insight into the size of that individual. So, I suppose if that 4th animal was a calf then the cow/calf observed later in the day may have been part of that group. However, the way the calf was surfacing so close to its mother, I think I would have seen that behavior in the previous sighting as I saw several surfacing sequences from the group of three to four. Despite the odd angle to the surfacings, all of those animals appeared to be the nearly the same size. So, my gut feeling is that the cow/calf were not part of the earlier group.

Appendix 3. Addendum 1 to Evidence Dossier
OBSERVATIONS ON VAQUITA SIGHTINGS 3 SEPTEMBER, 17, 19, & 27 OCTOBER 2019

Addendum with photographic evidence.
The original evidence dossier omitted photographs from sighting 001 that can be compared to sighting 006. Please read the original dossier for the other evidence for those sightings and then add the information below from photographs to your consideration of whether animals in these sightings could be the same individuals.

Sighting 001
Individual A (left and right sides)

Individual B

Together

Sighting 006
Individual A

Individual B

Appendix 4. Addendum 2 to Evidence Dossier

Photo addendum to Evidence Dossier
Note that photos were left with resolution so that you can zoom in on animals…but many will be small and pixilated. Both ID photos and behavioral photos are included.
Sighting 001
Identification photos:
Sighting 001
Individual A (left and right sides) Note this individual was identified in 2018 as an adult

Individual B

Together

Sighting 001 Behavior photos:
Taylor notes on photos below: In addition to the ID photos already shown, these photos show a bit more about position and behavior. The pair were not very close during all surfacings, so many photographs are of the same individual (with the wiggly fin) because it generally surfaced second in the trailing position and was the animal I focused on when both didn't fit in the same frame. The last surfacings both surfaced within the same frame, but none show either in the 'calf position'. 15-17 are one surfacing series.

Photo 15 

Photo 16

Photo 17

Photo 28

Photo 29

Photo 30

Sighting 002
Only single photo available

Sighting 003
No photos

Sighting 004

Photo 7

Photo 8

Photo 9

Photo 10

Photo 11

Photo 12

From Adam: This is the only sequence of photos I got from this sighting. The timestamps show the sequence started at 0915:14 and ended at 0915:27. Frames _2308 to the end are of particular interest to me for size, surfacing position, and surfacing timing, of two individuals. These images have been cropped a bit. The rest are as they came out of the camera.

Sighting 005 no photo

Sighting 006 identification photo

Sighting 006 Behavioral photos
Rosales photos that show 3 individuals surfacing simultaneously:
Photo 5 Photo 6

Adams: photos
_2652 - _2654 Upon quick glance it looks like these photos show a “calf position/timing” sequence but close inspection of the high res images shows the animal in the background/behind surfacing first and going down as the animal in the foreground/lead is coming up. So the physical spacing is tight/calf position but the timing sequence we’d expect for that is not correct.
Photo 52 Photo 53

Photo 54

_2693 - _2697 This sequence clearly shows a larger animal surfacing first with a

much smaller animal (both in body and fin size) surfacing in the correct time/space for what we'd expect for the "calf position". These animals were already being pursued so this behavior is similar to what was seen during chases during VaquitaCPR. The rest of the sequence by the panga shows a minimum of three animals at the surface that appear to be the same size. With the knowledge that there was an obviously smaller individual present three seconds before, there must be a fourth animal that's either underwater, behind a wave, or behind the panga.

Photo 93

Photo 94

Photo 95

Photo 96

Photo 97

Appendix 5. Expert Elicitation 1

Observer expert opinion on vaquita sightings on Sept 3, Oct 17, 19, & 27.

The objective of these questions is to quantify expertise on the number of vaquitas seen. Each observer who worked on the October effort should read the Observation document.

1. What is your name?
2. What dates did you work (maximum time is from Oct 14-Oct 28)
3. Assign 10 likelihood points to:
a. Sighting 2 was a mother and calf
b. Sighting 2 is not a mother and calf
4. Assign 10 likelihood points to:
a. Sighting 3 was a mother and calf
b. Sighting 3 is not a mother and calf
5. Assign 10 likelihood points to:
a. Sighting 2 and sighting 3 are separate sightings
b. Sighting 3 is a resight of sighting 2
6. Assign 10 likelihood points to:
a. The small individual in sighting 4 is a calf
b. The small individual in sighting 4 is not a calf
7. Assign 10 likelihood points to:
a. The individual described in sighting 5 is a new sighting
b. Sighting 5 is a resight of sighting 4
8. Assign 10 likelihood points to:
a. The individual in sighting 5 is a calf
b. The individual in sighting 5 is not a calf
9. Assign 10 likelihood points to:
a. The smaller individual in sighting 6 was a calf
b. The smaller individual in sighting 6 was not a calf
10. Assign 10 likelihood points to:
a. The smaller individual in sighting 7 was a calf
b. The smaller individual in sighting 7 was not a calf
11. Assign 10 likelihood points to:
a. Sighting 7 was a new sighting
b. Sighting 7 was a partial resight of sighting 6
12. Assign 100 likelihhood point to the plausible number of calves seen in October
a. 0 calves (none of the sightings included a calf of the year)
b. 1 calves
c. 2 calves
d. 3 calves
e. 4 calves
f. 5 calves
g. 6 calves (all of the putative sightings with calves were unique calves)
13. What is the lowest plausible number of vaquitas you believe were seen? (give a number)
14. What is the highest plausible number of vaquitas you believe were seen? (give a number)

15. Assign 100 likelihood points to the number of unique vaquita individuals you think were seen in this years' effort (both September and all of October through Oct 27) (use all 100 points, your response should range between your answers to question 13 and 14)

a. 4
b. 5
c. 6
d. 7
e. 8
f. 9
g. 10
h. 11
i. 12
j. 13
k. 14
l. 15
m. 16
n. 17
o. 18

Appendix 6. Expert elicitation 2

The summary and questions were sent independently to the 9 observer/experts. The probabilities in red differed for each observer (shown is one anonymous observer) and were calculated based on answers given by each observer in elicitation 1.

Short summary of an anonymous expert's responses to sightings

Sept 3
Sighting 001 2 adults

Oct 17
Sighting 002 2 animals (1 possible calf)
Sighting 003 2 animals (1 possible calf)

Your probability that sightings 2 & 3 are separate sightings = .80
Your probabilities for:
0 calves 0
1 calf .28
2 calves .72

Oct 19

Sighting 004 2 animals (1 possible calf)
Sighting 005 1 or 2 animals (could be a calf with adult not seen)

Your probability that 4 and 5 are separate sightings: .70
Your probabilities for:
0 calves 0

1 calf .75
2 calves .25

October 27

Sighting 006 2 animals (1 possible calf)
Sighting 007 2 animals (1 possible calf)

Your probability that 6 and 7 are separate sightings: .8
Your probabilities for:
0 calves 0
1 calf .44
2 calves .56

Appendix 6 – Example plot showing individual judgements and resulting RIO

Below are two figures relating to the number of unique vaquita calves sighted during the 2019 survey. The first (Figure A6-1) shows the individual judgements of all experts (anonymized) demonstrating the range of beliefs before group discussions. Note that much of the discussion following initial judgements involved better defining "calf", as it was clear that experts were interpreting this question differently – which resulted in a revised EE process. The second figure (A6-2) shows the final group consensus plot achieved via a RIO process. The difference in these plots highlights the value of ample time being available to discuss and clarify any issues that underpin the elicitation. In addition, it highlights the value of an experienced facilitator and the RIO approach to understand experts' perspectives and allow the group to explore the viability of different rationales before achieving a RIO consensus.

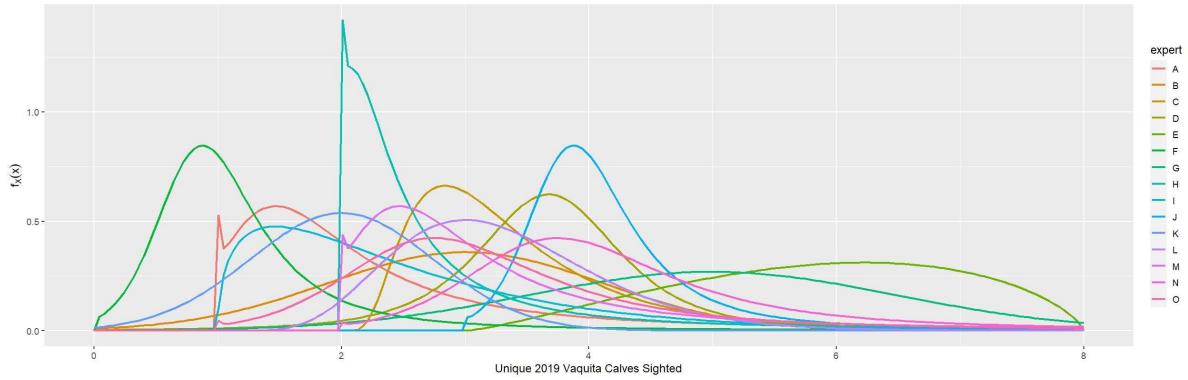


Figure A6-1. Elicited probability distributions of individual experts (each line is a different expert A-O) on the number of unique calves sighted during the 2019 survey.

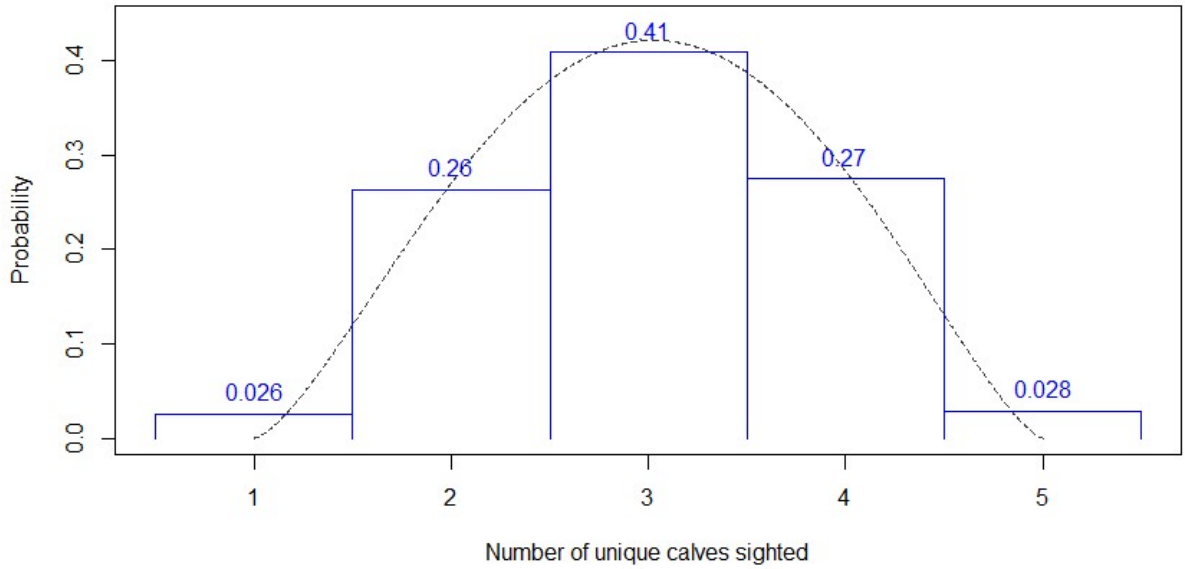


Figure A6-2. The elicited probability distribution for the number of unique vaquita calves sighted during the 2019 survey. The bins show the probability for each integer value and the dotted line shows the resulting continuous distribution.