An interactive decision-making tool for evaluating biological and statistical standards of migrating fish survival past hydroelectric dams

Alejandro Molina-Moctezuma1  |  Joseph Zydlewski1,2

1Department of Wildlife, Fisheries, and Conservation Biology, University of Maine, Orono, ME  
2U.S. Geological Survey, Maine Cooperative Fish and Wildlife Research Unit, University of Maine, Orono, ME

Correspondence  
Alejandro Molina-Moctezuma, Department of Wildlife, Fisheries, and Conservation Biology, University of Maine, Orono, Maine.  
Email: alejandro.molina@maine.edu

1INTRODUCTION

Migratory fish species have complex life cycles that require movements between freshwater and the ocean (Secor, 2015). Dams act as migratory barriers and can decrease fish survival for both upstream (e.g., adult salmons in the Columbia river; Caudill et al., 2007), and downstream migrating fish (e.g., juvenile Atlantic salmon; Norrgård, Greenberg, Piccolo, Schmitz, & Bergman, 2013; Stich, Bailey, & Zydlewski, 2014; Stich, Bailey, Holbrook, Kinnison, & Zydlewski, 2015).

Because of the mortality risks, the Federal Energy Regulatory Commission (FERC) requires hydroelectric dams to conduct “environmental measures to protect, mitigate effects on, or enhance environmental resources” as a condition for a project to be licensed (FERC, 2012, 2019).  
To protect migratory fish, FERC establishes conditions for licensing based on recommendations by the U.S. Fish and Wildlife Service (FWS), and the National Oceanic and Atmospheric Administration National Marine Fisheries Service (NOAA-NMFS; FERC, 2001), to increase efficient fish passage, including fishways, flow control, and
operational constraints. Oftentimes, one of the conditions of a project license is for operators to carry out assessments of fish survival to demonstrate efficacy (e.g., Milford Dam, in the Penobscot River, in Maine; Black Bear Hydro, 2013; FERC, 1998). There is much published on the design of such studies, particularly for salmon species, see Holbrook, Kinnison, & Zydlewski, 2011; Karppinen, Jounela, Huusko, & Erkina, 2014; Norrgård et al., 2012; Skalski, Townsend, Steig, & Hemstrom, 2010; Zydlewski, Stich, & Sigourney, 2017). The survival assessments usually consist of Cormack-Jolly-Seber (CJS) mark-recapture studies using either radio or acoustic-telemetry. These analyses provide an apparent survival estimate, with an associated error. Because of the high stakes related to meeting the survival standard, the selection of the standard can be a point of contention between regulators and operators when negotiating license requirements. Ideally, the established survival standard is based on the necessary survival to reach a conservation goal (here called a biological standard). However, there is a difference between this biological goal (the ecological target to meet a conservation goal), and the statistical survival standard (the value that the estimated survival needs to meet in order to pass and satisfy requirements). Therefore, in each analysis, there are four potential outcomes, in which the survival assessment indicates: (a) correctly that the standard is met ("true positive"), (b) incorrectly that the standard is not met ("false negative"), (c) correctly that standard is not met ("true negative"), or (d) incorrectly that the standard is met ("false positive"); Figure 1). These outcomes parallel theoretical statistical hypothesis testing.

A good test maximizes true detection of pass or fail, and minimized incorrect detections. For our purpose, we adopted terminology typically used in diagnostic tests; "sensitivity" is the probability of passing the standards given that the biological goal is met, while "specificity" is the probability of failing the standard, given that the real survival is below the goal (Figure 1). Based on their opposing objectives, sensitivity is most important for the licensee, while specificity is most important for the regulators. Therefore, balancing sensitivity and specificity is important during negotiations. However, some of the traditionally used standards may not be balanced.

A statistical standard in which the lower confidence limit of a survival estimate (e.g., 75 or 95%) has to be above a biological goal might seem appropriate to a regulator, as the approach has high specificity (thereby minimizing false positives). However, such a standard has low sensitivity, therefore it might not satisfy licensee interests and may not be considered a suitable approach by all parties. Conversely, a standard which requires only that the upper confidence limit of an estimate has to be above the biological goal might satisfy a licensee (because of the desired high sensitivity), but low specificity would mean there is a high probability of encountering false positives (undesirable from the regulators' point of view). Thus, there are inherent trade-offs in establishing a method of assessing a survival standard, and in finding a balance between the interests of operators and agencies.

Note that we are only considering how to find common ground with respect to the method of assessment—not the standard itself. Consensus on the method is necessary to meaningfully discuss the standard value (i.e., the "statistical standard"). Based on ongoing discussions and current FERC licensing in hydropower projects in the West Coast (e.g., Washington Dam, and Oregon Dam; FERC, 2017a, 2017b; Sumner, 2017), and East Coast (e.g., Milford and West Enfield Dams in the Penobscot River, Maine; Black Bear Hydro, 2012; FERC, 2013), we explored four main influences: (a) the method of assessment, (b) the actual value of the standard, (c) the number of years a criterion has to be satisfied consecutively, and (d) the effects of number of released fish and probability of detection, as they affect estimations in a CJS design (White & Burnham, 1999). Demonstrating the influence of these four factors that influence specificity and sensitivity may be a foundational step toward agreeing on a standard—and method of assessment—that satisfies opposing stakeholders on different sides of the table.

Based on observations from FERC relicensing documents, there are three methods of assessment that have been frequently used. One is the use of the "point estimate" (PE, the estimator [usually obtained by maximum likelihood] must be at or above the standard value), the upper confidence interval (UCI; the upper confidence limit must be greater than the standard value), and the lower confidence interval (LCI; the lower confidence limit must be at or above the standard value). For the UCI and LCI, the confidence interval may (theoretically) be defined at any value but typically 95 or 75% confidence intervals have been used in discussions (FERC, 2012) (Figure 2). In ongoing discussions for survival standards for projects in Maine, the three methods have been considered, and the 75% UCI is used in assessing several important projects. Importantly, we note that survival estimates produced by CJS methods have asymmetric error structures (Lebreton, Burnham, Clobert, & Anderson, 1992).

The next trade-off that influences the outcome is the value of the standard that has to be met. Different methods of assessment (i.e., PE, UCI, or LCI) would result in different outcomes for the same data, despite having the same value of the standard, and the same
We constructed a tool to evaluate the probabilities of passing a statistical standard given (a) an a priori true value of survival, (b) a defined method of assessment, and (c) defined study parameters (detection probability and sample size of the simulated population). By running a series of mark-recapture computer simulations, which modeled the studies that are often implemented by dam operators, we were able to explore the probabilities of passing a given standard. Each simulation was analyzed using the same analytical tools and methods used by operators (CJS) so that the implications of study design and data assessment can be compared. We applied these data in an interactive user interface (UI) to allow users to visualize the theoretical consequences of survival thresholds, the assessment technique, and study design. This tool is now available online: https://umainezlab.shinyapps.io/sims/.

2 | METHODS

2.1 | Simulations and survival estimations

Survival estimations are generally carried out by setting up a series of telemetry receiver stations in the river, including receivers above and below the dam. The analysis is the same as for a time-dependent model, but here travel distance in the river (like time) is assumed to be unidirectional (from upstream to downstream for salmon smolts). On either side of the dam, apparent survival ($\phi$) is estimated between stations and probability of detection ($p$) is estimated at each station except the last interval where $\phi$ and $p$ cannot be resolved and collectively are estimated as $\lambda$, the product of these two parameters (Zydlewski et al., 2017). The value of $\phi$ obtained for the interval including the dam is then used to make the decision of whether the standard is met or not. The estimates of $\phi$ and $p$ are strongly influenced by three parameters: (a) true survival ($S$; which is unknown), (b) true probability of detection ($D$; which is also unknown), and (c) the number of tagged fish released in each study ($N$; which is known). Thus, in the real world, there is no way to assess the precision and accuracy of $\phi$ and $p$. While this tool was developed with salmon smolts in mind, it can be used for any migrating species. See Table 1 for a description of abbreviations used.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Definitions of commonly used abbreviations in this publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Estimated probability of survival</td>
</tr>
<tr>
<td>$S$</td>
<td>True probability of survival</td>
</tr>
<tr>
<td>$P$</td>
<td>Estimated probability of survival</td>
</tr>
<tr>
<td>$D$</td>
<td>True probability of detection</td>
</tr>
<tr>
<td>$T$</td>
<td>Probability of passing the standard</td>
</tr>
<tr>
<td>PE</td>
<td>Point estimate (estimated value of survival)</td>
</tr>
<tr>
<td>LCL</td>
<td>Lower confidence limit</td>
</tr>
<tr>
<td>UCL</td>
<td>Upper confidence limit</td>
</tr>
</tbody>
</table>

Another influence on the outcome is whether a standard is required to be met once, or over consecutive years, in order to satisfy the licensing conditions. Regulators may require multiple years (usually sequential) to ensure that year to year environmental variation is captured in the assessment process. Operators may, therefore, need to pass the standard for three consecutive years (Figure 3). Even when the probability of meeting the standard is high, the probability of passing it several years in a row can be relatively low (e.g., a project with a 90% chance of passing a standard in a single year may have a 73% chance of passing it 3 years in a row). As a result, the number of consecutive years a standard must be passed lowers the sensitivity of the chance of passing it 3 years in a row. As a result, the number of consecutive years a standard must be passed lowers the sensitivity of the licensing conditions. Regulators may require multiple years (usually sequential) to ensure that year to year environmental variation is captured in the assessment process. Operators may, therefore, need to pass the standard for three consecutive years (Figure 3). Even when the probability of meeting the standard is high, the probability of passing it several years in a row can be relatively low (e.g., a project with a 90% chance of passing a standard in a single year may have a 73% chance of passing it 3 years in a row). As a result, the number of consecutive years a standard must be passed lowers the sensitivity of the standard. Finally, the number of fish released and the probability of detecting the released fish has an effect on the probability of passing a standard (White & Burnham, 1999).

The decision of whether a licensee satisfies the standard in a year is based on a single assessment. True values of survival are unknown, and the CJS method works by maximizing the likelihood. Therefore, stochasticity affects the outcome. However, if repeated many times, survival estimates vary from reality in predictable ways, which underlines the importance of selecting an appropriate standard. If too permissive, regulators may approve conditions that are incongruent with conservation goals. If too strict, the threshold may be unattainable and be at risk of being considered an arbitrary and capricious threshold. Thus, it benefits all stakeholders to define biological goals and understand the probabilities of measuring true success and detect true failure.

We constructed a tool to evaluate the probabilities of passing a statistical standard given (a) an a priori true value of survival, (b) a defined method of assessment, and (c) defined study parameters (detection probability and sample size of the simulated population). By running a series of mark-recapture computer simulations, which modeled the studies that are often implemented by dam operators, we were able to explore the probabilities of passing a given standard. Each simulation was analyzed using the same analytical tools and methods used by operators (CJS) so that the implications of study design and data assessment can be compared. We applied these data in an interactive user interface (UI) to allow users to visualize the theoretical consequences of survival thresholds, the assessment technique, and study design. This tool is now available online: https://umainezlab.shinyapps.io/sims/.

FIGURE 2  The three main methods for assessing a standard survival performance threshold ($T$) arbitrarily set to 0.96 (dashed line) for demonstration: PE (point estimate), LCI (lower confidence interval), and UCI (upper confidence interval). Filled symbols indicate successful passing of the standard, while the open symbols indicate failure.

FIGURE 3  Decision tree for a 3-year survival assessment with $P$ (passing the standard), and $F$ (failure to meet the standard) for a single year. The final checkmark represents passing the standard 3 years in a row, in which case, additional project survival studies are not required.
By using a simulated river, we are able to manipulate S and D and compare them to the estimators (ϕ and p). Our simulated mark-recapture study is loosely informed by studies carried out in the Penobscot River, Maine (HDR Engineering, Inc, 2016; Normandeau Associates, 2017, 2018). In this modeling approach, survival is estimated through a stretch of river that contains a dam, and survival is measured in space, rather than time intervals. Therefore, time intervals are not incorporated, while measures of distance are irrelevant (i.e., different distances between receivers do not change estimates of interval mortality), and hence, we defined arbitrary units of distance for convenience. Our study was in a river divided into five length units, with a release station at unit 0, and receiver stations at units 1, 2, 3, 4, and 5. A dam was present between Units 2 and 3 (Figure 4). For each simulation, we defined N, S, and D. So, the individual probability of being observed at unit Ui (U_i) is determined by the probability of surviving to Ui and the probability of being detected at that unit:

\[
\text{Observation at } U_i = S_i \times D_i
\]

in which S_i represents the true survival probability (unknown in real life, but defined in simulations) for the ith interval, while D_i represents the real probability of being detected (also unknown in real life but defined in simulations) at the ith receiver station, given that the individual survived. During the simulations, to determine whether an individual had successfully survived to Ui, a Bernoulli trial was carried out with two potential outcomes: success (survival, defined by S_i), and failure (1-S_i). If the individual survived to Ui, then, it had a chance of being detected at that station (given by D_i). If an individual was detected at a station, then it was assigned a 1 (meaning it survived and was detected), otherwise, it was assigned a 0 (meaning either that the individual is dead, or that the individual survived but was not detected). Individuals that survived to Ui then have a chance to survive to U_{i+1} (based on S_i) and be detected. If the individual fails to survive to U_i, it has no possibility to survive or be detected at any further units. The individual simulation continues until the individual is either dead or has left the system (after the last unit). As a result, a detection history is obtained for the individual. This process is repeated N times, as N represents the number of individuals used in the assessment. Therefore, a detection history for each individual of the virtual population is obtained, with all individuals having an initial 1 (representing initial release), and the detection history having six events.

Three parameters were defined before each simulation was run (Table 2): N, D, and S_3 (true survival for the interval with the dam). Values for N were set between 50 and 700 (by increments of 10, a total of 66 values), and values for D were set between 0.84 and 0.99 (from 0.84 to 0.92 with steps of 0.02, and from 0.93 to 0.99 with steps of 0.01 for a total of 12 values). Note that set values of D were equal among all receiver stations within a simulation (D_1 = D_2 = D_3 = D_4 = D_5). Values of S_3 were between 0.70 and 0.99 (0.7 to 0.85 with steps of 0.025 and from 0.86 to 0.99 by steps of 0.01 for a total of 21 values). The values for survival for the other intervals (i.e., S_1, S_2, S_4, and S_5; the intervals with no dam) were set to 0.99. This value is representative of survival observed in short intervals with no dams (Stich et al., 2014; Stich, Bailey, et al., 2015; Stich, Zydlewski, Kocik, & Zydlewski, 2015). All the values used in the simulations were included after consulting with potential users and decision-makers (e.g., NOAA, USFWS, Penobscot Nation, and Maine

![FIGURE 4](image-url) Iterative process by which the database for each combination of parameters was populated, analyzed, and the probability of passing a standard was obtained. For each combination of parameters, 1,000 simulations were run, in which N fish were released, with probability of detection P, and probability of survival S_3. Each simulation then was analyzed using a CJS model framework. The estimated apparent survival ϕ and error structure of each simulation were stored, and the probability and distribution of an outcome could be obtained after assessment rules were defined

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Increments</th>
</tr>
</thead>
<tbody>
<tr>
<td>True survival (S)</td>
<td>0.7</td>
<td>0.85</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>0.86</td>
<td>0.99</td>
<td>0.01</td>
</tr>
<tr>
<td>Number of “released” individuals (N)</td>
<td>50</td>
<td>700</td>
<td>10</td>
</tr>
<tr>
<td>Probability of detection (D)</td>
<td>.84</td>
<td>.92</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>.93</td>
<td>.99</td>
<td>.01</td>
</tr>
</tbody>
</table>

Note: Simulations were run 1,000 times for each parameter values combinations, totaling a total of 16,632,000 simulations.
Department of Marine Resources). We note that more values can be easily incorporated into the online interactive application. Based on the defined inputs above, the total number of parameter combinations is 16,632. Thousand simulations were run for each parameter combination, for a total of 16,632,000 simulations to inform the generated plots. A histogram is produced of either the PE, the LCI or UCI results from all the simulations (1,000) for the chosen parameter combination. The survival analysis was performed in program MARK, which uses maximum likelihood estimates (White & Burnham, 1999), via the RMark package in program R (Laake, 2013) (R version 3.43; R Core Team, 2019). Because we are only interested in the survival through the interval with the dam, we recorded each estimate of survival for the third interval ($\phi_3$) and the error structure (originally obtained on the logit scale, then back-transformed and presented in a probabilistic scale). Then, each set of recorded results for a parameter combination can be tested using a set of assessment rules (Figure 4).

2.2 Methods for assessing if the standard was met

We used each of the three assessment approaches (PE, LCI, and UCI) to evaluate whether the estimates of survival would pass or fail to meet a defined statistical standard (T). In order to pass, $\phi_3$ must be greater than or equal to T for the PE method. For the LCI approach, the standard would be met if the lower confidence limit for $\phi_3$ either 75% (LCI$_{0.75}$) or 95% (LCI$_{0.95}$) is equal to or greater than T. Similarly, for the UCI approach, the standard would be met if the upper confidence limit for $\phi_3$ (either 75 or 95%) is equal to or greater than T. In practice, these methods may not be mutually exclusive and a standard may incorporate multiple components (e.g., requiring the point estimate to be at or above 0.95, and the LCI$_{0.75}$ to be at or above 0.9). Thus, each individual simulation is assessed and it either passes or fails the standard under a user-defined assessment method and a user-defined T. By running 1,000 simulations for each combination, we can assess the probability of meeting a user-defined standard (T) and user-defined parameters (S, D, and N) under the each of the assessment methods (e.g., PE or LCI$_{0.75}$).

3 UI AS DECISION-MAKING TOOL

3.1 Use of interactive display

In order to present this complex database and associated results, an interactive user interface was developed using Shiny (version 1.3.0; Chang, Cheng, Allaire, Xie, & McPherson, 2019) in program R (R Core Team, 2019; user interface webpage: umaine2lab.shinyapps.io/sims). It is anticipated that this River Survival Simulator will be updated to meet users’ expectations. This article was written for version 1.0.1 of the River Survival Simulator tool, and an updated log of changes will be available in the readme section of the UI. The UI was built after considering the input of potential users and may be upgraded to fit the needs of others. Using this River Survival Simulator tool, the user can analyze: (a) a single standard test and (b) a multiple standard test.

3.1.1 Single standard test

The first panel of the River Survival Simulator tool allows the user to see the results of a single standard method of assessment. The user can select the S, D, N, and T, and the standard, and explore a suite of generated plots. A histogram is produced of either the PE, the LCI or the UCI results from all the simulations (1,000) for the chosen...
The histogram distributions show survival assessments and are independent of T (with defined S, D, and N). However, the plot incorporates T by graying out the area of the histogram that would pass the standard under a selected assessment method (Figure 5a). The user can also see a bar plot that represents the probability of passing the chosen standard for each of the five criteria for the selected scenario (Figure 5b).

**FIGURE 6** Comparison of frequency histograms of 1,000 simulated releases to estimate survival at a river reach with a dam (\(\phi_3\)) using CJS analysis. Probabilities of passing the standard for a scenario with a real survival (S) of 0.96, and a standard (T) of 0.96 (indicated by red vertical line). Because S and T have the same value, the green line for S is not visible in these graphs. In the user interface, when S and T values are equal, a star and a message under the plot notify the user (not shown). Panels (a–c) show results for a scenario with an N of 50, and a D of 0.84, while panels (d–f) depict a scenario with an N of 700, and a P of 0.99. Panel (a) and (d) show distribution of the point estimate (PE), while (b) and (e) demonstrate the distribution of the upper 75% confidence interval values (UCI\(_{0.75}\)). Panels and (c) and (f) are bar charts showing the probability of passing the standard for each of the methods of assessment described in the text: PE (point estimate), LCI\(_{0.75}\) and LCI\(_{0.95}\) (lower confidence interval at 75 and 95%), and UCI\(_{0.75}\) and UCI\(_{0.95}\) (upper confidence interval at 75 and 95%).

**FIGURE 7** Probability of passing a standard of 0.96 for each of the methods of assessment described in the text: PE (point estimate), LCI\(_{0.75}\) and LCI\(_{0.95}\) (lower confidence interval at 75 and 95%), and UCI\(_{0.75}\) and UCI\(_{0.95}\) (upper confidence interval at 75 and 95%) where a true survival of 0.91, and different numbers of individuals (N) are released under different detection probability(D) scenarios: (a) D = 0.88, N = 50, (b) D = 0.99, N = 50, (c) D = 0.88, N = 700, (d) D = 0.99, N = 700. As a higher N greatly reduces the confidence interval sizes, the increase in N had the highest response on the probability of passing the standard.
This single test allows the user to explore the sensitivity and specificity of different T values and methods of assessment. For example, a T of 0.96 is being applied to several projects under FERC (Black Bear Hydro, 2013; FERC, 2018). Therefore, for demonstration purposes, we will assume 0.96 is the biological standard (again, this value represents the biological goal) and the statistical standard (the value that will actually be tested) will also be 0.96 (but this can be set to any value). Using the interactive display, it is possible to explore the probability of passing the standard given under each scenario. When S is 0.96 (i.e., the biological goal is exactly met), the sensitivity (probability of passing given that the biological goal is met) depends on the method of assessment, the number of individuals used in the study (N), and the probability of detection (D). For example, when N and D are both low (e.g., N = 50 and D = 0.84), the probability of passing the standard for the point estimate method of assessment is 0.6, while the probability of passing the standard for the UCl0.75 is higher than 0.9 (Figure 6). However, if both N and D are increased (e.g., D = 0.99, and N = 700), the probability of passing the standard changes dramatically for the PE (decreasing to 0.5) but only a modest change for UCl0.75 (just below 0.9; Figure 6f).

The effects of the method of assessment, N, and D are even more evident when S is lower (keeping T at 0.96). When S is defined as 0.91, the specificity (probability of correctly ascribing failure) depends on the method of assessment and on N and D. For example, with a low D (0.88) and low N (50), the probability of passing the standard (i.e., a false positive) using the PE method is 0.20 (thus a specificity of 0.80). With these same parameters, but using UCl0.75 assessments, the probability of incorrectly passing is higher than 0.5 (resulting in an
We suggest the utility in exploring the conceptual difference between a biological goal (based on conservation objectives) and a statistical standard. An effective statistical standard that maximizes sensitivity and specificity may differ from the biological goal. A highly sensitive and specific standard may be perceived as fair by all stakeholders, as it is highly achievable when the conservation goal is truly being met, but difficult to achieve when the conservation goal is truly not being met. This tool provides a user interface that can aid in the design of studies, as it shows the importance of N, and D in providing estimates that reflect reality. Managers and operators have a mutual interest in providing an unbiased approach. Understanding how false positives and false negatives can be generated may inform statistical guidelines. Furthermore, this tool can be adapted or used “as is” for assessment of passing through different structures that impede movement and affect survival, such as culverts.

We caution that although setting clear survival goals is important to avoid negative consequences to populations, it is necessary to acknowledge that dams influence the success of migrating fish in complex ways (Caudill et al., 2007; Stich, Bailey, et al., 2015). Dams can cause delays (Ferguson, Absolon, Carlson, & Sandford, 2006), and nonlethal injuries to fish that can decrease the probability of survival later in their migration (Zydlewski, Zydlewski, & Danner, 2010). Survival is a critical component of fish passage regulations at dams, but is not the only performance metric observed to influence population viability.

ACKNOWLEDGMENTS

In-kind support was provided by the U.S. Geological Survey, Maine Cooperative Fish and Wildlife Research Unit. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Open Science Framework at https://doi.org/10.17605/OSF.IO/KBH4S.

ORCID

Alejandro Molina-Moctezuma https://orcid.org/0000-0002-9782-6460

REFERENCES

- Black Bear Hydro. (2013). Revised species protection plan and revised atlantic salmon passage study plan for project nos. 2710, 2712, 2534, 2600, and 2666 (Orono, Stillwater, Milford, West Enfield, and Medway hydroelectric projects). Retrieved from https://elibrary.ferc.gov/idmws/common/OpenNat.asp?fileID=13349181
HDR Engineering, Inc. (2016). Fish passage study report, west enfield project FERC no. 2600.

Data Citation
dataset. Molina-Moctezuma, A.; Zydlewski J.; 2019; Interactive UI Data; allist12222018.rds; https://doi.org/10.17605/OSF.IO/KBH4S