

## Supplementary Materials

### Rationale for considering single-species-sessions as suspicious.

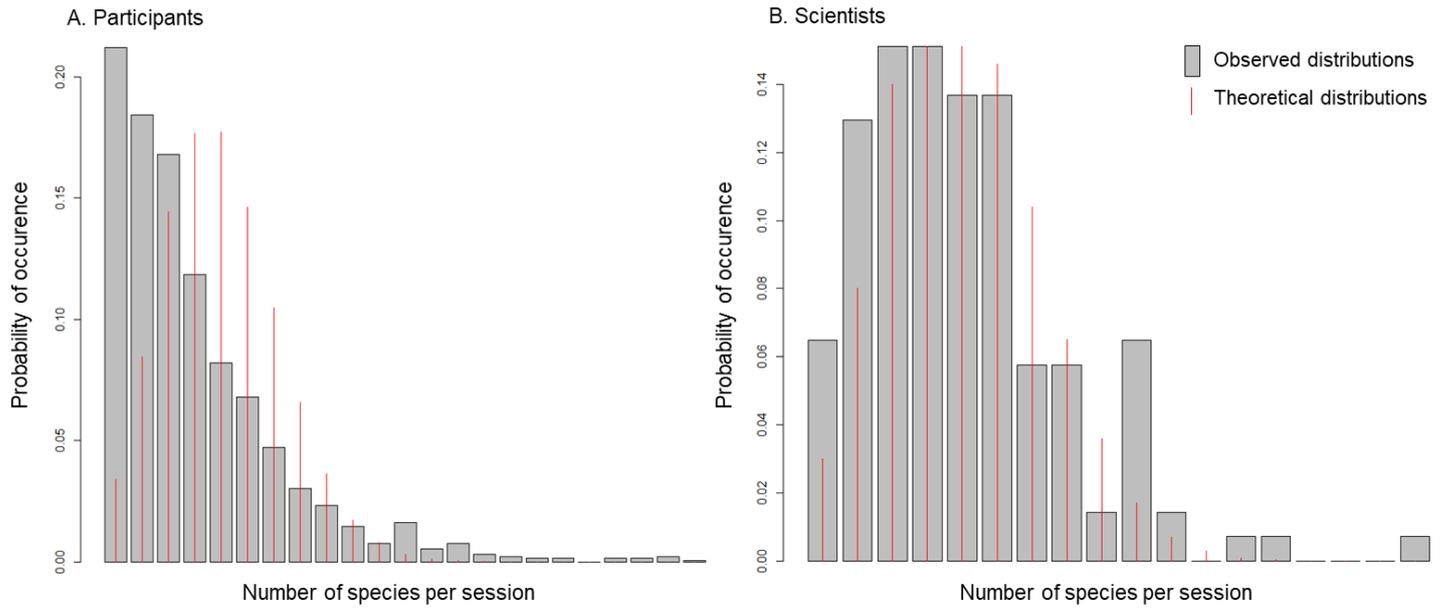
In our manuscript, we assumed that a session containing only a single insect picture was the result of a violation of the protocol. Indeed, regarding our knowledge of the Korean and French ecosystems, the personal experiences of experts of the National Museum of Natural History of Paris and Ewha Womans University, and already published literature (Deguines et al., 2016), we estimated that the probability of observing only one species during a period of 15 or 20 minutes is very low and can instead be used as an indicator of a violation of the protocol.

However, we acknowledge that among the single-species sessions, there are a few sessions for which there truly was only one species observed during 15 or 20 minutes. While we are not able to sort out which such sessions arose from a protocol violation and which did not, we below show that participants uploaded single-species sessions far more frequently than expected from the protocol, supporting our choice to consider them as ‘suspicious’ (i.e. violating the protocol).

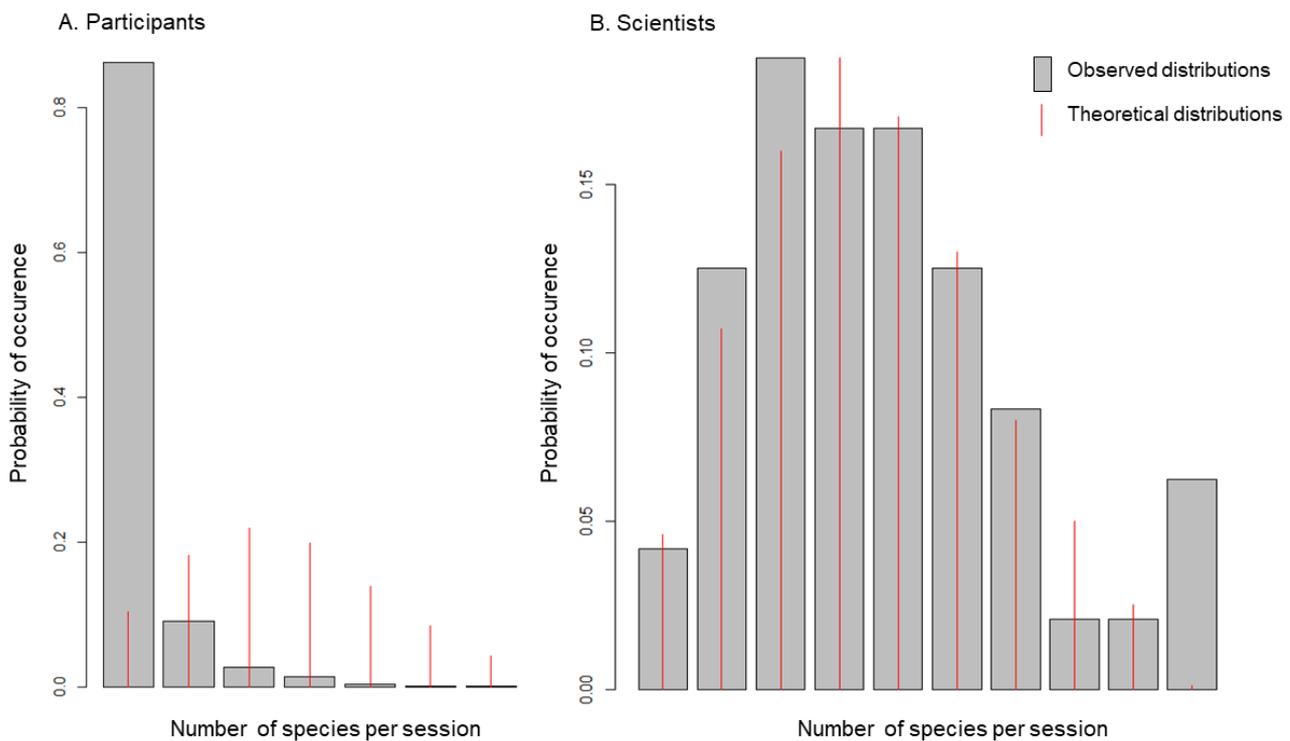
To know if the frequency of single-species sessions were over represented in the participants, we compared, for both programs, how the distributions of the number of species observed by sessions were fitting with the theoretical distributions that these data should follow, i.e. a zero-truncated Poisson law (zero-truncated because the websites would not allow uploading empty sessions, with no insect pictures). Additionally, we carried out these same comparisons using sessions done by scientists of each program for which we were sure that the sessions were strictly following the protocol. A theoretical zero-truncated Poisson law was generated using the package ExtraDistrib of R (version 1.0.143) and “rtpois” function. The  $\lambda$  (lambda) factor for the theoretical distributions were chosen according to session richness means computed over sessions observed by scientists: 5 species per session for the Spipoll (139 sessions) and 3.6 species per sessions for the K-Spipoll (48 sessions). Figures S1 and S2 illustrate the observed and theoretical distributions for sessions done by participants (S1A, S2A) and scientists (S1B, S2B) of Spipoll and K-Spipoll respectively.

A graphical assessments of figures S1 and S2 strongly suggests that, for both programs, the probability of single-species observations was over-represented in the sessions done by the participants compared to what it should theoretically be. For Spipoll, the probability of observing one species per sessions is reaching 21.2 % in the participant’s data even though the theoretical one should be 3.4 %. For K-Spipoll, these probability is reaching 86 % against 10.4 % for the theoretical one. Conversely, when scientists carried out the protocol, single-insect sessions were not over-represented.

This confirms that the vast majority of single-insect sessions made by participants likely was the result of a violation of the protocol.



**Figure S1.** Observed and theoretical distributions of the number of species per session for the Spipoll program. Panels A and B represent the distribution of species number in the sessions done by the participants and the scientists respectively.



**Figure S2.** Observed and theoretical distributions of the number of species per session for the K-Spipoll program. Panels A and B represent the distribution of species number in the sessions done by the participants and the scientists respectively.