Supplementary material B

Research Methods Details

Data collection

We observed 15 BioBlitzes led by The Natural History Museum in London, the California Academy of Sciences (San Francisco) and the Natural History Museum of Los Angeles County, with each NHM organising 4-7 events (for further information on the BioBlitzes see Supplementary material A). We used ethnographic field observations [Emerson et al., 1995] to capture participation data for 81 youth. The sampling method combined random purposeful and stratified purposeful sampling [Creswell, 2014]. We used a random sequence of numbers and then assigned each youth a number by counting them as they gathered at the start of the event or activity to decide which focal youth to observe on the day and in which order. We then checked if the selected youth were representative of the wider pool of youth at the event (e.g. age group, gender, ethnicity) and, if reasonable, skipped or substituted a given youth in the order to increase the range of observed subgroups. All demographic information (Table 1) are suppositions and were unconfirmed, as data was observational only, per ethics approval.

Table 1: Approximate age and presumed gender of the 81 focal youth (31 NHM London, 28 NHM Los Angeles County, 22 California Academy of Sciences) observed.

Gender	5-10 years (elementary school-age)	11-15 years (middle school-age)	15-19 years (high school-age)	Total
Female	20	11	9	40
Male	23	9	8	40
Unknown	1	0	0	1
Total	44	20	17	81

Our aim was to gather a variety of youth's BioBlitz experiences to get a broad picture of what youth do when attending BioBlitzes. We opted for focal sampling, observing and documenting all of the activities of the youth including conversations with others, solitary activities, behaviours and affects (such as boredom and excitement) for a particular individual during the entire sampling period [Altmann, 1974[1]; Pellegrini et al., 2004]. The observer followed the youth throughout the setting, so the youth could roam freely and decide in what activities they want to engage in (e.g. watch a puppet theatre or take part in pond dipping). To increase the number of focal youth that could be observed per event, we predetermined the duration of the observation interval. Taking into account how long the average CS data collection task would last and allowing time for youth to orient themselves within the setting and activity, we set the observation duration to 20 minutes. We are aware that in any one youth's participation, our study may have missed instances of activity that may have happened outside the observation interval. This was an unavoidable limitation unless we narrowed the number and range of youth observed in order to follow focal youth for the whole duration of the event.

[1] Altmann (1974) reports on two of her animal studies in which they found their observation capacity level when working with two observers together to be at 15 minutes when tracking non-social behaviour and its duration, and up to 40 minutes for studying social interaction but not tracking their duration.

Occasionally participation ended earlier or we had a chance to observe the same young person more than once, so actual observation times varied from 5-60 minutes with an average of 22.3 minutes. Researchers primarily took the approach of "observer as participant" [Creswell, 2014]. The observers were two postdoctoral researchers (one in the US and one in the UK) who were trained, experienced educational researchers and three project coordination officers (PCOs) who had a practitioner background (one at each of the NHMs) and were trained by the post-docs in both the theories and practice of qualitative research methods including taking ethnographic fieldnotes (Emerson, 1995), during this study. We had 1-2 observers (PCOs collected data on all BioBlitzes at their NHM, postdocs collected data at 4 BioBlitzes in total, two in the US and two in the UK) present at each of the events which resulted in observations of 2 to 14 focal youth per event, depending on the number of observers, the duration of the event and the number of young people attending. Our observation protocol ensured that methods were aligned across different settings and observers [Emerson et al., 1995].

Data analysis

To analyze the ethnographic fieldnotes, we adapted Creswell's steps of Data Analysis for Qualitative Research [2014]. We identified key action/ interaction episodes "key action/ interaction episode", defined a sequence of actions (when a young person is doing something by themselves) and/or interactions (when a young person is interacting with another person), that we determined to be evidence of when the young person was engaging in some way with science-related activities. Preparatory activities for citizen science activities are also considered relevant. For each key action/interaction episode, a researcher wrote a memo containing a claim about youth participation, a description of the type of participation observed and excerpts from fieldnotes as evidence to support the claim. At the start of the process, five researchers went through a calibration phase discussing and aligning the analytic approach in peer-review before progressing with data analysis. Working with a pool of 12 observations (15% of the data), each researcher wrote memos on their own observations and were assigned fieldnotes from the other writers. The memos written by different researchers about the same observation fieldnotes were then compared for discrepancies. After discussion of disagreements and aligning the process, the rest of the observation fieldnotes were analysed. The memo-writing process for each focal youth was iterative and involved at least two researchers per memo, involving a PCO and a postdoc working on each memo and ensuring that an observer who attended the event and an additional person agreed on the analysis.

Based on the scientific goal of the citizen science activity, we identified five essential steps to create a biological record [Pocock et al., 2015]. Therefore, we developed a priori codes for the different ways youth could potentially participate during the BioBlitz. Making nature observations ("What") in a habitat (Where) involves a person ("Who") finding organisms at a certain time ("When") [Isaac & Pocock, 2015], observable in the following types of participation:

- Exploring - Exploring nature to discover organisms, actively searching for organisms potentially involving tools such as binoculars or nets.
- Observing - Observing organisms in nature, using one's senses to find and study organisms

Figuring out the "What" aspect of an observation is a crucial piece of information for the biological record [Isaac & Pocock, 2015], we define this as:

- Identifying organisms - in the sense of finding out what organism (e.g. taxon or species) was observed

Through generation of evidence and making data accessible to others, the data about observed wildlife becomes a data point that can be used for research and monitoring. We consider these two types of participation separately, as not everything that is documented is necessarily shared:

- Documenting - Documenting the observations by generating evidence of the observation, such as a photograph or writing on a datasheet
- Recording - in the sense of making the documented observation available for biodiversity monitoring or research purposes, ideally providing the Who, When, Where and What aspects of a biological record [Isaac & Pocock, 2015]

More detailed and nuanced descriptions, as well as examples for the observed types of participation, are provided in the main article as well as additional types of participation that we discovered during the analysis. We used these types of participation for thematic coding [Saldaña, 2009] of the memos to enable a frequency analysis of participation types as well as the generation of a participation profile for each focal youth.

Two researchers, who both had been involved in all other data collection and data analysis steps, coded a sample (4% of the memos) including memos from each museum independently and achieved a substantial agreement (92% agreement, Cohen's κ = 0.73). The coders discussed disagreements and settled on coding-decisions for these cases [Creswell, 2014; Saldaña, 2009]. The coders then split up the remaining memos between them and coded them independently.

We report the frequency of focal youth that engaged in each type of participation and checked for differences in relation to age and gender based on those numbers (Table 5 & 6).

Table 5: Gender of our focal youth and types of participation (n(female)= 40, n(male)= 40, n(unknown)= 1).

Type of Participation	Female	Male	Unknown
Exploration	25	22	1
Observing	31	25	1
Identifying Organisms	24	21	0
Documenting	16	12	0
Recording	6	6	0

Table 6: Age groups of our focal youth and types of participation (n(elementary school)= 44, n(middle school)= 20, n(high school)= 17).

Type of Participation	5-10 years (elementary school-age)	11-15 years (middle school-age)	15-19 years (high school-age)
Exploration	30	14	4
Observing	31	16	10
Identifying Organisms	25	12	8
Documenting	14	6	8
Recording	3	3	6

Based on the results regarding the types of participations observed for youth, in the next step clusters were defined to propose participation profiles. The detailed descriptions and frequency of occurrence can be found in the main article.

References

Altmann, J. (1974). Observational Study of Behavior: Sampling Methods. *Behaviour*, 49 (3/4), 227-267.

Creswell, J. W. (2014). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. 4th ed. Thousand Oaks, CA: Sage Publications.

Emerson, R., Fretz, R. and Shaw, L. (1995). Writing Ethnographic Fieldnotes. Chicago: University of Chicago Press.

Isaac, N. J. B. and Pocock, M. J. O. (2015). 'Bias and information in biological records'. *Biological Journal Of The Linnean Society* 115 (3), pp. 522-531.

Pellegrini, A. D., Symons, F. J. and Hoch, J. (2004). Observing children in the natural worlds: A methodological primer. 2nd ed. Mahwah, NJ: Erlbaum.

Pocock, M. J. O., Roy, H. E., Preston, C. D., & Roy, D. B. (2015). 'The Biological Records Centre: A pioneer of citizen science'. *Biological Journal of the Linnean Society* 115 (3), pp. 475–493.

Saldaña, J. (2009). The Coding Manual for Qualitative Researchers. London: SAGE.