How deep do we dig?
Formal explanations as placeholders for inherent explanations

Susan A. Gelman[1]*, Andrei Cimpian[2]*, and Steven O. Roberts[3]

[1]University of Michigan
[2]New York University
[3]Stanford University

*Email: gelman@umich.edu (S. Gelman) and andrei.cimpian@nyu.edu (A. Cimpian)

SUPPLEMENTARY EXPERIMENT: FORMAL VS. NON-FORMAL EXPLANATIONS

The goal of the supplementary experiment is to test just how (un)satisfying formal explanations are. In Experiment 1, we showed that formal explanations are more natural-sounding than they are satisfying. However, without a comparison point, this result is difficult to interpret. In principle, people might find any simple, one-sentence explanation for complex properties such as having four legs to be unsatisfying. To address this point, in the current study, we compared formal explanations with non-formal (causal and teleological) explanations that we purposely selected to elicit high levels of agreement—which would then provide a yardstick for assessing how satisfying formal explanations are. Each question included the category label in order to focus on the explanatory value of the formal or non-formal explanation per se, independent of the informational value of identifying the category (which otherwise would differ for formal vs. non-formal explanations).

Method

Participants

Seventy-seven MTurk workers participated: 43 men, 34 women; M_{age} = 34.7 years, range 18-66; 60 White/Caucasian, 5 Black/African-American, 3 Asian/Asian-American, 4

Latino/Hispanic, 4 Multiracial, and 1 Other. None of the participants in this supplemental experiment had participated in the studies described in the main text.

Design

The study had a 2 (Explanation: Formal, Non-Formal) x 3 (Domain: Living kind, Artifact, Social) design, with Explanation and Domain as within-subject factors.

Materials and Procedure

The items were identical to those in Experiment 1 in the main text.

For each item, participants read a question (e.g., Abi asked, “Why are raincoats waterproof?”) followed by two different explanations, each provided by a different person. Each item paired a formal explanation (e.g., Person 1 replied, “Because they are raincoats.”) with a non-formal explanation (e.g., Person 2 replied, “Because they wouldn’t function otherwise.”). We chose the non-formal explanations in this study to be intuitively satisfying and thus to provide a rough upper bound of satisfyingness with which to compare the formal explanations. The order in which the formal and non-formal explanations appeared was randomized within participants.

We varied the non-formal explanations to fit the domain, and to provide two different options per item. Furthermore, they were designed to be roughly the same length as the formal explanations. Accordingly, for the living kinds, we included the following two non-formal explanations: “because they evolved that way” and “because of their DNA.” For the artifacts, we included the following two non-formal explanations: “because they wouldn’t function otherwise” and “because they were designed that way.” For the social categories, we included the following two non-formal explanations: “because they were trained to” and “because they want to.” Which non-formal explanation a participant received was randomized.

For each item, after reading both explanations, participants judged each of the two explanations for how satisfying it was, on a 7-point scale (1 = not at all satisfying; 7 = extremely satisfying). The instructions for satisfyingness ratings were the same as in Experiment 1.

Results and Discussion

We conducted a multilevel mixed-effects linear model with cross-classified random intercepts for participants and items. The predictors were Explanation (a within-subject [level-1], dichotomous variable), Domain (a within-subject [level-1], three-level categorical variable), and their interaction.

We obtained a main effect of Explanation ($M_{\text{formal}} = 2.81$ vs. $M_{\text{non-formal}} = 4.39$ on a 1–7 scale), Wald $\chi^2 = 917.99$, $p < .001$. As expected, formal explanations were far less satisfying than non-formal ones (see Table 4). We also observed an Explanation x Domain interaction, Wald $\chi^2 = 173.68$, $p < .001$. Importantly, however, simple effects tests confirmed that the difference in satisfyingness ratings between formal and non-formal explanations was significant in each of the category types tested separately, *p*s < .001 (see Table 4 below).[1] The main effect of Domain was non-significant, *p* = .14.

Table 4

Mean Satisfyingness Ratings as a Function of Domain and Explanation. SDs are in parentheses.

	Living Kind	Artifact	Social
Explanation:			
Formal	2.73 (1.92)	2.46 (1.77)	3.24 (2.01)
Non-Formal	4.30 (1.89)	4.88 (1.74)	3.98 (1.96)

[1] We also conducted post-hoc *t* tests to test, within each domain, how each causal explanation (e.g., for living kinds: “because they evolved that way,” “because of their DNA”) compared to formal explanations. For all six comparisons, the non-formal explanations received higher ratings than the corresponding formal explanations, *t*s(76) ranging from 2.80 to 13.79, *p*s ≤ .006.

These data supplement the findings from Experiment 1, which demonstrated that formal explanations are relatively less satisfying than they are natural-sounding, and further established that they are less satisfying than well-chosen non-formal explanations. Thus, these studies converge to provide evidence that formal explanations are not very satisfying, consistent with the hypothesis that for natural kinds, they serve as a pit stop for more detailed (and, as we show in the experiments reported in the main text, more satisfying) inherent explanations.