Supplemental Experiment 1

Xu and Tenenbaum (2007a) always presented the one example trials in the first trial block. This raises the possibility that performance on these trials might somehow impact performance on later trials with three subordinate examples. In Experiments 2 and 3, we opted to eliminate this possibility by presenting the three subordinate example trials first. Before making this change, however, we first determined whether the suspicious coincidence can be replicated when the order of trials is modified. Thus, this experiment was identical to Experiment 1 in all respects except that the three subordinate examples were presented during the first trial block. The remaining trial blocks were randomly ordered.

Participants. There were 19 adult participants. Participants received credit in a college course or monetary compensation for their participation. All participants reported normal vision and gave informed consent.

Results and Discussion. Participants generalized at the basic level at roughly the same percentage when they saw one example ($M = 24.56$) versus three subordinate examples ($M = 16.66$), $t(18)=1.14$, $p = .27$, two-tailed. When compared to performance in Experiment 1, this yielded a noticeable drop in basic-level generalization in the one example condition. A direct comparison across experiments showed comparable performance in the three subordinate condition, $t(36)= -.75$, $p =.46$, two-tailed, but a significant reduction in basic-level generalization across experiments in the one example condition, $t(36)=2.00$, $p = .05$, two-tailed. Thus, trial order has an impact on how participants generalize when presented with a single example. When the one example trials are presented first in the study (Experiment 1), participants generalize broadly. When these trials occur after participants have been shown three subordinate examples (Experiment S1), however, they generalize more narrowly.

Although the present findings suggest that there are factors that matter beyond the suspicious coincidence in this task, the key result still stands—participants generalized narrowly when presented with three subordinate examples. This is notable given our evidence that trial order matters for the one example condition. Even though we manipulated when participants saw the three subordinate examples in the present experiment, they showed the same response pattern. Given that trial order had little impact on the suspicious coincidence, Experiments 2 and 3 presented the three subordinate examples in the first block of trials. This is the most experimentally conservative choice because it ensures that there is no carry over from other trial blocks that might impact the central phenomenon of narrow generalization given three examples.

Supplemental Experiment 2

Experiments 1 and S1 establish the robustness of the suspicious coincidence in the three subordinate examples condition. Experiment S2 replicates Experiment S1 with one small change. In the original method, the generalization test items are in view when the exemplar is presented. Because this presentation method could interfere with the sequential presentations used in Experiment 2 and 3 and limit attention to the exemplars (i.e., participants may not fully study the exemplar set before deciding about generalization items), Experiment S2 presented the exemplar display at the bottom of the computer screen for 6 s of study time. Then, the full generalization set was presented along with the exemplars as before. All other experimental details were identical to Experiment S1.

Participants. There were 19 adult participants. Participants received credit in a college course or monetary compensation for their participation. All participants reported normal vision and gave informed consent.

Results. Results replicated the findings from Experiment S1. Participants generalized at the basic level with comparable percentages in the one example ($M = 15.79$) and three subordinate examples conditions ($M = 11.40$), $t(18)=.61$, $p = .55$, two-tailed. Comparisons across Experiments S1 and S2 for each condition separately revealed no significant differences, $t(36)=.95$, $p = .35$, two-tailed (one example); $t(36)=.65$, $p=.52$ (three subordinate examples). Moreover, as in Experiment S1, participants generalized at the basic-level less often in the one example condition relative to performance in Experiment 1, $t(36)=2.99$, $p <.01$, two-tailed. These data demonstrate that there were no differences in performance when we presented the exemplar display first, followed by the generalization set.