Supplemental Materials for "Learning absolute meaning from variable exemplars"

1. Additional analyses of participants' responses in Experiments 1 and 2

As discussed in Section 2.4, our analysis mainly focused on likelihoods with which participants chose the "Neither" option in the Test Phase. It is a diagnostic tool to assess whether they have learned *pelty* to mean "maximally tight-fitting". (This task was modeled after Syrett et al., (2010)'s investigation into children's and adults' interpretations of English absolute gradable adjectives). We expected that, if participants' semantic representation of *pelty* includes the maximum standard of comparison (i.e., maximally tight-fitting), they should be more likely to choose the "Neither" option in Ambiguous test trials. It is, however, also important to evaluate how reliably participants chose the tighter option in the Ambiguous and Unambiguous test trials. In particular, their selection of the tighter option in an Unambiguous trial critically corroborates that they are indeed associating the meaning of *pelty* with the maximum standard of comparison: when presented with a maximally tight-fitting option and a less tight-fitting option, participants should choose the former as a referent of *pelty*.

To assess participants' selection of the tighter-fitting option, we constructed mixed effect regression models similar to what we describe in result sections (Section 2.4). As in the model assessing choices of the "Neither" option, the analysis was conducted using the *glmer* function of the *lme4* package (Bates, Maechler, Bolker & Walker, 2015) in R (R Core Team, 2016). The dependent measure was participants' responses coded as a: "Tighter-fitting" = 1 and all others = 0. The model included Trial types (Unambiguous vs. Ambiguous), Explanation manipulation (1 = With-Explanation vs. 0 = No-Explanation), and Experiments as fixed effects as well as by-participant and by-item (at Test) as random effects.

Table 1. Model fixed and random effects predicting "tighter-fitting" responses for both Experiments 1 and 2.

Experiment 1 and 2: Fixed Effects				
	β	Std. Error	z-value	p-value
Intercept	0.49	0.15	3.30	**< 0.001**
With-Explanation Condition	0.44	0.15	2.93	**0.003**
Trial (Unambiguous vs. Ambiguous)	1.21	0.14	8.46	**< 2.00e-16**
Experiment (1 or 2)	0.22	0.15	1.51	0.13
With-Explanation Condition x Trial	0.71	0.13	5.37	**7.78e-08**
With-Explanation Condition x Experiment	-0.13	0.15	-0.87	0.38
Trial x Experiment	0.16	0.13	1.24	0.21
With-Explanation Condition x Trial x Experiment	0.11	0.13	0.82	0.41
Experiment 1 and 2: Random Effects				

Groups		Variance	Std. Dev.	
Subject	(Intercept)	7.20e-01	8.49e-01	
Item	(Intercept)	1.56e-16	1.25e-08	
Observations: 504, Subjects: 126, Item Type: 2				

The results of the analysis are summarized in Table 1. They revealed the main effects of Explanation condition (With- vs. No-Explanation, $\beta = .44$, $z = 2.93$, $p < .01$) and Trial types (Unambiguous vs. Ambiguous, $\beta = 1.21$, $z = 8.46$, $p < .001$) such that the tighter-fitting option was more likely to be chosen: 1) when the contextual explanations were provided in Exposure; and 2) when there is an unambiguously tight-fit option was available in a test trial. The interaction between Trials and Explanation manipulation was also significant ($\beta = .71$, $z = .13$, $p < .001$). Participants were more likely to pick the tighter-fitting option in an Unambiguous test trial when in the With-Explanation condition. Notably, none the following terms including Experiment was significant:

- main effect of Experiment (Experiment 1 vs. 2),
- the two-way interaction between the Explanation conditions and Experiment,
- the two-way interaction between the Trial types and Experiment
- the three-way interaction between Explanation conditions, Trial types and Experiment

This suggests that participants' choice of the tighter-fitting option was largely intact when the number of absolute exemplars was halved (i.e., Experiment 2).

Together with the choice patterns for the "Neither" option, these results likely indicate that, in both experiments, participants in the With-Explanation condition inferred a more stringent meaning for the novel adjective: for an object to be *pelty* it has to be maximally tight-fitting (rather than sufficiently tight-fitting).

2. Follow-up experiment: Learning the meaning of *pelty* with no absolute exemplars

We conducted a follow-up experiment to test if participants can infer the maximum standard of comparison for the novel word (*pelty*) exclusively from non-absolute exemplars. We replicated Experiment 2 while replacing all the visually unambiguous (maximally tight-fitting) visual examples with visually ambiguous ones. As a result, all the participants in this experiment watched 24 visually ambiguous exemplars. Half of them were labeled as "*pelty*" and the other half were labeled as "not *pelty*". As in Experiments 1 and 2, we used a between subject manipulation and assigned participants randomly to With-Explanation and No-Explanation conditions.

2.1 Participants

99 subjects were recruited to participate from Amazon Mechanical Turk (https://www.mturk.com/mturk/). These participants all identified as monolingual, English-speaking adults. They were assigned randomly to one of two between-subject conditions; With-Explanation or No-Explanation. Due to poor catch trial performance (n=1) and past participation in similar experiments (n=25), we excluded 26 participants in total. All remaining 73 participants (34 in With-Explanation and 39 in the No-Explanation conditions) were included in our analyses.

2.2. Stimuli

In Exposure, participants saw a total of 24 ambiguously tight-fitting examples of *pelty*. These were constructed from six object types (T-shirt, bracelet, bookshelf, shoe, laptop case, *card*). Each object type was presented with two unique videos (totaling 12 unique videos). Each video was shown twice while paired with a different description as well as a label of *pelty* or not *pelty*.

The videos were identical across two between-subject conditions. However, in the With-Explanation Condition, participants received contextual explanations that described the intended use of the novel adjective as an absolute adjective (e.g., "I do a lot of hands-on work and don't want to worry about my bracelet falling off. This bracelet is *pelty*"). In the No-Explanation Condition, the accompanying audio did not provide such information, and instead described the visual scene or irrelevant information (e.g., "This is my sister's bracelet and she likes this very much. This bracelet is *pelty*." The make-up of videos used in the Exposure phase in Experiments 1 through 3 is summarized in Table 2.

Table 2. The number of example tokens used in the Exposure phase in Experiments 1 – 3.

	Visually unambiguous		Visually ambiguous		Total
	labeled as *pelty*	labeled as *not pelty*	labeled as *pelty*	labeled as *not pelty*	
Experiment 1	6	6	6	6	24
Experiment 2	3	3	9	9	24
Experiment 3	**0**	**0**	**12**	**12**	**24**

2.3 Procedure

The procedure was identical to Experiments 1 and 2.

2.4 Planned analysis

We sought to determine if the participants learn the meaning of *pelty* to include a maximum standard of comparison without any exposure to an absolute example. As in Experiment 1, we analyzed participants' choice patterns of "Neither" option in the Unambiguous as well as in the Ambiguous Test trials. To do so, we coded: "Neither" = 1 and all others = 0 as our dependent variable. We used mixed-effects logistic regression (Breslow & Clayton, 1993; Jaeger, 2008), applying the *glmer* function of the *lme4* package (Bates, Maechler, Bolker & Walker, 2015) in R (R Core Team, 2016). The model included Explanation manipulation (With-Explanation = 1 vs. No-Explanation = 0), Trial type (Unambiguous vs. Ambiguous) as fixed effects. We also random effects of by-participant and by-item (at Test).

2.4 Results

Participants were overall most likely to choose the tighter-option regardless of Trial types (i.e., Unambiguous and Ambiguous) and the between-subject Explanation conditions (i.e., With- and No-Explanation conditions) (Figure 1). One notable pattern in the responses in Experiment 3, in comparison to those in Experiments 1 and 2, was the increased proportion of the not-tighter option chosen both in the Unambiguous and in the Ambiguous trials. As a result, participants'

responses were distributed over the three options in Experiment 3, indicating a great deal of uncertainty and variability in the meaning representations inferred in the Exposure phase.
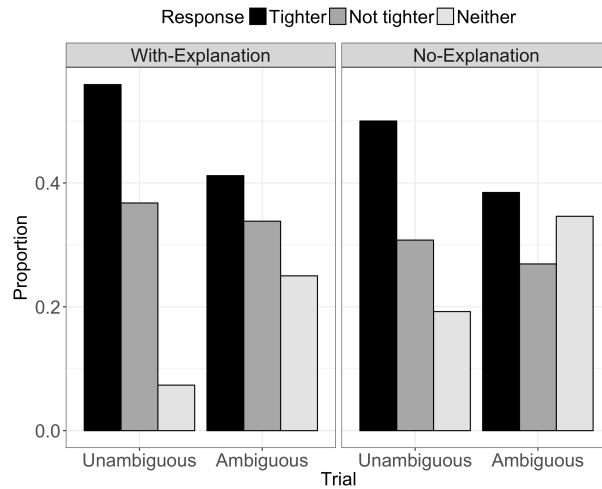

Figure 1. Proportion of responses at Test in the follow-up experiment (Experiment 3).

Our planned analysis on participants' choice of the "Neither" option revealed the main effects of explanation conditions (β = -0.45, z = -2.03, p < .05) and Trial types (β = 0.76, z = 4.60, p < .0001) (Table 3). Unlike in Experiments 1 and 2, however, the interaction between these two fixed effect variables was not significant. These results suggest that, even without any exposure to absolute exemplars, participants were capable of relating the meaning of *pelty* to the positive valency of tightness of fit. Moreover, at least some subjects seem to have been able to infer the maximum standard of comparison. The effectiveness of contextual explanations, however, significantly dissipated in this experiment as compared to Experiments 1 and 2. The presence of contextual explanations did not appear to have supported explaining away of visually ambiguous (non absolute) exemplars.

Table 3. Model fixed and random effects predicting "Neither" responses for Experiments 2-3.

Experiment 3: Fixed Effects				
	β	Std. Error	z-value	p-value
Intercept	-1.69	0.25	-6.77	**<1.32e-11**
With-Explanation Condition	-0.45	0.22	-2.03	**<.05**
Trial (Ambiguous)	0.63	0.18	3.56	**<.001**
With-Explanation Condition x Trial	0.16	0.17	0.9	0.36
Experiment 3: Random Effects				
Groups		**Variance**	**Std. Dev.**	
Subject	(Intercept)	0.99	0.99	

Item	(Intercept)	0	0	
Observations: 544, Subjects: 166, Item Type: 2				

The results of this follow-up experiment illuminate the importance of visually unambiguous (absolute) exemplars in inferring the maximum standard of comparison. In its absence, participants were unlikely to infer the absolute meaning (maximally tight-fitting) for the novel adjective introduced in the Exposure phase. Together with the results from Experiment 1 and 2, the current result suggests that explaining away comes to effect when some absolute exemplars are available to guide attention to an end point on a relevant scale.