Novel labels increase category coherence, but only when people have the goal to coordinate

Citation for published version:
Suffill, E, Branigan, H & Pickering, M 2019, 'Novel labels increase category coherence, but only when people have the goal to coordinate', *Cognitive Science*, vol. 43, no. 11, e12796. https://doi.org/10.1111/cogs.12796

Digital Object Identifier (DOI):
10.1111/cogs.12796

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Cognitive Science

COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 43 (2019) e12796

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12796

Novel Labels Increase Category Coherence, But Only When People Have the Goal to Coordinate

Ellise Suffill, Holly Branigan, Martin Pickering

Department of Psychology, University of Edinburgh

Received 10 October 2018; received in revised form 28 August 2019; accepted 24 September 2019

Abstract

From infancy, we recognize that labels denote category membership and help us to identify the critical features that objects within a category share. Labels not only reflect how we categorize, but also allow us to communicate and share categories with others. Given the special status of labels as markers of category membership, do novel labels (i.e., non-words) affect the way in which adults select dimensions for categorization in unsupervised settings? Additionally, is the purpose of this effect primarily coordinative (i.e., do labels promote shared understanding of how we categorize objects)? To address this, we conducted two experiments in which participants individually categorized images of mountains with or without novel labels, and with or without a goal of coordination, within a non-communicative paradigm. People who sorted items with novel labels had more similar categories than people who sorted without labels only when they were told that their categories should make sense to other people, and not otherwise. We argue that sorters' goals determine whether novel labels promote the development of socially coherent categories.

Keywords: Categorization; Coherence; Communicative goals; Coordination; Labeling

1. Introduction

From infancy, we recognize that labels denote category membership and help us to identify the critical features that objects within a category share. Labels not only reflect how we categorize, but also allow us to communicate and share our categories with others. Labels are thus closely tied to communication. Given the special status of labels as both an indicator of category membership and a coordinator of categories between people, how does the process of labeling affect the similarity of people's categories, and are any effects tied specifically to settings that have a *coordinative goal* (i.e., a goal in which people try to integrate their behavior with others)? To address this question, we

Correspondence should be sent to Ellise Suffill, Department of Psychology, University of Wisconsin-Madison, 1202 W Johnson St, Madison, WI 53706, USA. E-mail: suffill@wisc.edu

carried out two experiments that investigated the effects of novel labels on category coherence when acting with a coordinative goal (i.e., when participants were told that their categories should make sense to another person, and moreover that their categories would be compared to a partner's categories), and without a coordinative goal (i.e., when participants were told that their categories should make sense to themselves). If category coherence is a general consequence of labeling, then people who use novel labels should show greater category coherence than people who do not, regardless of the goal of sorting. But if the effects of labeling are specific to situations in which people have the goal of coordinating their categories with others' categories, then the labeling effect for category coherence should occur only in settings where people also have a coordinative goal.

1.1. Linguistic categorization and category coherence

People sort objects in similar ways when the categories they form are built upon shared features (i.e., features that are common to all sorters in a group), such as perceptual or functional information about the objects. Rosch and Mervis (1975) posited that intuitive category structure maximizes the within-category similarity of members while minimizing between-group similarity. To an extent, language relies upon this category coherence, because people need to have a similar understanding of the world—and how objects within it are labeled—in order to successfully communicate about the world. Children presumably learn this shared understanding through the feedback that they receive as they acquire language; for example if a child incorrectly labels a lemon as an orange, an adult can produce the correct label, "lemon," and so provide feedback about the underlying conventional category and its associated label, in a process of "supervised categorization" (Pothos, Edwards, & Perlman, 2011).

However, other work shows that categorization is also affected by language: People categorize objects differently when they use word labels to sort objects (*linguistic categorization*), compared with when they sort objects without labels (*non-linguistic categorization*). While these linguistic categories vary across speakers of different native languages (Malt, Sloman, Gennari, Shi, & Wang, 1999), they appear relatively consistent for native speakers of the same language (Laskowski, 2011). Therefore, language can increase the coherence of people's categories, when those people speak the same language and, so use a shared set of linguistic labels for the items they are sorting.

Labels can also change the way in which we perceive and remember objects, and this may contribute to the convergence of categories across native speakers of the same language. Lupyan (2008, 2012) proposed that existing labels might affect categorization by causing a shift in how people represent categories, by distorting which object features people successfully store in memory when categorizing and subsequently recalling objects. That is, when speakers apply a conventionalized word label to an object, it causes their representation of the object to become a mix of its idiosyncratic features and features that are typically associated with the relevant category denoted by the label. Lupyan argued that this representational shift supports categorization by helping people to select more generalizable dimensions that work well across a range of objects within the category.

Sorters are therefore able to avoid forming categories using too fine and too many dimensions that do not generalize well across a range of objects, and which would lead to an unhelpfully large number of categories. And when people use fewer, more generalizable dimensions, it becomes more likely that their categories will overlap and be similar to each other. And indeed, in "unsupervised categorization" (i.e., where people sort stimuli into self-determined categories without feedback), sorters tend to show a preference for simpler dimensional sorts (e.g., unidimensional; Ashby, Queller & Berretty, 1999), though this preference appears somewhat task and stimulus dependant (Pothos & Close, 2008; Pothos, Perlman, Bailey, Kurtz, Edwards, Hines, & McDonnell, 2011).

1.2. Effects of novel labels on category coherence

As well as existing word labels with conventionalized meanings, people also frequently encounter novel labels for new or even familiar objects, and such novel labels can also influence category learning in both infants and adults. Waxman and Markow (1995) argued that novel labels actively guide infants' attention to the relevant perceptual similarities across objects, in a way that promotes category learning. For example, non-word labels (e.g., "Look at the Timbo!") increasingly direct infants' attention toward shared object features (i.e., features that all objects within a category have) over dissimilar features, which in turn leads to enhanced category learning for novel objects (Althaus & Mareschal, 2014; Althaus & Plunkett, 2016).

There is also evidence that the presence of novel labels can draw adults' attention to perceptual features across objects, in a way that supports category learning. Lupyan and Casasanto (2014) had adults categorize novel "alien" stimuli into two pre-determined categories using existing word labels (i.e., "smooth" vs. "pointy") or non-word labels (i.e., "foove" vs. "crelch"). In this supervised categorization task, participants performed equally well in learning to label smooth-headed aliens as "smooth" or "foove," and equally well in learning to label pointy-headed aliens as "pointy" or "crelch," with participants in both of these conditions performing better then participants who categorized without any labels. In other words, they learnt to assign stimuli to the relevant categories equally well using existing or novel labels. However, the effects may depend on the specific labels used. For example, "foove" both sounds like "smooth" and may have a sound–symbolic relation to smooth objects (as a consequence of the vowel; Köhler, 1929), whereas "crelch" may have a sound–symbolic relation to pointy objects.

Similarly, Lupyan, Rakison, and McClelland (2007) investigated the efficacy of printed and spoken non-words as labels compared with non-linguistic, location-based cues as labels when learning to categorize novel "alien" stimuli as "friendly" or "unfriendly". Location-based cues involved the alien moving vertically onscreen to indicate "where it lived" (e.g., whether the alien came from the "friendly" part of the planet, or not). Both printed and spoken word labels facilitated category learning, but location-based cues did not. Lupyan et al. (2007) argued that the novel word labels (like existing labels) were effective because they simplified the distinction between the categories. That is, using word labels meant that participants could categorize objects under a single term (here, "leebish" or "grecious") that represented multiple category dimensions and also made these dimensions more concrete (and, so, easier to access in upcoming trials).

However, the majority of work examining the effects of novel labels on adult categorization (e.g., Barnhart et al., 2018; Lupyan & Casasanto, 2014; Lupyan et al., 2007) involves supervised categorization paradigms, in which people categorize objects using externally specified dimensions and receive feedback as they learn to do so. Pothos, Edwards, et al. (2011) argued that this feedback promotes representational flexibility in how people choose their dimensions for categorization; that is, the way in which people categorize can be changed over time through error-based feedback (Goldstone, 1994). In contrast, unsupervised categorization (e.g., Malt et al., 1999) is spontaneous and can be unpredictable in how people choose dimensions for categorization (Pothos & Chater, 2005). Pothos and Chater (2002) suggest that unsupervised categorization resembles perceptual organization, in that the categorization of novel objects is driven by the goal to simplify the similarity relations between objects in a category, and is not susceptible to the representational flexibility that is found in supervised settings as a consequence of feedback; Pothos, Edwards, et al., 2011).

1.3. Labels and coordination during unsupervised categorization

Despite the assumption that unsupervised categorization is not subject to representational flexibility, there is some evidence that labeling can enhance coherence between sorters in paired tasks involving unsupervised categorization (e.g., Markman & Makin, 1998; Suffill, Branigan, & Pickering, 2016). Therefore, the act of labeling may in itself produce changes in the way that people select dimensions for categorization and represent their categories (and, in doing so, increase the coherence of their categories). However, this evidence has come from studies involving a coordinative goal (i.e., involving tasks which promote the development of greater coherence between partners).

For example, Markman and Makin (1998) had participants use plastic building blocks to build a small structure (i.e., a car or a spaceship). Participants in a pair developed shared conventional labels for the different types of building block and used these labels to build the structure collaboratively. Pairs who sorted with shared labels developed greater category coherence than pairs who sorted without establishing shared labels, and pairs who built the structures individually (before being yoked for analysis). As such, the goal that participants had was coordinative (i.e., because the coordination of partners' categories and category labels was beneficial to their performance in the building task).

Similarly, Suffill et al. (2016) had pairs of participants sort morphed shapes with or without novel labels and with or without exposure to a partner. Regardless of exposure between partners, participants who sorted with novel labels became more similar to one another in how they categorized the objects, than those who sorted without labels. This result suggests that using novel labels can affect unsupervised adult categorization, and that it does so in a way that increases the coherence of partners' categories. However, the paired structure of the task again reinforced the need for coordination between partners. Thus, these results suggest that when people categorize spontaneously, their categories can be representationally flexible to some extent, but we cannot tell whether these effects are due to language, the goal of sorting or a combination of the two.

How, then, might such goals affect the way that people use novel labels to categorize? One possibility is that labels directly reorganize how people select dimensions for categorization in a way that makes them categorize more similarly to each other. In that case, people would tend to categorize in similar ways to each other whenever they used labels, compared to when they did not use labels, and this would occur irrespective of the goal they had. Alternatively, labels might affect how people select dimensions for categorization in a way that is tied specifically to the communication of those categories. That is, labels might act as a device for coordinating categories with others only when people sort with goals that require coordination (Clark, 1996). In that case, novel labels might increase category coherence specifically for goals that involve coordination between people.

In addition, as categories that would make sense to one partner are likely to also make sense to other people, using labels with a coordinative goal might lead people to develop greater category coherence not simply with one partner, but also with a group of people undertaking the same task (*social coherence*; see Garrod & Doherty, 1994, for evidence of how interaction with a partner can enhance coherence across a group of people). Under this account, labels might increase category coherence between people who sort with a coordinative goal even if they are not directly exposed to each other's categories, thus, leading to greater social coherence of the groups' categories.

1.4. Current study

We have seen that novel labels—like existing labels—can affect the way that people categorize, with evidence from both supervised (Barnhart et al., 2018; Lupyan & Casasanto, 2014; Lupyan et al., 2007) and unsupervised (Suffill et al., 2016) categorization. However, it is not clear whether the effects of novel labels on category coherence occur with any goal, or whether they are specific to goals by which there is a need for people to coordinate their categories with others. To test between these possibilities, we carried out two experiments, in which we investigated the effects of novel, non-word labels on category coherence during unsupervised categorization of real world items (i.e., grayscale images of mountains) when participants did or did not have the goal of forming categories that would make sense to another person. Specifically, we examined the effects of novel labels that lack both conventionalized meanings (i.e., within the mental lexicon; Jackendoff, 2002) and pre-defined associations to non-arbitrary features (e.g., such as the word form-shape associations in Lupyan & Casasanto, 2014). In Experiment 1, participants sorted with or without these novel labels, with a non-coordinative goal. In contrast, in Experiment 2, participants sorted with or without novel labels, but with a coordinative goal.

We predicted that novel labels would change the way that people sorted under at least some circumstances, by affecting which dimensions they selected for their categories, so that using labels would lead them to select more generalizable dimensions based on shared, perceptual features of the items and, therefore, to instantiate more socially coherent categories. Our interest was in the specific circumstances under which such effects would occur. We can identify three possible patterns of effects. First, if labels increase category coherence regardless of participants' beliefs about the purpose of sorting, then people who sorted with novel labels would show greater category coherence than people who sorted without labels, and crucially this effect would hold to the same extent in Experiments 1 and 2. Alternatively, if participants' beliefs about the purpose of sorting (i.e., having a coordinative context) additively increase a basic effect of labeling on coherence, then we would expect participants who sorted with novel labels to show greater category coherence than participants who sorted without labels in both experiments, but that people who sorted with novel labels in Experiment 2 (involving a coordinative context) would show greater coherence than people who sorted with novel labels in Experiment 1 (involving a non-coordinative context). Finally, if the effect of novel labels on category coherence depends on participants' beliefs about the purpose of sorting, such that labels affect the selection of category dimensions in a way that relates specifically to communicating those categories to others, then participants who sorted with novel labels would show greater category coherence than participants who sorted without labels in Experiment 2 only; there would be no difference between participants who sorted with and without labels in Experiment 1.

2. Experiment 1

2.1. Methods

2.1.1. Participants

There were 200 monolingual English speakers (121 female) from the USA and UK. Ages ranged from 18–35 years ($\bar{X} = 27.01$, $SD = 5.04$). Participants were recruited through Prolific Academic (https://www.prolific.ac.uk). The University of Edinburgh's Ethics Committee approved this study.

2.1.2. Stimuli

The stimuli comprised 72 greyscale images of mountains (see Fig. 1 for a subset). We chose images of mountains because they are real-world items that are unlikely to be recognized, and which vary on a relatively small number of salient dimensions (e.g., height, sharpness, number of peaks, presence of snow). We created three sets (A–C, see "Stimuli" in Data S1), each comprising 24 items (i.e., images). Participants sorted one set per round. Additionally, we generated 50 pairs of non-word labels with a consonant-vowel-consonant structure (e.g., "rah," "jib"; see "Labels" in Data S1). Each of these label pairs was used by only two participants within the experiment (and, as we explain in the results, these two participants were not specifically compared to each other in the analysis). We extensively varied the labels to ensure that any labeling effects could not be due to non-arbitrary characteristics of specific labels (e.g., word form–shape associations).

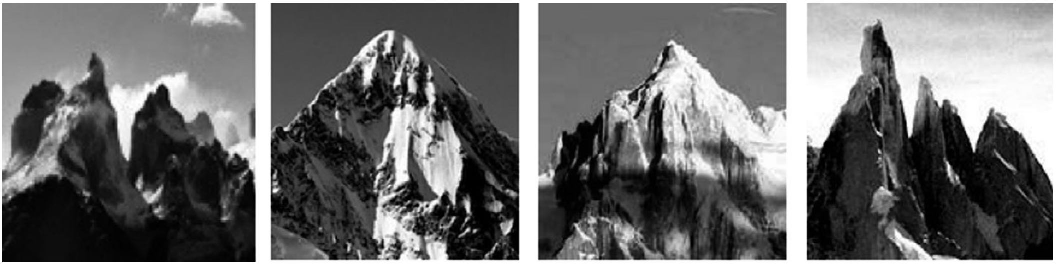

Fig. 1. Example stimuli (from Set A).

2.1.3. Design

For the factor of Labels (between-participants; with-labels vs. no-labels), we assigned 100 individuals to each of the two conditions. In the with-labels condition, participants sorted stimuli into two groups using novel, non-word labels. In the no-labels condition, participants sorted stimuli into two groups without using labels. As data were collected at two different time points (i.e., we collected data for 100 participants in the first instance, and another 100 participants in a second instance), we also included a variable of Time (i.e., Time 1 vs. 2) as a covariate in order to check that participants from Time 1 and 2 did not perform differently to one another. The dependent variable was group category coherence, which was the extent to which participants' categories were coherent with those of other participants in the same condition (i.e., the extent to which they put the same items in the same categories).

2.1.4. Procedure

The experiment was run using Qualtrics (https://www.qualtrics.com). Instructions told participants that they would see images of mountains (see Fig. 2). These images appeared on the left-hand side of the screen and were individually randomized for each participant. They were asked to sort these images into two categories (i.e., "*Please sort them in a way that makes sense to you.*") by dragging the items into one of two predefined spaces onscreen. Participants could drag items across to these spaces in any order and could change their categories freely during a round.

In the with-labels condition, each of these spaces was labeled with a non-word label (and these labels remained the same across rounds for each participant). In the no-labels condition, the two predefined spaces were unlabeled. Participants were told that they could have any number of items in each group, as long as each group contained between 8 and 16 items when they had finished. They were also instructed that they must sort every item (i.e., they could not leave any items uncategorized). Participants sorted set A in round 1, set B in round 2, and set C in round 3 (set order was fixed to allow item comparisons across all participants within each condition). Participants could not return to a round after proceeding and were told that they had 15 min to complete the study.

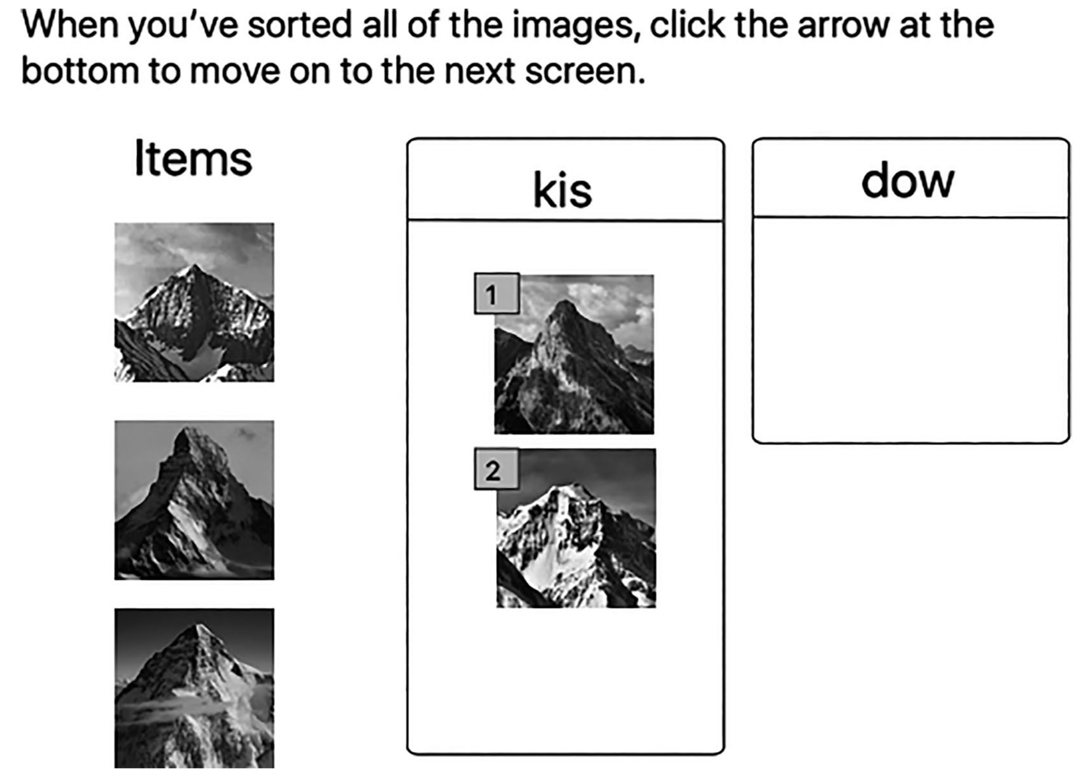

Fig. 2. Example screenshot of online sorting task.

2.2. Results

2.2.1. Analysis

We implemented the cultural consensus model (CCM; Romney, Batchelder, & Weller, 1986) to calculate category coherence. We took every participant from the with-labels condition and compared their data to every other participant from the with-labels condition; similarly, we took every participant from the no-labels condition and compared their data to every other participant from the no-labels condition. We compared participants specifically within each instance of Time (1, 2) to allow us to compute the covariate of Time for use in the analysis (see Section 2.1.3).

To compute these CCM scores, for every participant, we coded whether they put each possible pair of items ($24 \times 23/2 = 276$ item pairs per round) into the same category or not. If a participant placed two items into the same category, that item pair was coded as 1; if not, it was coded as 0. We used this data to compute a proportional measure of association between each possible pairing of participants (see Fig. 3). For example, if a pairing of participants matched on all 276 item pairs within a round, the pairing was given a score of 1. If a pairing matched on 207 item pairs, it was given a score of $207/276 = 0.75$. Hence under this approach, a change in 0.01 CCM score would reflect a

Person 1				
Item *a*	$a_{(0,1)}$	Item *b*	$b_{(0,1)}$	*a=b*
1	1	2	0	0
1	1	3	1	1
1	1	4	0	0
2	0	3	1	0
2	0	4	0	1
3	1	4	0	0

Person 2				
Item *a*	$a_{(0,1)}$	Item *b*	$b_{(0,1)}$	*a=b*
1	1	2	1	1
1	1	3	1	1
1	1	4	0	0
2	1	3	1	1
2	1	4	0	0
3	1	4	0	0

	Person 1 vs. 2		
	Person 1	Person 2	*CCM*
$a=b_1$	0	1	0
$a=b_2$	1	1	1
$a=b_3$	0	0	1
$a=b_4$	0	1	0
$a=b_5$	1	0	0
$a=b_6$	0	0	1
Average *CCM* (3÷6) =			0.5

Fig. 3. Subset of 4 items demonstrating how CCM was calculated between participants. Scores were next averaged by participant and by round.

change in the categorization of 2.76 item pairs in the task. This approach allowed us to compare the category coherence of groups of people, rather than pre-defined, yoked pairs (i.e., as in previous work examining the effects of labels on category coherence; e.g., Markman & Makin, 1998; Suffill et al., 2016).

These pair coherence scores (for all possible pairs) were subsequently averaged by participant (yielding each participant's average similarity to the group, as a measure of social coherence) to avoid multiple comparisons being entered into the analysis. As participants sorted three sets of items (one per round), this produced three average CCM scores for each participant (i.e., average coherence in round 1, round 2 and round 3). Data included Round (1–3) for modeling random effects, but not as a fixed effect (note that we always used Set A, B, and C in Round 1, 2, and 3 respectively, to allow item comparisons across all participants within each condition).

2.2.2. Descriptive statistics

Average CCM scores (*SD*) across conditions are summarized in Table 1 and visualized in Fig. 4.

2.2.3. Linear mixed-effects models (LME) analysis

To normalize the sampling distribution of the proportional CCM scores, the scores were *Z*-transformed with $Z = 0.5 \times \ln[(1 + r)/(1 - r)]$. To test for the effects of Labels on category coherence, data were analyzed in R 3.2.1 (R Core Team, 2015), with the

Table 1
Experiment 1: Average CCM scores (*SD*)

Round	No Labels	With Labels
1	0.29 (0.05)	0.28 (0.04)
2	0.28 (0.03)	0.27 (0.03)
3	0.28 (0.03)	0.29 (0.06)
$\bar{X}$	0.29 (0.04)	0.28 (0.05)

lme4 package, version 1.1-8 (Bates, Maechler, Bolker, & Walker, 2014). This approach allowed us to account for random variance due to differences across participants, and across rounds. The threshold for statistical significance was set at $t > 2$ (Baayen, 2008). A backwards, stepwise elimination approach was used to select fixed effect factors for the final model, using likelihood ratio tests to compare models (for the full analysis procedure, see "Experiment 1 LME markdown" in Data S1). The factors of Labels, Time (1 vs. 2), and Round were centered about 0 using R's default scale function. The maximal converging model included random slopes and intercepts for Labels (no-labels vs. with-labels) by Participant, and for Labels by Round (1–3). The model included Labels as a fixed effect and Time as a covariate, with the reference level set as no-labels:time 1. This

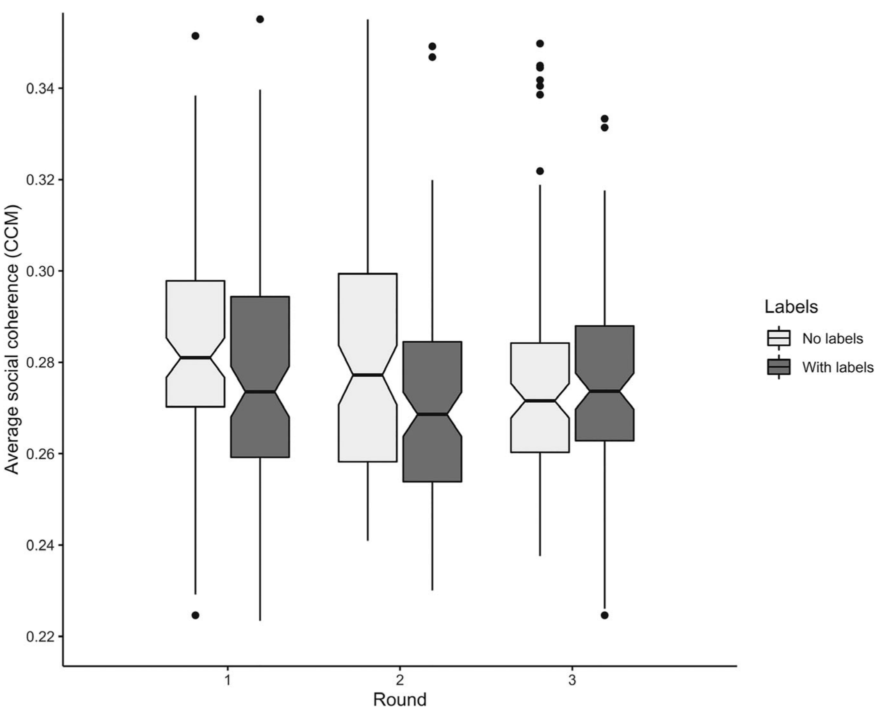

Fig. 4. Experiment 1: Average social coherence by Round and Labels. Notches reflect an approximated 95% confidence interval about the median.

model was not a significantly better fit than the null model ($p > .05$). Hence, the null model was the model of best fit.

2.3. Summary of Experiment 1

Experiment 1 demonstrated no effect of labeling on category coherence: People who sorted with novel labels did not categorize more similarly than those who sorted without labels. However, this experiment did not involve a coordinative goal (i.e., there was no requirement for participants to sort in a way that would make sense to another person). Hence, in Experiment 2, we investigated the effect of labeling on category coherence when participants sorted with a coordinative goal.

3. Experiment 2

3.1. Methods

3.1.1. Participants

These were a further 200 native monolingual English speakers (119 female) from the USA and UK, who did not take part in Experiment 1. Ages ranged from 18–35 years ($\overline{X} = 27.48$, $SD = 4.87$).

3.1.2. Stimuli

The stimuli were identical to Experiment 1.

3.1.3. Design

The design of the experiment was identical to Experiment 1.

3.1.4. Procedure

The procedure was identical to Experiment 1, except for amendments to the instructions. Participants in the no-labels condition were given the additional instructions: "*Please sort them in a way that makes sense to you, but that would also make sense to another person*" and "*You'll be assigned a partner who will separately sort the same items—we'll then compare the way you sorted the items to how they sorted them.*" Participants in the with-labels condition were given the additional instructions: "*Please sort them in a way that makes sense to you, but that would also make sense to another person*" and "*You'll be assigned a partner who will separately sort the same items—we'll then compare the way you sorted the items using the given labels, to how they sorted them.*"

3.2. Results

3.2.1. Analysis

The approach to analysis was identical to that of Experiment 1.

Table 2
Experiment 2: Average CCM scores (*SD*)

Round	No Labels	With Labels
1	0.27 (0.02)	0.30 (0.05)
2	0.27 (0.03)	0.29 (0.03)
3	0.27 (0.02)	0.28 (0.03)
$\bar{X}$	0.27 (0.03)	0.29 (0.04)

3.2.2. Descriptive Statistics

Average CCM scores (*SD*) across conditions are summarized in Table 2 and visualized in Fig. 5.

3.2.3. LME analysis

The treatment of the data was identical to that of Experiment 1 (for the full analysis procedure, see "Experiment 2 LME markdown" in Data S1). The initial model included Labels as a fixed effect and Time as a covariate, with the reference level set as no-labels:time 1. This model was a significantly better fit than the null model ($\chi^2(9) = 53.98$, $p < .001$). Under this model, there was a significant effect of Labels, such that participants who sorted with labels had greater category coherence than those who sorted without labels ($\beta = 0.49$, $SE = 0.19$, $t = 3.31$). There was not a significant effect of Time, and so no evidence that participants performed differently across the two data collection times.

3.3. Summary of Experiment 2

In contrast to Experiment 1, Experiment 2 demonstrated an effect of labels on category coherence, with people who sorted with non-word labels having more similar, or socially coherent, categories than those who sorted without labels (i.e., more coherent by around 5.5 item pairs overall; see Ameel, Storms, Malt, & Sloman, 2005, in which similarly sized differences are indicative of significant effects). To assess whether this labeling effect was specific to the coordinative goal of Experiment 2, we conducted an analysis on the combined data from Experiments 1 and 2.

4. Combined analysis

4.1. Results

4.1.1. Analysis

We included Labels (no-labels vs. with-labels) and Goal (non-coordinative [Experiment 1] vs. coordinative [Experiment 2]) as predictors. Again, Time (Time 1 vs. 2) was included as a covariate.

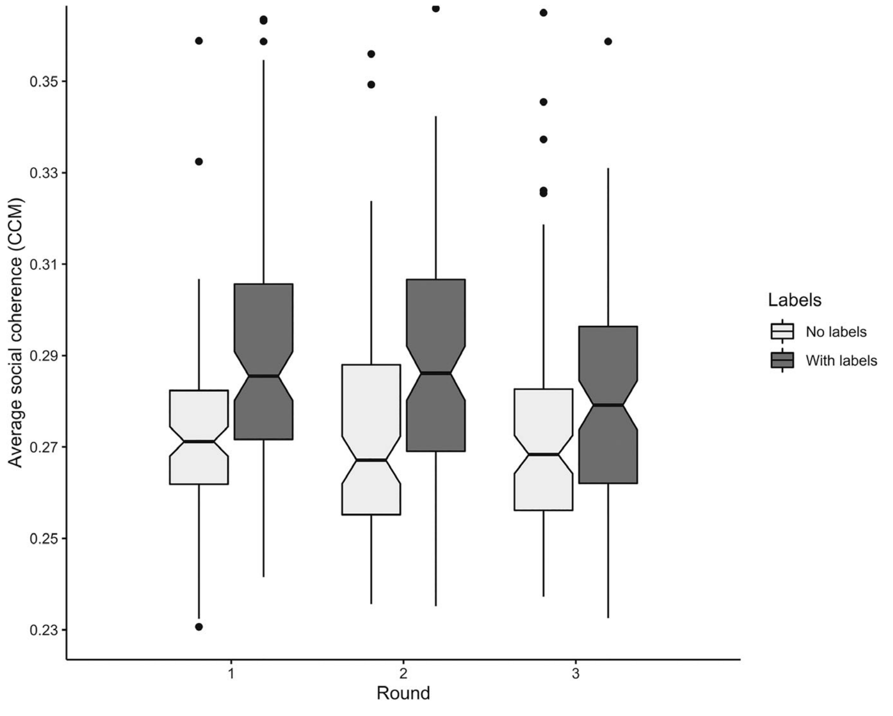

Fig. 5. Experiment 2: Average social coherence by Round and Labels. Notches reflect an approximated 95% confidence interval about the median.

4.1.2. LME analysis

Data treatment were identical to that of Experiments 1 and 2 (for the full analysis procedure, see “Combined LME markdown” in Data S1). The maximal, converging model included random slopes for the factors Labels, Goal and Time by Participant and for Goal and Time by Round. The model included Labels and Goal as fixed effects, as well as the interaction term between factors, plus Time as a covariate, with the reference level set as no-labels:non-coordinative:time 1. This model was a significantly better fit than the null model ($\chi^2(7) = 22.09$, $p < .01$). Removing the interaction term significantly reduced model fit ($\chi^2(1) = 13.96$, $p < .001$). Therefore, the model of best fit was the full model. Under this model, there was a significant effect of Labels, such that participants who sorted with labels had greater category coherence than those who sorted without labels ($\beta = 0.18$, $SE = 0.07$, $t = 2.57$). Importantly, there was also a significant interaction between Labels and Goal, such that participants who sorted with labels and with a coordinative goal had greater category coherence than those who sorted without labels and with a non-coordinative goal ($\beta = 0.51$, $SE = 0.14$, $t = 3.75$). Again, there was no significant effect of Time, hence no evidence that participants performed differently across the two data collection times.

5. Discussion

In two experiments, participants categorized grayscale images of mountains into two, self-determined categories. They sorted items either without using labels, or using two novel, non-word labels. Individual analyses of Experiments 1 and 2 compared the social category coherence of groups of people and the results suggested that categorizing with labels led to greater category coherence only in situations involving coordinative goals. A combined analysis across Experiments 1 and 2 confirmed the specificity of this labeling effect to coordinative settings.

5.1. Coordinative goals and novel labels

In Experiment 1, participants who sorted with novel, non-word labels did not differ in their category coherence from those who sorted without labels. These results suggest that in the absence of a coordinative goal, labeling did not impact how people categorized. Conversely, we suggest that in Experiment 2, the coordinative nature of the goal caused the label to become a potential means for communication, and—through this—a focus for the coordination of people's categories, as is the case with existing labels (Clark, 1996). Lupyan (2008, 2012) argued that labels lead people to focus on dimensions that generalize well across a range of objects. In addition, Lupyan et al. (2007) posited that labels can simplify category distinctions, by allowing people to categorize objects under a single term that represents multiple category dimensions and makes these dimensions more concrete, and easier to access when deciding whether a new object fits into a category.

We propose that for novel labels with no conventionalized meaning, the goal of categorization also plays an important role in participants' selection of dimensions for categorization. That is, people recognize existing (conventional) labels as indicators of category membership—but we suggest that coordination (as a primary feature of language in communication) is necessary for adult sorters to view novel labels in a similar way. As such, when sorters have a coordinative goal, novel labels influence them not only to select certain dimensions, but moreover to select those dimensions that would be sensible to another person. As such, sorters apply the label to the category members in a way that would make sense to another person doing the same task (Clark, 1996; Clark & Brennan, 1991).

It is also possible that the goal of Experiment 2 might have caused people to more actively engage with the labels during sorting (i.e., because they were asked to use the labels to sort the items in a way that would make sense to another person) than in Experiment 1, in which participants may have passively placed items into groups without considering the application of the labels to items (see Duff, Hengst, Tranel, & Cohen, 2006, for an account of how active labeling in amnesic patients creates lasting referential terms for novel shapes, but passive labeling does not). And because people more actively engaged with the labels during categorization, the labels played a stronger role in coordinating the way in which they sorted (i.e., again by influencing them to select dimensions that were also sensible to others).

5.2. Coordination without interaction

Unlike previous research examining how people coordinate labels for objects and linguistic categories (e.g., Clark & Brennan, 1991; Garrod & Doherty, 1994; Markman & Makin, 1998; Suffill et al., 2016), our participants did not encounter each other. Moreover, we varied the pairs of non-word labels across participants in order to reduce the effects of specific non-words (e.g., through sound symbolism or word-form associations) on people's categorization, and people who did share the same labels were never compared with one another. Despite this, people's categories were more socially coherent (i.e., more similar to a group of other people) when they used novel labels during sorting with a coordinative goal than without a coordinative goal.

Given that participants in each condition did not provide feedback to each other, it is likely that the category dimensions people selected when using labels for coordination were based on shared, perceptual features of the stimuli that were sensible to other sorters. In addition, as the majority of participants who sorted with labels had different word label pairs, this labeling effect is unlikely to be due to any non-arbitrary features of the specific label pairs. This reiterates the importance of language as a tool for coordinating with others (Clark, 1991), even in the absence of direct communication.

Lastly, our results have interesting implications for theories of unsupervised categorization. In particular, they suggest that the process of unsupervised categorization may not always be as straightforward as largely assumed: If sorters picked simply the most intuitive dimensions for categorization (e.g., unidimensional sorting; Ashby et al., 1999), then we might have expected sorters to uniformly categorize items in identical ways to each other, irrespective of contextual factors (i.e., using novel labels, or having a coordinative goal). But instead we found that these contextual factors modulated the categories that people used, and their social coherence. Thus, while our findings cannot tell us precisely which dimensions people selected for sorting, they do suggest that the dimensions that people choose during unsupervised categorization are somewhat flexible, in ways that are by definition independent of error-based feedback (Goldstone, 1994; Pothos, Edwards, et al., 2011), and conversely are at least partially dependent on factors such as language and the context (or goal) of sorting. We note however that our participants were limited to categorizing the items into two groups only (i.e., rather than choosing their own number of categories; Pothos, Edwards, et al., 2011). In future work, it would be beneficial to examine the effects of novel labels alongside contextual factors (such as labels and the context of sorting) on category flexibility and social coherence when sorters are free to determine their own number of categories.

5.3. Conclusion

Like existing labels, novel labels can influence which dimensions people select for categorization, but we have shown that they can also increase the social coherence of people's categories. Moreover, we have shown that the goal of sorting is crucial to this labeling effect. When people sort with a goal that does not promote coordination, novel labels do not increase category coherence between sorters. But when people sort with a goal that does promote coordination, novel labels can increase category coherence even without communication between sorters. As such, our results not only provide evidence that the act of labeling with novel labels increases the potential for the coordination of people's categories, but crucially they show that labels do so specifically when the goal of sorting encourages coordination among sorters.

Acknowledgments

The research leading to these results has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 613465. This research was carried out as part of the *Advancing the European Multilingual Experience* project.

References

Althaus, N., & Mareschal, D. (2014). Labels direct infants' attention to commonalities during novel category learning. *PLoS ONE*, *9*(7), e99670.

Althaus, N., & Plunkett, K. (2016). Categorization in infancy: Labeling induces a persisting focus on commonalities. *Developmental Science*, *19*(5), 770–780.

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, *53*(1), 60–80.

Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, *61*(6), 1178–1199.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Barnhart, W. R., Rivera, S., & Robinson, C. W. (2018). Effects of linguistic labels on visual attention in children and young adults. *Frontiers in Psychology*, *9*, 358.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version, 1.1.-4.

Clark, H. H. (1991). Words, the world, and their possibilities. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The perception of structure: Essays in honor of Wendell Garner*. Washington, DC: American Psychological Association.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press, *952*, 274–296.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on Coordinationly Shared Cognition*, *13*(1991), 127–149.

Duff, M. C., Hengst, J., Tranel, D., & Cohen, N. J. (2006). Development of shared information in communication despite hippocampal amnesia. *Nature Neuroscience*, *9*(1), 140.

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, *53*(3), 181–215.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*(2), 178.

Jackendoff, R. S. (2002). *Foundations of language: Brain, meaning, grammar, and evolution*. Oxford, UK: Oxford University Press.

Köhler, W. (1929). *Gestalt psychology*. New York, NY: Liveright.

Laskowski, C. (2011). The role of language in conceptual coordination. Unpublished PhD thesis, University of Edinburgh.

Lupyan, G. (2008). From chair to "chair": A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General*, *137*(2), 348–396.

Lupyan, G. (2012). What do words do? Toward a theory of language-augmented thought. In B. H. Ross (Ed.), *The psychology of learning and motivation* (vol. *57*, pp. 255–297). San Diego, CA: Academic Press.

Lupyan, G., & Casasanto, D. (2014). Meaningless words promote meaningful categorization. *Language and Cognition*, *7*(2), 167–193.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking redundant labels facilitate learning of novel categories. *Psychological Science*, *18*(12), 1077–1083.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Coherence and the linguistic categorization of artifacts. *Journal of Memory and Language*, *40*(2), 230–262.

Markman, A. B., & Makin, V. S. (1998). Referential communication and category acquisition. *Journal of Experimental Psychology: General*, *127*(4), 331.

Pothos, E. M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science*, *26*(3), 303–343.

Pothos, E. M., & Chater, N. (2005). Unsupervised categorization and category learning. *The Quarterly Journal of Experimental Psychology*, *58*(4), 733–752.

Pothos, E. M., & Close, J. (2008). One or two dimensions in spontaneous classification: A simplicity approach. *Cognition*, *107*(2), 581–602.

Pothos, E. M., Edwards, D. J., & Perlman, A. (2011). Supervised versus unsupervised categorization: Two sides of the same coin? *The Quarterly Journal of Experimental Psychology*, *64*(9), 1692–1713.

Pothos, E. M., Perlman, A., Bailey, T. M., Kurtz, K., Edwards, D. J., Hines, P., & McDonnell, J. V. (2011). Measuring category intuitiveness in unconstrained categorization tasks. *Cognition*, *121*(1), 83–100.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: The R Foundation for Statistical Computing. Available at: http://www.R-project.org/.

Romney, A. K., Batchelder, W. H., & Weller, S. C. (1986). Recent applications of cultural consensus theory. *American Behavioral Scientist*, *31*(2), 163–177.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605.

Suffill, E., Branigan, H. P., & Pickering, M. J. (2016). When the words don't matter: Arbitrary labels improve categorical coordination through the anchoring of categories. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1715–1720). Austin, TX: Cognitive Science Society.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive Psychology*, *29*(3), 257–302.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Data S1. See supporting materials for R markdowns (i.e., of Experiments 1, 2 and the Combined analysis), and a list of labels and stimuli used in the experiments.