Journal of Experimental Social Psychology xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Reports

# Abandoning a label doesn't make it disappear: The perseverance of labeling effects ☆

Francesco Foroni *, Myron Rothbart

*SISSA Trieste, Italy*
*University of Oregon, United States*

## HIGHLIGHTS

- ► We explore if labeling effects continue even after labels are abandoned.
- ► Participants judge similarity of drawings of female body types with and without labels.
- ► Silhouettes sharing the same label are perceived as more similar than those having different labels.
- ► When the labels are removed, the labeling effects are diminished but they persist.
- ► It did not matter whether the labels are simply abandoned or had their validity challenged.

## ARTICLE INFO

*Article history:*
Received 9 March 2012
Revised 23 July 2012
Available online xxxx

*Keywords:*
Labeling effect
Perseverance
Categorization

## ABSTRACT

Labels exert strong influence on perception and judgment. The present experiment examines the possibility that such effects may persist even when labels are abandoned. Participants judged the similarity of pairs of silhouette drawings of female body types, ordered on a continuum from very thin to very heavy, under conditions where category labels were, and were not, superimposed on the ordered stimuli. Consistent with earlier research, labels had strong effects on perceived similarity, with silhouettes sharing the same label judged as more similar than those having different labels. Moreover, when the labels were removed and no longer present, the effect of the labels, although diminished, persisted. It did not make any difference whether the labels were simply abandoned or, in addition, had their validity challenged. The results are important for our understanding of categorization and labeling processes. The potential theoretical and practical implications of these results for social processes are discussed.



## Introduction

The tendency to classify objects into categories, and to attach labels to those categories, is ubiquitous. Labeled categories can function as psychological 'equivalence classes' and affect our perception and our judgments of their members (e.g., Eiser & Stroebe, 1972; Goldstone, 1995; Hacking, 1995; Kenney, 2002; Maruna, LeBel, Mitchell, & Naples, 2004; Rosch, 1988; Tajfel & Wilkes, 1963; Whorf, 1956). Early research by Tajfel and Wilkes (1963) investigated how the application of category labels would distort the judgment of simple objects, and social psychologists have dedicated considerable effort to understand how the content or meaning of category labels structure perception and behavior (e.g., Allport, 1954; Darley & Fazio, 1980; Rosenhan, 1973; Rosenthal & Jacobson, 1968; Rothbart, Davis-Stitt, & Hill, 1997).

Rothbart et al. (1997) experimentally varied the placement of both category boundaries and labels and assessed their effects on the judgment of similarity of pairs of individuals. The results showed that – compared to the baseline condition – both boundaries and labels had additive effects increasing the similarity of within-category pairs and decreasing the similarity of between-category pairs. More recently, there have been attempts to separate the content or meaning of the labels from the nature of the labeling process itself (e.g., Carnaghi et al., 2008; Foroni & Rothbart, 2011; Gelman & Heyman, 1999; Walton & Banaji, 2004). This research demonstrates the greater power of nouns over verb predicates in making behavioral inferences (Gelman & Heyman, 1999) as well as the greater impact on person perception of nouns compare to adjectives (Carnaghi et al., 2008). Recently, Foroni and Rothbart (2011) examined the influence of labels' semantic content on the strength of categorization effects using a paradigm in which participants are simultaneously given information about an object's position along a meaningful continuum, as well as the classifying labels associated with regions of that continuum. In their research the social context and source of the labeling, and the semantic content of the classifying labels were manipulated. The evidence consistently showed that judgments of categorized objects

☆ We dedicate this paper to our colleague and friend, Professor Luciano Arcuri, University of Padova, for his long-term interest and support of our research relating language to social cognition. The present work was presented at the conference "Social Perception Cognition and Language in honor of Arcuri". University of Padova, May 20th, 2011. We would like to thank David Hamilton for his useful and constructive comments. This research was supported by the National Institute of Mental Health grant MH40662 to the second author.

* Corresponding author at: SISSA International School for Advanced Studies, Cognitive Neuroscience Sector, Via Bonomea 265, 34136 Trieste, Italy.
*E-mail address:* francesco.foroni@sissa.it (F. Foroni).

0022-1031/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jesp.2012.08.002

are strongly affected by the presence and strength of labels, independent of the context, source and consequences of the labels.

The present paper builds on these findings and investigates the perseverance of labeling effects, that is, whether the effects of a set of category labels continue to exert its influence even after the labels themselves have been removed.

Consider the often-heard complaint by statistics instructors about the common practice of converting continuous measures into discrete categories, where information loss increases as the number of grouping categories decreases (Glass & Hopkins, 1996; Levin, 1987). In personality research, for example, the dimension introversion–extraversion is measured as a continuous variable, but frequently dichotomized in data analysis. This practice functionally ignores the large variation within each category, and exaggerates the differences between individuals who are barely on opposite sides of the boundary. Despite admonitions to let continuous measures be continuous, the siren call to group continuous data into discrete categories proves difficult to resist. Although categorizing and labeling continuous data has the virtue of simplifying a data set, it has been suggested that such categorization may not only temporarily "distort" the meaning of the categorized objects (Darley & Fazio, 1980; Fiske & Neuberg, 1990; Hamilton, Sherman, & Ruvolo, 1990; Miller & Turnbull, 1986), but may also produce a more stable change in the way we perceive the categorized objects — even when the arbitrariness of reducing continuous data into simplifying categories is explicitly acknowledged (e.g., Harnad, 1987, 2003).

The phenomenon by which continuous variation along a physical dimension is nonetheless perceived as a series of discrete regions (e.g., in speech perception) is referred to as Categorical Perception (CP), and has been extensively investigated by cognitive psychologists (e.g., Harnad, 1987, 2003). Harnad (1987) argues that applying a categorization system onto a continuum is the same as an "analog-to-digital conversion" (p. 542) that records a continuous region of physical variation into a discrete, labeled equivalence class even when there is no discontinuity. Such induced conversion due to the use of the categorical system (cf. Snowdon, 1987) may result in a loss of information that can't easily be regained once the categorization system is in some way abandoned or nullified. It has been argued that similar reasoning applies to social category labels that may be able to change the psychological representation of the categorized objects (Goldstone, Lippa, & Shiffrin, 2001; Kikutani, Roberson, & Hanley, 2008).

In support of such claims, Rothbart et al. (1997) found that when arbitrary category boundaries were superimposed on – and pitted against – meaningful continuous information, greater use was made of the former than the latter. Foroni and Rothbart (2011) found further, using a much richer stimulus set, that category labels had a power of their own that was not moderated by the expert source of the labels, or the consequences of applying the labels. Based on the disproportionate power of category labels to influence judgment suggested by this research, we expect that such labels would continue to exert their influence over previously categorized stimuli, even when the category labels have been removed.

Thus, the present research addresses the following questions: Once the effects of category labels have been established, do those effects disappear when the category labels are removed by exposing participants to the stimulus objects in its original continuous information? Or does it take a more active intervention to negate the role of category labels? In other words, does a category label system, once established, produce long-term effects as suggested by Harnad (1987)?

## Overview

In this study, we investigated the perseverance of category labels effects using a paradigm developed by Foroni and Rothbart (2011). Foroni and Rothbart presented participants with silhouette drawings of female body types, ordered along a continuum from very thin to very heavy, and asked them to judge the degree of similarity between pairs of silhouettes (i.e., baseline). The task was then repeated in a subsequent second phase where category boundaries and labels were experimentally introduced and present at the time of the second judgments. The present experiment used a paradigm that is a modification of the one successfully implemented by Foroni and Rothbart (2011) and consisted of three phases, rather than two as the original one.

In the first phase all participants made similarity judgments between selected pairs of silhouettes that were presented together at their corresponding positions on the continuum.

In the second phase participants were assigned to one of three conditions: Control, labels absent, labels challenged. Participants in the control condition repeated the similarity judgment exactly as Phase 1, while category labels were present for both label conditions.

In the third phase of the experiment, all participants again judged the similarity between silhouettes along the continuum exactly as done in Phase 1. Participants of the labels challenged condition, but not those of the other two conditions, were exposed to a written rationale challenging the validity of the labels.

## Method

### *Participants*

One hundred-twenty-one students (71 females), participated in this computer-based experiment as partial fulfillment of a research requirement for an Introductory Psychology course.

### *Stimulus material and design*

The stimulus material consisted of 9 female silhouettes, frontal view, at roughly equally-spaced intervals along a continuum ranging from very thin to very heavy (Fig. 1, derived from Furnham & Alibhai, 1983). The silhouettes were selected to allow the continuum to be divided into 3 equally-spaced regions with 3 silhouettes in each region.

Participants were seated in front of a personal computer where they could self-administer the experiment. Standardized instructions presented on the screen described the experiment as one involving social perception. The silhouettes were described as "Varying along a single continuum called the Body Ratio Index (or BRI) […]".

The experiment consisted of three phases separated by a short break (approximately 10 min of intervening tasks). All participants in Phase 1 were presented with the continuum and asked to express similarity judgments between selected pairs of silhouettes. In total 15 pairs of silhouettes, either one- or two-units apart from each other on the continuum were presented. The two silhouettes of each pair were presented together at their corresponding positions on the continuum and stayed on the screen until a judgment was expressed. Participants judged the similarity between the two silhouettes on a 9-point Likert-scale (from 'not at all similar' to 'extremely similar') on three different dimensions ('personality,' 'life style,' and 'body type'). After participants entered the rating for the target pair, the next trial started with the subsequent pair presented on the continuum. In Phase 2 the nature of the continuum on which the stimuli were presented was systematically varied between subjects and participants again judged the similarity between silhouettes. In Phase 3, all participants judged again the similarity between silhouettes along the continuum exactly as was done in Phase 1. In each phase participants were informed that the silhouettes to be judged might be different from the previous phase.

Participants were assigned randomly to one of three conditions:

(a) Control condition: Participants were presented with the same exact continuum in Phase 1, 2, and 3, with no category boundary and no labels.
(b) Labels absent condition: Participants were presented with a continuum with no category boundaries and no labels in Phase

Please cite this article as: Foroni, F., & Rothbart, M., Abandoning a label doesn't make it disappear: The perseverance of labeling effects, *Journal of Experimental Social Psychology* (2012), http://dx.doi.org/10.1016/j.jesp.2012.08.002

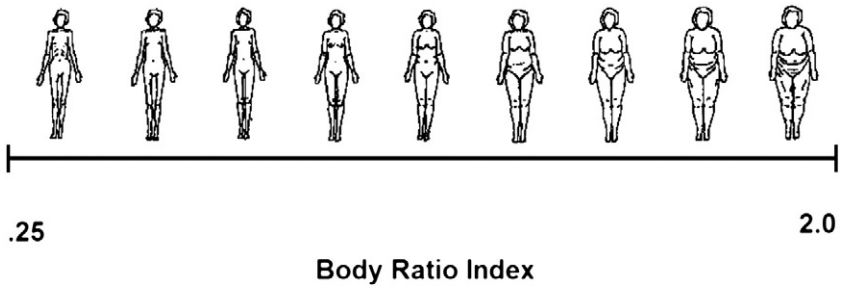


**Fig. 1.** Set of silhouette images used as stimuli adapted from Furnham and Alibhai (1983).

1. In Phase 2 the continuum was divided by tick marks into three equally-spaced regions, with three labels attached ('anorexic,' 'normal,' and 'obese').[1] During Phase 3 the continuum was exactly as in Phase 1.

(c) Labels challenged condition. Participants were presented with a continuum with no category boundaries and no labels in Phase 1. In Phase 2 the continuum was divided by tick marks into three equally-spaced regions, with three 'strong' labels attached ('anorexic,' 'normal,' and 'obese') as the labels absent condition. However, just before starting Phase 3, participants in this condition read that "There is recent evidence from the Bulletin of American Nutritionists […], that the most important information is represented by each silhouette's actual score on the BRI continuum rather than a silhouette's placement within a category. […] Any medical treatment would be based solely on a person's individual BRI score, and would never be based on their category membership […]" (cited from the experiment instructions). The previous rationale supported the invalidity of the label system.[2] The two label conditions differed only in the presence (or absence) of the above rationale for invalidity of the label system. The differences between the two label conditions may be thought of as the difference between a passive and an active abandonment of category labels: in both cases the category labels are absent, but in the labels challenged condition the participants are instructed to focus on the continuous rather than (the "invalid") categorical information.

Importantly, the debriefing phase confirmed that (a) participants believed that the labels were generated by expert nutritionists, (b) they accepted the validity of the 'challenge to the label system', and (c) they showed no signs that they have guessed the rationale and goals of the experiment.

### *Hypotheses*

First, we predicted, in line with the results found by Foroni and Rothbart (2011) that category labels should result in greater perceived similarity for silhouettes sharing the same label than for those having different labels. This difference should be apparent in comparisons between Phase 1 and Phase 2. Second, we predicted that these labeling effects will be significantly reduced in magnitude after the labels were removed. Third, albeit reduced in magnitude, we expect that the labeling effects would continue to be present in Phase 3. Finally, we predicted that any reduction in the strength of labeling effects would be greater when the category labels are actively challenged (labels challenged condition) than if they are merely passively removed (labels absent condition).

## Results

### *Data reduction*

Each participant made three similarity judgments for each of the 15 pairs (in each phase). The three dimensions of similarity (physical, personality, and lifestyle similarity) were highly correlated (average within-subject correlation of $r = .50$ for Phase 1; $r = .69$ for Phase 2, and $r = .72$ for Phase 3). Thus, the three similarity judgments were averaged to create a single similarity measure for each pair of silhouettes (separately for each phase). The 15 pairs varied according to the distance between the members within a pair (1 vs. 2 units apart: first and third silhouette from left in Fig. 1 being an example of 2-unit apart pair), and to whether the members of the pair existed within the same or different categories (used in Phase 2: first and third silhouette from left being an example of within-category pair), resulting in four different pair-types for each phase: 1-unit apart/across-category boundary, 2-unit apart/across-category boundary, 1-unit apart/within-category boundary, and 2-unit apart/within-category boundary. The similarity judgments for the pairs were averaged within each of these four pair-types creating four different similarity scores one for each pair-type (separately for each phase). Then, the 1-unit-apart and 2-unit-apart similarity scores were further averaged to yield a single within-category similarity score and a single across-category similarity score separately for each phase.[3] For simplicity, we created a single index subtracting the across-boundary-similarity score from the within-boundary-similarity score: single categorization index (separately for each phase).

The primary evidence for categorization is present in the difference between within- and across-boundary similarity judgments as we move from an uncategorized state (Phase 1) to a categorized state (Phase 2). That is, the same pairs of silhouettes are judged first in the absence and then in the presence of category boundaries, and if category boundaries and labels have an effect, the within-category pairs should be judged as more similar, and across-category pairs should be judged as less similar (increasing the categorization index). This categorization index was used as the main dependent variable for all the analyses across Phases 1, 2, and 3. Increase in this index from Phase 1 to Phase 2 in the label conditions, for instance, would indicate an increase of the difference between within- and across-boundary pairs and thus evidence of categorization.[4]

[1] Foroni and Rothbart (2011) demonstrated that the strength of the label is an important moderating factor of the magnitude of the labeling effects. A set of strong labels was therefore chosen for this experiment.

[2] 'Doctors from the American Nutritionists Association' were described as the source of the label system and of the information regarding its invalidity to maximize the power of the intervention.

[3] Analyses including the factor distance (1-unit vs. 2-unit) did produce parallel results, thus, this factor was no longer considered.

[4] Since within-boundary pairs (1–2,2–3,1–3,4–5,5–6,4–6,7–8,8–9,7–9) are different from the across-boundary pairs (2–4,3–4,3–5,5–7,6–7,6–8) we anticipate that the baseline similarity of within- and across-boundary pairs may be different. This difference was also present in previous research (see Foroni & Rothbart, 2011) and it is most likely due to the presence of the extreme silhouettes 1 and 9 which are present for the within-boundary pairs, but absent for the across-boundary pairs. The presence of these baseline differences (responsible for the negative categorization indexes baseline, see Fig. 2) are assessed in phase 1 control condition and the key theoretical issue is how the categorization index changes as a result of the superimposition of category labels in Phase 2 and of their removal in Phase 3 in comparison to the control.

Please cite this article as: Foroni, F., & Rothbart, M., Abandoning a label doesn't make it disappear: The perseverance of labeling effects, *Journal of Experimental Social Psychology* (2012), http://dx.doi.org/10.1016/j.jesp.2012.08.002

*Data analyses*

We first analyze the full three-phase design and then we turn to analyze specifically the change in the categorization index from Phase 1 to Phase 2 (i.e., to test the effect of the presence of the labels). Then we will focus on the comparison between Phase 2 and Phase 3 to determine if abandoning the labels system reduces the effect of the labels. Finally, we will analyze the change in the categorization index from Phase 1 to Phase 3. This latter analysis will assess whether any categorization effect remains in Phase 3, after the categories and labels have been removed.

*Phase 1, Phase 2, and Phase 3*

A mixed 3×3 ANOVA was conducted on the categorization index, with the first factor within subjects (Time: Phase 1 vs. Phase 2 vs. Phase 3) and the second between subjects (Condition: control vs. labels absent vs. labels challenged).[5] The means and standard errors are presented in Fig. 2.

There was a significant effect of Time ($F(2, 203) = 35.62$, $p < .001$, $\eta^2 = .23$[6]) showing a general increase in categorization index over time, both from Phase 1 to Phase 2, and from Phase 2 to Phase 3. In addition, the interaction between Time and Condition was also significant, $F(4, 236) = 8.70$, $p < .001$, $\eta^2 = .13$. The presence of the labels affected the categorization index across the three phases; namely, when labels are present they induce categorization effects. These results are in line with the findings reported by Foroni and Rothbart (2011) where the presence of labels differentially affects similarity judgments for within- and across-boundary pairs. We now turn to the comparison between Phase 1 and 2, then Phase 2 vs. Phase 3, and finally Phase 1 vs. Phase 3. In the following analyses, p-values are Bonferroni-corrected, when necessary, to correct for multiple comparisons.

*Labeling effect: Phase 2 vs. Phase 1*

The categorization index was subjected to a mixed 2×3 ANOVA with the first factor within subjects (Time: Phase 1 vs. Phase 2) and the second between subjects (Condition: control vs. labels absent vs. labels challenged).[7]

The factor time was significant indicating that the categorization effect in Phase 2 was significantly greater than in Phase 1, $F(1, 118) = 50.79$, $p < .001$, $\eta^2 = .30$. As expected, this effect was qualified by the interaction between Time and Condition $F(2, 118) = 12.23$, $p = .001$, $\eta^2 = .17$. Both label conditions show a significant increase from Phase 1 to Phase 2 ($t(34) < 4.57$, $p = .003$ and $t(44) < 7.24$, $p = .003$ for labels absent and labels challenged respectively) while the control condition shows no change, $t(40) < 1$, *ns*. In addition, the two label conditions did not differ from each other: namely, there is no difference in the magnitude of the reduction of the categorization effect ($F < 1$, *ns.*).[8]

Thus, the presence of labels affects similarity judgments. Judging the similarity of pairs of silhouettes in the presence of labels increases the categorization index: namely, it makes the pairs sharing the same label (within-boundary pairs) appear more similar than those having different labels (across-boundary pairs). These results replicate and provide clear additional support for the biasing effect of the labels on judgments of similarity (Foroni & Rothbart, 2011; Rothbart et al., 1997).

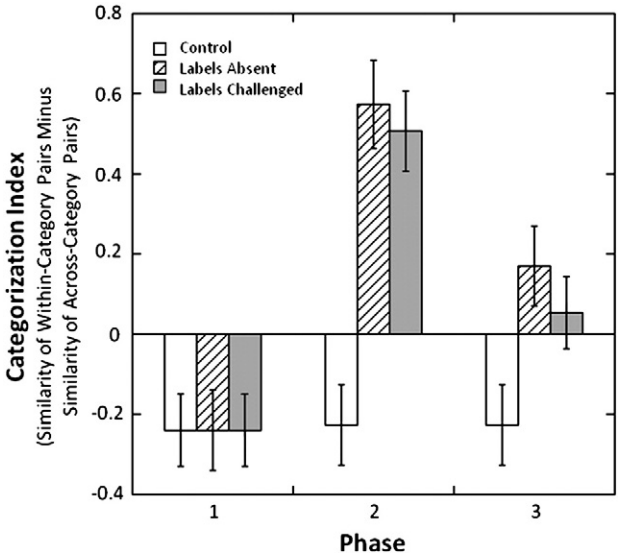


**Fig. 2.** Categorization index by Condition and by Experimental Phase. Increase in the categorization index from Phase 1 to Phase 2, for instance, would indicate an increase of the difference between within-boundary and across-boundary pairs and thus evidence of categorization. Bars represent SE of the mean.

*Abandoning the labels: Phase 3 vs. Phase 2*

Both label conditions showed a significant increase in categorization effect from Phase 1 to Phase 2. Now we test whether abandoning the labels system in Phase 3 will reduce the categorization effect observed in Phase 2. The two label conditions (labels absent and labels challenged) will also be compared to test whether challenging the validity of the labels will be more powerful than simply removing them. Thus, the categorization index was subjected to a mixed 2×3 ANOVA with the first factor within subjects (Time: Phase 2 vs. Phase 3) and the second between subjects (Condition: control vs. labels absent vs. labels challenged).

The factor time was significant indicating that the categorization effect in Phase 2 was significantly greater than in Phase 3, $F(1, 118) = 27.98$, $p < .001$, $\eta^2 = .19$. As expected this effect was qualified by the interaction between Time and Condition $F(2, 118) = 7.25$, $p = .001$, $\eta^2 = .11$. Both label conditions show a significant decrease from Phase 2 to Phase 3 ($t(34) < 3.93$, $p = .003$ and $t(44) < 5.02$, $p = .003$ for labels absent and labels challenged respectively) while the control condition showed no change, $t(40) < 1$, *ns*. In addition, the two label conditions did not differ from each other: namely, there is no difference in the magnitude of the reduction of the categorization effect ($F < 1$, *ns.*).

Thus, abandoning a previously-used labeling system reduces its biasing effect as indicated by the significant reduction in the categorization index, a reduction that is not present in the control. It does not matter if the labels are passively abandoned (labels absent) or are actively challenged in their validity (labels challenged).

*The perseverance of labeling effect: Phase 3 vs. Phase 1*

Both label conditions showed a significant reduction in categorization effect from Phase 2 to Phase 3. However, evidence for the persisting effects of category labels would be present if in Phase 3 the label conditions still show a larger categorization index compared to Phase 1. The two label conditions (labels absent and labels challenged) will also be compared to test whether challenging the validity of the labels will be more powerful than simply removing them. The categorization index was subjected to a mixed 2×3 ANOVA with the first factor within subjects (Time: Phase 1 vs. Phase 3) and the second between subjects (Condition: control vs. labels absent vs. labels challenged).

[5] The original ANOVA included Gender of the participants as a between-subject factor, but as Gender was not significant either as main effect or in interaction with other variables, it was not considered further in the analyses.

[6] The number of degree of freedoms of this test of significance has been adjusted using Greenhouse–Geisser correction criteria to correct for the violation of Sphericity.

[7] In this comparison (Phase 1 vs. Phase 2) as well as the following ones (Phase 2 vs. Phase 3 and Phase 1 vs. Phase 3) the results are the same if, instead, a covariate analysis of variance is used.

[8] Note that the two label conditions (labels absent and labels challenged) differ from each other only in Phase 3. Thus, no difference is expected between them during phase 1 and 2.

Please cite this article as: Foroni, F., & Rothbart, M., Abandoning a label doesn't make it disappear: The perseverance of labeling effects, *Journal of Experimental Social Psychology* (2012), http://dx.doi.org/10.1016/j.jesp.2012.08.002

The factor time was significant indicating that the categorization index in Phase 3 was significantly greater than in Phase 1, $F(1, 118) = 17.36$, $p < .001$, $\eta^2 = .13$. As expected this effect was qualified by the interaction between Time and Condition $F(2, 118) = 4.16$, $p = .018$, $\eta^2 = .07$. Both label conditions still show in Phase 3 a larger categorization index compared to Phase 1 ($t(34) < 3.37$, $p = .006$ and $t(44) < 3.06$, $p = .012$ for labels absent and labels challenged respectively) while the control condition shows no change, $t(40) < 1$, *ns*. In addition, the two label conditions did not differ from each other: namely, there is no difference in the magnitude of the perseverance of the categorization effect ($F < 1$, *ns*.).

Even though abandoning a previously-used label system reduces its biasing effect, the labeling effects nonetheless persist. Moreover, it does not seem to matter whether the label system is passively abandoned or is actively challenged.[9]

## Discussion

Labels have strong influences on participants' perceived similarity of pairs of silhouettes. When labels are present (Phase 2) participants judged individuals sharing the same label as more similar than those having different labels. These results replicate those reported by Foroni and Rothbart (2011) and support the notion of language-induced effects similar to category perception (Harnad, 2003). When the silhouettes are subsequently shown again without labels (Phase 3), the labeling effects are reduced in magnitude but they nevertheless persist. It did not matter if the labels were simply abandoned (labels absent condition) or, in addition, had their validity challenged (labels challenged condition). Although differences between the passive and active abandonment of category labels were expected, none were found in this research. Perhaps even an active challenge to the legitimacy of the category labels is not sufficient to override categorization effects. Future research should try to identify possible ways to reduce the persistence of the labeling effects in the same way that debiasing research tested the limits of cognitive processes (e.g., Galinsky & Moskowitz, 2000); extending the time period between 'Phase 2' and 'Phase 3' or providing more explicit and powerful attempts to discredit the labels are good candidates for future research.

The data reported here are consistent with the view that categories influence the representation of categorized objects (cf. Goldstone et al., 2001), and also consistent with Harnad's (1987) suggestion that a categorization system produces a loss of information that can't easily be regained once the categorization system is discharged or abandoned.

In our view, the visual judgment paradigm used here provides a stringent test of our hypotheses. First, familiar stimuli were used and pairs of stimuli were simultaneously present when judgments were made. Then, each silhouette's position on the continuum is apparent, visible, and hard to ignore, making it easy for participants to ignore category labels when no longer present. Nevertheless, participants continue to show significant labeling effects. Thus, we should expect even stronger effects when the labeling system is less easily ignored (e.g., after a long history of usage or in the absence of evidence for its invalidity) or the situation becomes more complex and ambiguous that characterizes daily life (cf. Corneille, Klein, Lambert, & Judd, 2002).

In line with this argument, in an interesting experiment with stock portfolio managers (Kida, Smith, & Maletta, 1998), participants studied two sets of stocks presented sequentially. In List 1, one stock (stock A) stood out from all the others as superior on a number of indexes. Later, participants studied List 2, where all stocks were roughly comparable in quality, where some stocks might be higher of some indexes while others were higher on different indexes. When then asked to choose between stock A from List 1 and say stock G from List 2 – where stock G was slightly but consistently higher than stock A on the various indexes – participants showed a strong preference for stock A over G. Presumably stock A had been categorized as 'best' within its List 1, and was preferred over stock G which was categorized (at best) as 'above average' within its List 2. In short, the participants were focused on the evaluative labels, which favored stock A over G, and ignored the far more relevant continuous information, which favored stock G over A.

How much of the present effects are attributable to experimenter demand, or strategic approaches to the experiment (Goldstone et al., 2001)? Although demand characteristics cannot be ruled out with certainty, several arguments are relevant to this concern. First, post-experimental debriefings indicated that participants accepted the validity of the challenge by expert nutritionists and tried to ignore the categorization system. Second, the present findings are entirely consistent with the results obtained with absolute estimates of weight used in Foroni and Rothbart (2011), a measure that is less susceptible to experimenter demands (cf. Kobrynowicz & Biernat, 1997). Third, if the participants were sensitive to experimenter demands, the effects of category labels should have been much less in the labels challenged condition than in the labels absent condition, yet the effects across the conditions were virtually identical ($F < 1$).

Once objects along a continuum are grouped into a small number of subtended categories, those objects become equated with the categories and lose the properties that originally differentiated one category member from another. The simple act of naming a section of a continuum will imply that a category is richly structured (Hall & Moore, 1997; Markman, 1989), when in reality it is not and this may stimulate essentialistic assumptions about category member (see also Carnaghi et al., 2008). The present results suggest that such effects may be hard to eliminate. Once someone is categorized as part of the category "obese", for instance, he or she will face the implications of this membership longer after the labels are removed or longer after the labels no longer apply (e.g., after a successful diet and/or physical exercise) as the effects of stigma on social perception persist long after the category label is no longer applicable as implied by Rosenhan's (1973) elegant research.

However, there may be even more pernicious implications of such persistence. Using the same continuum, labels, and stimuli, in fact, Foroni (2006) showed that the labels affect also female participants' judgments of their own body — modifying their own body image. Thus this research may have also implications for mental health (cf. Grogan, 1999).

The perseverative effects observed here have direct implications for social perception. The link between the perseverative effects of category labels and the de-individualization of group members is an important one and deserves further investigation (cf., Hamilton et al., 1990). More generally, however, it should be clear that the type of categorization effects demonstrated here is intimately linked to the process of stereotyping. It is implicit in the definition of a stereotype that a category label minimizes within-category differences and accentuates between-category differences, and this research may help explaining the persistence of stereotyping effects, often in the absence of confirming information (e.g., Rothbart & John, 1985).

The importance of this issue for social perception and judgment cannot be over-estimated, as many theories of stereotype change are predicated on the assumption that members originally perceived as 'only' category members come to acquire a complex set of individuating characteristics, which in turn increases the perceived complexity of the category as a whole (Fiske & Neuberg, 1990). These findings suggest

[9] For simplicity and because of the focus of the present paper we implemented as main dependent variable a single index (single categorization index) for each phase. We also conducted 2 separate 3×3 ANOVAs on within-boundary-similarity and across-boundary-similarity ratings with the first factor within subjects (Time: Phase 1 vs. Phase 2 vs. Phase 3) and the second between subjects (Condition: control vs. labels absent vs. labels challenged). The results parallel the one reported here for the combined index. Labels significantly increase within-boundary similarity (Time effect: $F(2, 216) = 40.21$, $p < .001$, $\eta^2 = .25$; Time by Condition effect: $F(4, 236) = 3.59$, $p < .001$, $\eta^2 = .06$) as well as decrease across-boundary similarity (Time effect: $F(2, 221) = 8.47$, $p < .001$, $\eta^2 = .07$; Time by Condition effect: $F(4, 236) = 3.37$, $p = .01$, $\eta^2 = .05$). Once the labels are removed (regardless if a rationale has been provided) both labeling effects are reduced but not eliminated.

Please cite this article as: Foroni, F., & Rothbart, M., Abandoning a label doesn't make it disappear: The perseverance of labeling effects, *Journal of Experimental Social Psychology* (2012), http://dx.doi.org/10.1016/j.jesp.2012.08.002

that the de-categorization process may be difficult, especially when a set of category labels has a long history of use. To the extent that category labels change the representation of social objects, it may argue in favor of those models of stereotype change based on re-categorization. Re-categorizing outgroup members as ingroup members, for example, through the use of a common, superordinate category label has been effective (e.g., Gaertner & Dovidio, 1986), where a more recent favorable category label can be substituted for an older less favorable label.

## References


Allport, G. W. (1954). *The nature of prejudice.* Cambridge, MA: Addison-Wesley.

Carnaghi, A., Maass, A., Gresta, S., Bianchi, M., Cadinu, M., & Arcuri, L. (2008). Nomina sunt omina: On the inductive potential of nouns and adjectives in person perception. *Journal of Personality and Social Psychology*, *94*, 839–859.

Corneille, O., Klein, O., Lambert, S., & Judd, C. M. (2002). On the role of familiarity with units of measurement in categorical accentuation: Tajfel and Wilkes (1963) revisited and replicated. *Psychological Science*, *13*(4), 380–383.

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *The American Psychologist*, *35*, 867–881.

Eiser, J. R., & Stroebe, W. (1972). *Categorization and social judgment.* New York: Academic Press.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, *Vol. 23.* (pp. 1–74)San Diego, CA: Academic Press.

Foroni, F. (2006). Labeling and categorization: evidence for a mere labeling effect, its modulating factors, and characteristics. *D.A.I: Section B: The Sciences and Engineering*, *66*(7-B), 3997.

Foroni, F., & Rothbart, M. (2011). Category boundaries and category labels: When does a category name influence the perceived similarity of category members? *Social Cognition*, *29*(5), 547–577.

Furnham, A., & Alibhai, N. (1983). Cross-cultural differences in the perception of female body shapes. *Psychological Medicine*, *13*, 829–837.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio, & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). Orlando, FL, USA: Academic Press, Inc.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*, 708–724.

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science*, *10*, 489–493.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education & psychology.* Boston: Allyn & Bacon.

Goldstone, R. L. (1995). Effects of categorization on color perception. *Psychological Science*, *6*(5), 298–304.

Goldstone, R., Lippa, Y., & Shiffrin, R. (2001). Altering object representations through category learning. *Cognition*, *78*(1), 27–43.

Grogan, S. (1999). *Body image: Understanding body dissatisfaction in men, women, and children.* London; New York: Routledge.

Hacking, I. (1995). The looping effects of human kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 351–394). New York, NY: Clarendon Press/Oxford University Press.

Hall, D. G., & Moore, C. E. (1997). Red bluebirds and black greenflies: Preschoolers' understanding of the semantics of adjectives and count nouns. *Journal of Experimental Child Psychology*, *67*(2), 236–267.

Hamilton, D. L., Sherman, S. J., & Ruvolo, C. M. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues*, *46*(2), 35–60.

Harnad, S. (1987). *Categorical perception: The groundwork of cognition* (Ed.). New York, NY, US: Cambridge University Press.

Harnad, S. (2003). *Categorical perception.* Encyclopedia of cognitive science. : Nature Publishing Group/Macmillan.

Kenney, J. S. (2002). Victims of crime and labeling theory: A parallel process? *Deviant Behavior*, *23*(3), 235–265.

Kida, T., Smith, J. F., & Maletta, M. (1998). The effects of encoded memory traces for numerical data on accounting decision making. *Accounting, Organizations and Society*, *5*(6), 451–466.

Kikutani, M., Roberson, D., & Hanley, J. M. (2008). What's in the name? Categorical perception for unfamiliar faces can occur through labeling. *Psychonomic Bulletin & Review*, *15*(4), 787–794.

Kobrynowicz, D., & Biernat, M. (1997). Decoding subjective evaluations: How stereotypes provide shifting standards. *Journal of Experimental Social Psychology*, *33*(6), 579–601.

Levin, R. I. (1987). *Statistic for management* (3rd edition). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Markman, E. M. (1989). *Categorization and naming in children.* Cambridge, MA: MIT Press.

Maruna, S., LeBel, T. P., Mitchell, N., & Naples, M. (2004). Pygmalion in the reintegration process: Desistance from crime through the looking glass. *Psychology, Crime & Law*, *10*, 271–281.

Miller, D. T., & Turnbull, W. (1986). Expectancies and interpersonal processes. In M. R. Rosenzweig, & L. W. Porter (Eds.), *Annual review of psychology*, *Vol. 37.* (pp. 233–256) Palo Alto, CA: Annual Reviews, Inc.

Rosch, E. (1988). Principles of categorization. In A. M. Collins, & E. E. Smith (Eds.), *Readings in cognitive science: A perspective from psychology and artificial intelligence* (pp. 312–322). San Mateo, CA: Morgan Kaufmann, Inc.

Rosenhan, D. L. (1973). On being sane in insane places. *Science*, *179*, 250–258.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom.* Unpublished York: Holt, Rinehart & Winston.

Rothbart, M., Davis-Stitt, C., & Hill, J. (1997). Effects of arbitrarily placed category boundaries on similarity judgments. *Journal of Experimental Social Psychology*, *33*, 122–145.

Rothbart, M., & John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. *Journal of Social Issues*, *41*(3), 81–104.

Snowdon, C. T. (1987). A naturalistic view of categorical perception. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition*. New York: Cambridge Univerity Press.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgment. *The British Journal of Social Psychology*, *54*, 101–114.

Walton, G. M., & Banaji, M. R. (2004). Being what you say: The effect of essentialist linguistic labels on preferences. *Social Cognition*, *22*, 193–213.

Whorf, B. (1956). Language, thought, and reality: Selected writings of Benjamin Lee Whorf. In J. B. Carroll (Ed.), Cambridge, MA: MIT Press.


Please cite this article as: Foroni, F., & Rothbart, M., Abandoning a label doesn't make it disappear: The perseverance of labeling effects, *Journal of Experimental Social Psychology* (2012), http://dx.doi.org/10.1016/j.jesp.2012.08.002