Off the top of my head: Malleability and stability in natural categories

Article *in* Acta Psychologica · February 2018

DOI: 10.1016/j.actpsy.2018.02.002

Tomás Palma
University of Lisbon

Ana Sofia Santos
University of Lisbon

Leonel Garcia-Marques
University of Lisbon

Acta Psychologica 185 (2018) 104–115

Contents lists available at ScienceDirect

Acta Psychologica

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

Off the top of my head: Malleability and stability in natural categories

Tomás A. Palma*, Ana Sofia Santos, Leonel Garcia-Marques

CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal

ARTICLE INFO

Keywords:
Category malleability
Perceptual fluency
Typicality
Perceived relevance
Time pressure

ABSTRACT

Previous research has found that category representations are highly malleable knowledge structures, varying widely across different contexts and individuals. However, it has also been found that such malleability does not apply equally to all types of category information. The present research further investigates the representational malleability versus stability of natural taxonomic categories. Using perceptual fluency as means to induce malleability, we explored whether malleability is moderated by the degree of typicality of category information. In the first experiment, we found that fluency-based malleability only occurs for non-typical category information. In follow-up experiments, we investigated the boundary conditions under which such fluency-based malleability occurs. Namely, in Experiment 2, we showed that the effect of fluency on non-typical features disappeared when there is a sensory modality mismatch between study and test phases. Finally, in Experiment 3, we demonstrated that this effect reappears in the modality mismatch condition when participants are given a response deadline. The implications of these findings to current theories of category representation and the perceptual fluency literature are discussed.

PsycINFO codes

2340 cognitive processes
2343 learning & memory
3040 social perception & cognition

1. Introduction

Categorization is one of our most fundamental cognitive abilities. Without being able to categorize it would be impossible to make sense and manage the vast amount of different information one is confronted with in everyday life. Category representations are thus at the heart of many other cognitive abilities such as the learning of new information, inference-making, problem-solving, judgment, or language use (Barsalou, 2008; Lakoff, 1987; Medin & Smith, 1984).

In the last several decades, research has taught us a lot about the inner-workings of category representations. One aspect of category representations that seems now undisputed is their great malleability depending on the context (for reviews, see Barsalou, 1987, 1993). For example, the representation that one constructs for *eagle* at one occasion (e.g., when watching a wildlife documentary) might include the features *wings*, *long claws*, and *fearless*, whereas the concept constructed at another occasion (e.g., when watching an eagle performing in a bird show at a zoo) might include the features, *wings*, *fast*, and *imposing*.

However, malleability is not unlimited. Some studies have shown that not all category information is equally permeable to contextual influences (e.g., Barsalou, 1982; Garcia-Marques, Santos, & Mackie, 2006; Santos et al., 2012, 2017). That is, category representations seem to possess two broad types of information with different degrees of context dependency: high-accessibility information that is readily activated and incorporated into one's representation of a category, irrespective of the context; and low-accessibility information, which is information that is incorporated into representations only when cued by relevant contexts. These two types of information provide category representations with the capacity of being both malleable and stable.

In the current paper, we further explored these two complementary facets of category representations. For this purpose, we used perceptual fluency (i.e., the experienced ease of processing low-level features that deal with the recognition of the stimulus' physical identity) as a mean to induce malleability in category representations. Perceptual fluency has been shown to be a contextual variable with powerful effects on cognition (for a review, see Alter & Oppenheimer, 2009). Thus, our first goal was to investigate whether one could induce category malleability through perceptual fluency. To our knowledge, only one study has demonstrated effects of fluency on category representations (Oppenheimer & Frank, 2008). Our second goal was then to understand whether fluency-based malleability would extend to all category information or, instead, is limited to low-accessibility information. Our third, and last, goal was to investigate another possible boundary condition of fluency-based category malleability. Namely, we explored

* Corresponding author at: Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisbon, Portugal.
E-mail address: tapalma@psicologia.ulisboa.pt (T.A. Palma).

https://doi.org/10.1016/j.actpsy.2018.02.002
Received 12 July 2017; Received in revised form 24 January 2018; Accepted 7 February 2018

whether for perceptual fluency to cause malleability it would be necessary that people would perceive fluency as a relevant contextual cue (e.g., Westerman, Lloyd, & Miller, 2002).

In the next paragraphs, we start by briefly reviewing relevant evidence for representational malleability and stability of categories. Then, we introduce perceptual fluency as a potential source of malleability. Finally, we present our goal and hypotheses.

1.1. Malleability in category representation

Within a category, not all items are perceived as good examples of that category. Some items are seen as better examples than others. That is, all categories possess a graded structure in which the different members of a category differ with respect to how representative or typical of the category they are (e.g., Armstrong, Gleitman, & Gleitman, 1983; Barsalou, 1983, 1985; Rosch & Mervis, 1975; Rothbart & John, 1985). For example, people usually consider *robin* as a typical member of the natural taxonomic category *birds* and *ostrich* as a non-typical one (e.g., Rosch, 1973). Several different factors have been identified as important determinants of an item's typicality. For example, Rosch and Mervis (1975) showed that an item is typical to the extent that it is similar to the category's central tendency, namely, to the most prototypical items. On the other hand, Barsalou (1983, 1985) showed that, for goal-derived categories, the most typical items are the ones that better serve the goal of the category.

Over the last decades, a large body of evidence accumulates to show that the graded structure of categories is greatly constrained by contexts, goals, and recent experiences. In other words, categories are malleable. For example, it has been shown that different linguistic contexts give rise to different typicality orderings of items within the same category (Roth & Shoben, 1983), or that taking different points of view on a category results in substantially different graded structures for the same category (Barsalou & Sewell, 1984). Recently, it has been shown that even the type of action associated with an item can influence how this item will then be judged regarding typicality. Namely, faces associated with a selection response during a study phase were subsequently judged more typical of their gender category than faces associated with an avoidance response (Ramos, Garcia-Marques, Santos, & Carneiro, 2015). Thus, this research suggests that graded structures are not fixed structures but instead that they reflect people's goals, needs, and experiences. As argued by Barsalou (1985, 1987), changes in the graded structure of a category reflect changes in its current representation. Thus, graded category structure is a crucial variable in the study of category representation.

Further evidence for the malleability of graded structures comes from research looking at the stability of category representations over time, within the same individual. For example, in a study exploring the stability of common categories, McCloskey and Glucksberg (1978) asked participants to make category membership judgments for the same exemplar-category name pairs in two sessions one month apart. Their results showed that participants often changed their minds about whether an exemplar was or was not a member of a category.

But, researchers also obtained evidence for malleability in the graded structure of social categories. In a series of experiments, Garcia-Marques et al. (2006) assessed within-individual stability in the content and use of social category representations across sessions with 2 to 4 weeks apart. Their experiments consistently found moderate correlations between participants' responses (typicality ratings and property verification) in the two sessions. Importantly, they showed that such moderate correlations could be explained by context instability. Namely, when the context was kept constant across two experimental sessions, held 2 weeks apart, there was a high correlation between participants' responses in first and second sessions. However, when the context created in the first session differed from the context of the second session, the obtained correlation dropped substantially.

The research described above shows that graded structures are highly malleable (for more examples, see Barsalou, 1987, 1993). Such malleability is an obvious quality of category representations as it allows them to adaptively respond to a constantly changing environment. However, too much malleability can also be dysfunctional. If category representations were to be completely permeable to contextual influences, it would be impossible for them to ensure some of its main functions such as, for example, saving the social perceiver from coping with infinite information and allow him or her to go beyond the information given (e.g., Crocker, Fiske, & Taylor, 1984; Fiske & Taylor, 1984; Taylor, 1981).

And, indeed, both McCloskey and Glucksberg (1978) and Garcia-Marques et al. (2006), found that not all category content is equally malleable. Namely, McCloskey and Glucksberg (1978) observed that participants had changed their minds more often for non-typical category members (e.g., bookends – furniture) than for typical members (e.g., chair – furniture). Similarly, Garcia-Marques et al. (2006) showed that traits rated in the first session as more central to the category (e.g., dishonest – gypsy) were both more stable across sessions and more likely to be retrieved first than less central traits.

Convergent with the idea that not all category content exhibits the same degree of malleability, it has been shown (Santos et al., 2012) that social category unrelated traits (e.g., friendly - computer programmers) have a higher probability of being chosen as good descriptors of certain target categories if they have been previously subtly primed by an irrelevant task. However, the same was not observed for typical traits (e.g., intelligent - computer programmers), which were always chosen as good descriptors of a category independently of the priming task (see also Santos et al., 2017).

The research just cited suggests that typical category items are more resistant to contextual influences than non-typical items. One possible reason for this may reside on the high accessibility in memory of typical information. Several studies have shown that, for example, typical items are recalled earlier (Mervis, Catlin, & Rosch, 1976; Rosch, 1975; Uyeda & Mandler, 1980) and responded faster (e.g., Fujihara, Nageishi, Koyama, & Nakajima, 1998; Hampton, 1979; Larochelle & Pineu, 1994; McCloskey & Glucksberg, 1979; Rips, Shoben, & Smith, 1973; Smith, Shoben, & Rips, 1974) than non-typical items. According to Barsalou 1982, see also (Barsalou, 1989, 2017), high-accessibility information has a higher likelihood of being incorporated into an individual's current category representation, irrespective of the context, than low-accessibility information. As further suggested (Barsalou & Medin, 1986), information that is highly diagnostic of a category, and information that is functionally important, are likely to be processed frequently and thereby to become more resistant to contextual influences.

As for category information that possesses an overall low-accessibility in memory, it has been suggested that its inclusion in category representations depends largely on whether it is cued by relevant contextual variables or not (Barsalou, 1982). For example, *three legs* is not the type of feature that immediately comes to mind when we read the word *chair* unless it is cued by a relevant context such as modern art museums. Low-accessibility information can either be stored in long-term memory along with other category information or, instead, be inferred online, namely from specific tasks or information currently under one's focus of attention (see Barsalou, 1989).

The research described so far shows that category representations can vary greatly depending on a number of contextual factors. However, such variability or malleability seems to be restricted when a subset of relatively low-accessibility items is made contextually accessible, such as, for example, non-typical or non-central category items (e.g., category-unrelated traits, see Santos et al., 2012). Note that we are not arguing that high-accessibility typical or central items are always included into representations and are always stable. Research, in fact, has shown that also these items are affected by contextual factors (e.g., Garcia-Marques et al., 2006). The point we would like to make is simply that relatively low-accessibility non-typical or non-central items have a higher probability of being affected by contextual factors than high-accessibility items.

1.2. Perceptual fluency as a potential source of malleability

Perceptual fluency has been shown to be a relevant cue for a wide variety of judgments (for a review, see Alter & Oppenheimer, 2009). For example, people typically judge fluent stimuli as more familiar (Jacoby & Dallas, 1981; Jacoby & Whitehouse, 1989; Whittlesea, 1993), more pleasant (Reber, Winkielman, & Schwarz, 1998), more truthful (Reber & Schwarz, 1999), more trustworthy (Alter, Oppenheimer, Epley, & Eyre, 2007; Kelley & Lindsay, 1993), more famous (Jacoby, Kelley, Brown, & Jasechko, 1989; Topolinski & Strack, 2009), more coherent (Topolinski & Strack, 2009), and more likely to be encountered in the future (Palma, Santos, & Garcia-Marques, 2017) than less fluent stimuli.

According to an influential account of fluency, the attributional account (Jacoby, Kelley, & Dywan, 1989; Whittlesea & Williams, 2001a, 2001b; Unkelbach & Greifeneder, 2013), fluency is not directly linked to the dimension that is being judged. In this view, when a stimulus is unexpectedly fluent, people attempt to find a cause to assign the experience of fluency. Thus, fluency is seen as a neutral experience that can be interpreted as familiarity, pleasantness, truth, and so forth depending on a number of contextual and subject-specific variables (for an alternative account, see Winkielman, Schwarz, Fazendeiro, & Reber, 2003; see also T. Garcia-Marques, Mackie, Claypool, & Garcia-Marques, 2004, 2010).

Recently, perceptual fluency has also been linked to judgments about natural taxonomic categories. Namely, in a series of experiments, Oppenheimer and Frank (2008) manipulated the perceptual fluency of exemplar-category (e.g., dog – mammal) name pairs and feature-category name pairs (e.g., barks – dog) using a font type manipulation. Half of the participants received a questionnaire printed in an easy to read font (control condition) while the other half received the same questionnaire printed in a difficult to read font (disfluency condition). Results showed that disfluent exemplars and features were considered less typical of the target categories than its fluent counterparts.

The results obtained in Oppenheimer and Frank's (2008) research suggest that people relied on perceptual fluency as a useful cue for deciding on the degree of typicality of exemplars and features for a given category. These results are thus the first demonstration that perceptual fluency causes malleability in the graded structure of categories and, therefore, on category representations. An open question though is whether all category items are affected by perceptual fluency. Although the authors report that category items ranged greatly in typicality, they did not report any analysis separating the typical from the non-typical items. Therefore, one does not know whether this fluency-based malleability applies equally to all category items or, instead, is mainly driven by non-typical or non-central items (see Garcia-Marques et al., 2006Santos et al., 2012).

In fact, there is already some evidence in the literature suggesting that the influence of perceptual fluency on categorization judgments does not apply to all category items. Using a category-learning task, Miles and Minda (2012) presented participants with a study phase consisting of a set of distortions of a prototype shape. They were told that these shapes belonged all to the same category. In the test phase, participants saw a new set of distortions of the prototype and a new set of unrelated distortions (created from another shape) and had to respond whether each shape was a member of the previously learned category. Critically, perceptual fluency of the shapes was enhanced through repetition priming (Jacoby & Whitehouse, 1989). That is, in some of the trials, the target shapes were subtly primed with identical shapes while in other trials there was no priming. Results showed that unrelated shapes primed with identical shapes (i.e., fluent shapes) were judged as better category members than unprimed shapes. No effect of fluency was observed for category members (distortions of the prototype).

1.3. The present research

The present research aimed at further investigating the role of perceptual fluency (enhanced through repetition priming; Jacoby & Whitehouse, 1989) on the malleability of natural taxonomic categories (Oppenheimer & Frank, 2008). Specifically, in the first experiment, we investigated whether repetition priming effects are moderated by the degree of typicality of category items. Based on previous research (e.g., Garcia-Marques et al., 2006; Miles & Minda, 2012; Santos et al., 2012), we predicted that only non-typical items should be affected by perceptual fluency induced by repetition priming. Unlike typical category items, which are easily retrieved from long-term memory and included in current category representations, non-typical items are highly dependent on whether they are made relevant by contextual variables such as perceptual fluency. In the second and third experiments, we sought to investigate the boundary conditions to these effects. In Experiment 2, we tested whether a necessary pre-condition for repetition priming to affect judgments, namely perceived relevance, also need to be to induce category malleability. Drawing on previous research (Westerman et al., 2002), we decreased the perceived relevance of the perceptual fluency by manipulating the match between the sensory modalities of a study phase and a test phase. We hypothesized that the repetition priming effects on non-typical items would be minimized or even eliminated when there was a sensory modality mismatch between study and test. Finally, in Experiment 3, we investigated whether time pressure potentiates the use of perceptual fluency on categorization judgments, even when the relevance of perceptual fluency is decreased by the modality mismatch manipulation.

2. Experiment 1

The goal of Experiment 1 was to test our hypothesis that the influence of perceptual fluency on categorization judgments depends on the degree of typicality of category items. For this purpose, we used a robust perceptual fluency manipulation - repetition priming (e.g., Brown & Marsh, 2009; Huber, Clark, Curran, & Winkielman, 2008; Jacoby & Whitehouse, 1989; Palma et al., 2017). Participants were presented with typical and non-typical category features names and had to decide whether these features were characteristic (typical) of the category that preceded them. We predicted an interaction between feature type and prime type. That is, we expected to observe an effect of perceptual fluency on non-typical features, with fluent features being judged more typical of the category than non-fluent features, but not on typical features.

2.1. Method

2.1.1. Participants

Forty-seven students from the University of Lisbon (34 women and 13 men) participated in this experiment. They were paid in exchange for participation. Their mean age was 23.51 years ($SD = 5.15$; one participant did not report her age).

In all experiments, participants gave informed consent prior to participation. The Ethics Committee of the Faculty of Psychology of the University of Lisbon approved this research. We report all data exclusions, all manipulations, and all measures. The data from all studies are available at the Open Science Framework website (https://osf.io/gtrph).

2.1.2. Materials and procedures

The stimuli were a pool of 15 animal names and 60 features names (see Appendix A). Half of these features were true and typical (e.g., *bat – blind*) and half were true but less typical, that is non-typical (e.g., bat – *swims*). Each animal exemplar had two features of each type. These features were selected so that they were unique to the target animals. Typical features were selected from a semantic feature norm corpus

(McRae, Cree, Seidenberg, & McNorgan, 2005), based on two measures of feature informativeness. Firstly, we selected the most *distinguishing* features, that is, the ones that were associated with only one or two animals. Secondly, we checked the degree of *distinctiveness* (a measure that reflects the continuum from truly distinguishing to highly shared) and chose the most distinctive ones. As for the non-typical features, they were selected from relevant books and manuals on animal anatomy and physiology. We avoided features with more than one word. Furthermore, we checked if these non-typical features did not appear in the above-mentioned norms (McRae et al., 2005), to assure that they were not at some extend spontaneously produced by normal adults. Thus, these features belonged to the target animals, but they are not characteristic of them.

The experiment was run in sessions of up to 8 participants at a time. Participants were seated in front of 60 Hz CRT computer screens in individual workstations. All instructions and stimuli were presented on the computer. Initial instructions informed participants that they would be presented with several different pairs of animal-feature and that, for each pair, they should decide whether the feature was characteristic (typical) to the animal in question or not. Each trial started with a fixation point (+) displayed in the center of the screen for 1000 ms. Then, an animal name appeared embedded in a mask (e.g., xxxx bat xxxx) during 1000 ms being immediately followed by a prime feature (e.g., swims). The prime was shown for 30 ms. This presentation time is in line with previous studies (e.g., Lloyd, Westerman, & Miller, 2003; Palma et al., 2017; Westerman et al., 2002). After the prime, the animal name that had been presented before was again displayed (e.g., xxxx bat xxxx) for another 1000 ms, giving the impression that it had never left the screen. Finally, the prime feature was again presented as a target and remained on the screen until a response was made. Participants were instructed to respond "yes" (key P) if they thought the feature was typical to the previously presented animal and "no" (key Q) if they thought the feature was not typical of the previously shown animal. Each animal name was presented four times, two times paired with a typical feature and two times with a non-typical feature. The order of the pairs was randomized for each participant. Upon completing the task, experimenters thanked and debriefed participants.

2.2. Results

In this and the following experiments, we report partial eta-squares as effect sizes for omnibus ANOVA. For planned contrasts involving a single degree of freedom in the F-ratio's numerator, we report Cohen's *d* and the respective 95% confidence interval (CI 95%).[1] We adopted a significance criterion of 0.05 in all statistical tests.

A 2 (feature type: typical, non-typical) × 2 (prime type: primed, not primed) repeated-measures analysis of variance (ANOVA) was performed on the proportion of "yes" responses. All cell means are shown in Table 1. This analysis revealed a large main effect of feature type, $F(1, 46) = 1175.70$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.96$, showing that typical features were considered better category members ($M = 0.90$, $SD = 0.11$) than non-typical features ($M = 0.24$, $SD = 0.16$). The main effect of prime type did not reach statistical significance, $F(1, 46) = 3.27$, $MSE = 0.01$, $p = 0.077$, $\eta_p^2 = 0.07$. There was also a significant interaction between feature type and prime type, $F(1, 46) = 4.96$, $MSE = 0.01$, $p = 0.031$, $\eta_p^2 = 0.10$. Planned contrasts showed that while non-typical features were judged more frequently as typical of the category, $F(1, 46) = 8.38$, $MSE = 0.01$, $p = 0.006$, $d = 0.37$ 95% CI [0.08, 0.67], typical features were not influenced by the priming manipulation, $F(1, 46) = 0.02$, $MSE = 0.01$, $p = 0.898$, $d = 0.02$ 95% CI [− 0.26, 0.31].

[1] Cohen's *d* and 95% CI were calculated based on the formulas provided by Cooper, Hedges, and Valentine (2009, p. 228–229). These formulas correct the effect size for within-participant comparisons with the correlation between measures.

Table 1
Mean proportion (and standard deviations) of "yes" responses as a function of feature type and prime type (Experiment 1).

	Feature type	
Prime type	Typical	Non-typical
Primed	0.90 (0.11)	0.27 (0.15)
Unprimed	0.90 (0.09)	0.22 (0.13)

2.3. Discussion

Experiment 1 showed that repetition priming has an effect used as a valid contextual cue to decide upon the category typicality of non-typical but not of typical items, which is in accordance with our initial hypothesis. Thus, these results can be seen as an extension of the results previously obtained by Oppenheimer and Frank (2008) in the sense that they show that the role of perceptual fluency on category malleability is restricted to non-typical category items.

3. Experiment 2

In Experiment 1, we showed that indeed perceptual fluency affects feature typicality judgments, but only of non-typical items. In the second experiment, we explored the limits of this effect, namely, whether perceptual fluency always influences non-typical items.

Although perceptual fluency effects are widely documented in the literature, previous research has also shown that there are certain factors that discourage the use of fluency on judgments (e.g., Gallo, Perlmutter, Moore, & Schacter, 2008; Lloyd et al., 2015; Luka & Barsalou, 2005; Westerman, 2001; Westerman et al., 2002; Westerman, Miller, & Lloyd, 2003; Whittlesea & Williams, 1998; Wolk, Gold, Signoff, & Budson, 2009). One of such factors is the perceived relevance of perceptual fluency for the judgment at hand. Namely, it has been demonstrated that fluency only affects responses when people perceive it as a relevant cue to the judgments at hand, such as when the same perceptual modality is used both for study and test (e.g., Gallo, Weiss, & Schacter, 2004; Miller, Lloyd, & Westerman, 2008; Thapar & Westerman, 2009; Westerman et al., 2002; Westerman et al., 2003).

For example, research by Westerman et al. (2002) tested the hypothesis that people should interpret fluency as a relevant cue for the previous experience if they are being tested in the same sensory modality in which they previously experienced the test items. On the other hand, participants should not interpret fluency as a valid cue for their previous experience when the test items were experienced in a different modality. This hypothesis was based on the assumption that people's reliance on perceptual fluency as a cue to memory stems from their knowledge that previously experienced stimuli are perceived more fluently than are novel stimuli (see Jacoby & Dallas, 1981). Across several studies, they found that when the sensory modality of the study and test phases matched (i.e., were both visual), the perceptual fluency manipulation affected judgments. However, when the modalities mismatched, there was no effect of fluency.

Interestingly, Westerman et al. (2002) showed that this interaction between sensory modalities and perceptual fluency persisted even when there were no real target items presented during the study phase. Namely, in their third experiment, these authors manipulated the modality of a counterfeit study list (visual vs. auditory). The counterfeit list was supposed to comprise subliminal test items according to what participants were told, however, it was composed only of visual or auditory noise. In the recognition test, always presented visually, participants had to guess which words appeared on the counterfeit list. Some of the words were made fluent through repetition priming (Jacoby & Whitehouse, 1989). In the condition where the study and test modalities mismatched, results showed that the J-W illusion was greatly reduced.

In a related vein, Lloyd et al. (2003) told their participants that some of the words to be presented in a recognition test had been subliminally shown in a previous (counterfeit) study list. To half of the participants, they told each word appeared five times while to the other half they told the words appeared only once. The goal of these different instructions was to manipulate participants' expectations about the amount of perceptual fluency that should be associated with test words. The results showed that the influence of the priming manipulation (embedded on the recognition test) was greater on the latter than the former group, supposedly because these participants were expecting a high amount of fluency to accompany targets, and the priming manipulation was too weak to influence their responses significantly.

Although the mechanism responsible for these modality results is not yet entirely clear, some research points to discounting, namely participants discount or disregard fluency in the modality mismatching conditions (Kurilla & Gonsalves, 2012). Thus, taken together, the findings describe above provide strong support to the argument that people develop expectations for the amount of perceptual fluency associated with a prior study episode and that these expectations moderate the interpretation that subjects adopt for the source of enhanced fluency.

In Experiment 2, we extended Westerman et al.'s (2002) ideas to the context of our task and tested whether manipulating the perceived relevance of perceptual fluency would impact the typicality judgments of typical and non-typical features. For this purpose, we manipulated the sensory modality of a counterfeit study list. We told half of the participants that they would be presented with a list of features that people, in general, consider typical of animals. For some of these participants, the counterfeit study list was visual, while for others it was auditory. They were then given a visual test in which they were presented with primed and unprimed features and had to decide whether these belonged or not to the animals in question. We predicted that, when the counterfeit list was presented visually, participants should interpret the perceptual fluency (due to repetition priming) associated with some non-typical features as a sign that those features were shown in the previous study list composed by typical animal features. As a result, they should be positively influenced by perceptual fluency ("if it feels fluent then it might have been presented in the visual list composed by typical features"). On the other, when the counterfeit list was presented auditorily, people should not attribute the fluency associated with some non-typical features to the study list because this list was presented in a different modality. Thus, fluency should be partially or totally discounted from their judgments. We did not predict typical features to be influenced by fluency nor modality.

To the other half of the participants, we asked for judgments of familiarity instead of typicality. That is, we asked them to guess whether the target words (features) had been presented in the (counterfeit) study list or not, exactly as in Westerman et al. (2002; Experiment 3). In this condition, we expected to replicate Westerman et al., namely people should interpret fluency as a sign of previous experience only when there was a match between study and test modalities.

3.1. Method

3.1.1. Participants

One hundred and sixty-two University of Lisbon students participated in this experiment for course credit (104 women and 58 men). Their mean age was 21.78 years ($SD = 6.81$). Participants were randomly assigned to one of four between-participants conditions (n's ranging between 40 and 41).

3.1.2. Materials and procedures

The stimuli were the same of Experiment 1. All instructions and stimuli were presented on the computer. The experiment consisted of two phases. In the first phase, participants were presented with a counterfeit study list. In the modality-match conditions, the counterfeit list was visual. Participants were told that they were going to be presented with a subliminal study list. They were further told that these words would appear embedded in visual noise and would be presented so quickly they would not be able to detect them; however, they should keep their eyes on the screen during the entire presentation, even if they could not identify the words. They saw 60 presentations of a visual mask (#$%&/#"&$#). In the categorization condition, participants were told that the words corresponded to features people, in general, consider typical of animals. In the familiarity condition, nothing was told about the nature of the words. In the modality-mismatch conditions, the counterfeit list was auditory. Participants in these conditions were told they were going to hear several words embedded in static noises, similar to radio static. These noises were created from the conversion of real word sound files (personality traits not included in the judgments task; e.g., smart, strong, meticulous) into static noise sound files, using a specialized software package (Voice processor Alex, a text-to-speech system, from Mac OS X Leopard, 2007). They heard 60 noises played through headphones with 2 s of delay between each. The noises had between 388 ms and 1049 ms of duration. In the categorization condition, they were told that each noise corresponded to a different feature that people, in general, consider being typical of certain animals. In the familiarity condition, they were only told each noise corresponded to a different word.

In the second phase, a judgment task, consisting of 60 trials, immediately followed the presentation of counterfeit study list. In the categorization condition, the task was the same used in Experiment 1. In the familiarity condition, the repetition priming task had some minor differences comparing to repetition priming task used in categorization condition. Participants first saw a blank screen for 1500 ms followed by a fixation point (+) for 1000 ms. Then, a pre-mask consisting of a row of number signs (#########) appeared for 250 ms, followed by a prime feature for 30 ms. A post-mask was then displayed for 250 ms. Following a 50 ms blank screen, the target feature was presented. Participants were asked to respond whether they thought the target feature had been presented in the subliminal study list or not. In this condition, they were told that some words had been presented in the list and that some were new. They were further instructed that the words presented subliminally may "feel familiar" and that they should respond "yes" to those words, even if they did not remember having seen them. Upon completing the task, experimenters thanked and debriefed participants.

3.2. Results

A preliminary analysis revealed that six participants from the familiarity judgment condition responded "no" to all items, and, therefore, their data were not included in the main analyses. To facilitate the interpretation of the results, we performed separate ANOVAs for each condition (categorization and familiarity).

For the categorization condition, a 2 (list modality: visual, auditory) × 2 (feature type: typical, non-typical) × 2 (prime type: primed, not primed) mixed model ANOVA (see Table 2 for all cell means) revealed a large main effect of the type of feature, $F(1, 80) = 1314.50$, $MSE = 0.02, p < 0.001, \eta_p^2 = 0.94$, showing that typical features were

Table 2
Mean proportion (and standard deviations) of "yes" responses in the typicality judgment condition as a function of modality, feature type, and prime type (Experiment 2).

Prime type	Visual modality		Auditory modality	
	Feature type		Feature type	
	Typical	Non-typical	Typical	Non-typical
Primed	0.88 (0.11)	0.28 (0.12)	0.87 (0.11)	0.24 (0.15)
Unprimed	0.89 (0.10)	0.21 (0.11)	0.85 (0.15)	0.24 (0.15)

indeed considered more typical ($M = 0.87$, $SD = 0.10$) of the animals in question than non-typical features ($M = 0.24$, $SD = 0.11$). The main effect of prime type was statistically significant, $F(1, 80) = 4.09$, $MSE = 0.03$, $p = 0.046$, $\eta_p^2 = 0.05$, with more "yes" responses overall to primed ($M = 0.57$, $SD = 0.09$) than unprimed features ($M = 0.55$, $SD = 0.09$). There was a significant three-way interaction between list modality, feature type, and prime type, $F(1, 80) = 5.67$, $MSE = 0.01$, $p = 0.020$, $\eta_p^2 = 0.07$. As a reminder, we predicted an interaction between feature type and prime type in the visual modality condition but not on the auditory modality condition. Planned contrasts indeed showed the predicted interaction in the visual modality condition, $F(1, 80) = 5.76$, $MSE = 0.01$, $p = 0.019$, $\eta_p^2 = 0.07$. That is, priming had a significant effect on non-typical, $F(1, 80) = 9.58$, $MSE = 0.01$, $p = 0.003$, $d = 0.54$, 95% CI [0.21, 0.88], but not on typical features, $F(1, 80) = 1.62$, $MSE = 0.01$, $p = 0.207$, $d = 0.23$, 95% CI [−0.08, 0.54]. Regarding the auditory modality condition, the interaction between feature type and prime type was not statistically significant, ($F < 1$, $p = 0.336$). See Table 2.

For the familiarity condition, a 2 (list modality: visual, auditory) × 2 (feature type: typical, non-typical) × 2 (prime type: primed, not prime) mixed ANOVA was also conducted on the number of affirmative responses (see Table 3 for all cell means). This analysis yielded two significant main effects and one interaction. There was a main effect of the type of feature, $F(1, 71) = 13.77$, $MSE = 0.01$, $p < 0.001$, $\eta_p^2 = 0.16$, with typical features being more recognized ($M = 0.42$, $SD = 0.19$) than non-typical ones ($M = 0.37$, $SD = 0.15$). The main effect of prime type was also significant, $F(1, 71) = 17.25$, $MSE = 0.08$, $p < 0.001$, $\eta_p^2 = 0.19$. Overall, primed features were given more "yes" responses ($M = 0.46$, $SD = 0.23$) than unprimed features ($M = 0.33$, $SD = 0.20$). This main effect was though qualified by an interaction with the counterfeit list modality, $F(1, 71) = 10.64$, $MSE = 0.08$, $p = 0.002$, $\eta_p^2 = 0.13$. Planned contrasts showed that the effect of the prime on familiarity judgments only occurred in the condition where the study list was presented visually, $F(1, 71) = 27.88$, $MSE = 0.08$, $p < 0.001$, $d = 1.14$, 95% CI [0.67, 1.61], and not in the condition where the study list was presented auditory, $F(1, 71) = 0.40$, $MSE = 0.08$, $p = 0.534$, $d = 0.10$, 95% CI [−0.22, 0.42]. No other interactions or main effects were significant.

3.3. Discussion

The results of our second experiment extend the findings obtained in the first experiment in that they showed that the influence of perceptual fluency on category membership judgments of non-typical features is moderated by modality congruence. Namely, people used fluency to judge non-typical features only when study and test modalities were congruent. In the familiarity judgment condition, the results successfully replicated previous findings by Westerman et al. (2002; Experiment 3). This replication is also worthwhile highlighting as it is, to our knowledge, the first one obtained by an independent lab.

4. Experiment 3

The results of Experiment 2 showed that perceptual fluency is sensitive to a modality mismatch between the different experimental phases. Namely, participants discounted the fluency, created by repetition priming, from typicality judgments when the counterfeit study list was presented auditorily. In the present experiment, we explored the conditions under which such discounting process takes place, that is, whether it occurs whenever study and test phases mismatch.

Research by Westerman et al. (2002; Experiment 2) had already shown that participants do not always discount visual fluency from their decisions when previously exposed to an auditory study list. Namely, when the modality of the study list was manipulated within-participants (i.e., participants were presented with two study lists, one presented visually and one auditorily), the interaction between study list modality and perceptual fluency did not occur. These results suggest that as long as some modality mismatch is not diagnostic at the test, participants will not disregard perceptual fluency and therefore be influenced by it.

Importantly, Gallo et al. (2008) suggested that this perceptual mismatch effect may not be due to the reliance on a post-retrieval monitoring strategy, that is, to an evaluation of the information that comes to mind concerning their relevance for current goals. Instead, the effect might be caused by an overall pre-retrieval orientation strategy to avoid perceptual fluency as a cue when modality mismatch occurs. Thus, contrary to a post-retrieval monitoring process, which would occur on an item-by-item basis, ignoring the fluency coming from priming can be considered a pre-retrieval orientation process adopted for the entire test. This possibility was based on a study in which Gallo et al. (2008) examined the role of the distinctiveness heuristic on the use perceptual fluency. The distinctiveness heuristic is a strategy of responding based on distinctive details of an experience people expect to remember. When a novel item lacks the expected distinctive information, individuals use this absence of distinctive information to reject the item. Thus, the distinctiveness heuristic is a kind of post-retrieval monitoring strategy for reducing misattribution errors.

In their study, these authors asked participants to study a list of words for a subsequent memory test. In one condition each word was paired with a picture (mismatch condition; for a similar manipulation, see Westerman et al., 2003), while in the other condition each word was followed by the same word spoken over the computer speakers. After this study phase, participants were given a recognition test in which they had to decide if the words had been presented before or not. Some of the words (targets and lures) were made fluent trough repetition priming (Jacoby & Whitehouse, 1989). Critically, half or the participants were given a response deadline. They predicted that participants would expect more distinctive recollections from pictures and, therefore, would rely on this distinctiveness heuristic to make their judgments rather than on perceptual fluency. Results showed that fluency effects on familiarity judgments were substantially attenuated in the mismatch condition when participants had time to decide, but, also when participants were under time pressure. The results in the latter condition were surprising because the distinctiveness heuristic is assumed to be a time-consuming process and thus it should only occur when participants are not under time pressure. This unexpected result led the authors to conclude that the attenuation of the perceptual fluency effects in the mismatch condition might have occurred through some process other than a post-retrieval process.

Note, however, that results of Gallo et al. were obtained on recognition and not on typicality judgments. It is therefore important to examine whether the effects of modality mismatch on typicality judgments are more compatible with a post-retrieval monitoring or with a pre-retrieval orientation strategy.

Moreover, a large body of research has shown that manipulations that place participants under sub-optimal processing conditions, such as a concurrent task, tend to amplify people's reliance on subjective experiences, such as perceptual fluency, over other possible sources of information (e.g., Benjamin & Craik, 2001; De Houwer & Smith, 2013; Jacoby, 1991; Jacoby, Woloshyn, & Kelley, 1989).

Table 3
Mean proportion (and standard deviations) of "yes" responses in the familiarity judgment condition as a function of modality, feature type, and prime type (Experiment 2).

Prime type	Visual modality		Auditory modality	
	Feature type		Feature type	
	Typical	Non-typical	Typical	Non-typical
Primed	0.52 (0.28)	0.48 (0.26)	0.46 (0.24)	0.40 (0.20)
Unprimed	0.26 (0.22)	0.24 (0.21)	0.44 (0.25)	0.36 (0.20)

Drawing on the research described above, in Experiment 3, we investigated whether time pressure moderates the modality mismatch effect obtained in Experiment 2. If putting participants under time pressure diminishes their capacity to use other sources of information to inform judgments, then one can predict an effect of perceptual fluency on typicality judgments not moderated by the modality of the counterfeit study list. That is, when asked to respond quickly, participants should not be able of easily dismissing perceptual fluency as irrelevant to the typicality judgments in the auditory modality. As a result, they should be influenced by repetition priming. Thus, in this condition (speeded response condition), we expected a main effect of repetition priming and no interaction with the modality. Conversely, when not asked to respond quickly (non-speeded response condition), participants should not be affected by repetition priming in the auditory modality condition, precisely as in Experiment 2.

4.1. Method

4.1.1. Participants

One hundred and fifty-nine University of Lisbon students participated in this experiment in return for course credit or money (97 women and 62 men). Their mean age was 22.74 years (*SD* = 8.07; one participant did not report his age). They were randomly assigned to one of four between-participants conditions (39 participants in the auditory non-speeded condition and 40 in each of the remaining conditions).[2]

4.1.2. Materials and procedures

The stimuli were the same used in the previous two experiments and the procedures followed those described in Experiment 2, with two exceptions. First, in this experiment, all participants made category membership judgments. Second, we manipulated the response deadline. Namely, half of the participants were instructed to respond as quickly as possible. They were told they should try to make the category membership judgments within 1 s. When a response took longer than 1 s, the message "Respond Now!" appeared on the screen. This message remained on the screen until a response was given. Thus, the computer recorded the responses given before and after the 1 s deadline. For the other half of the participants, no response deadline was given (like in the previous experiments). Upon completing the task, experimenters thanked and debriefed participants.

4.2. Results

An independent *t*-test confirmed that participants in the speeded response condition were faster (*M* = 1407 ms, *SD* = 682 ms) to judge the features than participants in the non-speeded response condition (*M* = 2112 ms, *SD* = 959 ms), $t(157) = -5.34$, $p < 0.001$. To facilitate the interpretation of the results, we performed separate ANOVAs for the non-speeded response condition and the speeded-response condition.

For the non-speeded response condition, a 2 (list modality: visual, auditory) × 2 (feature type: typical, non-typical) × 2 (prime type: primed, not primed) mixed model ANOVA (see Table 4 for all cell means) showed a large main effect of feature type, $F(1, 77) = 1571.40$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.95$, with typical features being considered more typical (*M* = 0.88, *SD* = 0.15) than non-typical features (*M* = 0.23, *SD* = 0.16). The main effect of prime type was also significant, $F(1, 77) = 6.26$, $MSE = 0.005$, $p = 0.014$, $\eta_p^2 = 0.07$. Overall, participants gave more "yes" responses to primed (*M* = 0.56, *SD* = 0.13) than unprimed features (*M* = 0.54, *SD* = 0.13). There was also a significant interaction between list modality and prime type, $F(1, 77) = 5.74$, $MSE = 0.005$, $p = 0.019$, $\eta_p^2 = 0.07$. Planned contrasts showed that priming had an effect on typicality judgments in the visual modality condition, $F(1, 77) = 12.15$, $MSE = 0.005$, $p < 0.001$, $d = 0.39$, 95% CI [0.07, 0.70], but not in the auditory modality condition ($F < 1$, $p = 0.940$. Surprisingly, the three-way interaction between list modality, feature type, and prime type was not statistically significant ($F < 1$, $p = 0.719$). In order to understand why this interaction did not reach significance, we performed planned contrasts within each modality condition. In the visual modality condition, there was a main effect of prime type, $F(1, 78) = 12.27$, $MSE = 0.005$, $p = 0.001$, $\eta_p^2 = 0.14$ but no interaction between this factor and feature type ($F < p = 0.528$). This means that repetition priming occurred both for typical and for non-typical features, in contrast to the results of the previous studies in which repetition priming was only effective for non-typical features. We will address this discrepancy later. In the auditory modality condition, there was no main effect of prime type ($F < 1$, $p = 0.940$) nor an interaction between this factor and feature type ($F < 1$, $p = 0.908$). See Table 4.

Table 4
Mean proportion (and standard deviations) of "yes" responses in the non-speeded response condition as a function of modality, feature type, and prime type (Experiment 3).

Prime type	Visual modality		Auditory modality	
	Feature type		Feature type	
	Typical	Non-typical	Typical	Non-typical
Primed	0.88 (0.13)	0.25 (0.13)	0.89 (0.13)	0.23 (0.13)
Unprimed	0.83 (0.12)	0.22 (0.12)	0.89 (0.12)	0.23 (0.12)

For the speeded response condition, a 2 (list modality: visual, auditory) × 2 (feature type: typical, non-typical) × 2 (prime type: primed, not prime) mixed ANOVA revealed a main effect of list modality, $F(1, 78) = 4.53$, $MSE = 0.03$, $p = 0.036$, $\eta_p^2 = 0.05$, showing more affirmative responses in the auditory modality (*M* = 0.57, *SD* = 0.09) than the visual modality (*M* = 0.52, *SD* = 0.09). The main effect of feature type was again highly significant, $F(1, 78) = 1042.42$, $MSE = 0.03$, $p < 0.001$, $\eta_p^2 = 0.93$; typical features were indeed judged as more typical (*M* = 0.84, *SD* = 0.16) of the animals in question than non-typical features (*M* = 0.25, *SD* = 0.19). The expected main effect of prime type was also significant, $F(1, 78) = 10.47$, $MSE = 0.01$, $p = 0.002$, $\eta_p^2 = 0.12$, showing that primed features were given more "yes" responses (*M* = 0.56, *SD* = 0.14) than unprimed features (*M* = 0.53, *SD* = 0.15). There was no interaction between feature type and prime type ($F < 1$, $p = 0.334$) nor interaction between these two factors and list modality ($F < 1$, $p = 0.969$). See Table 5.

4.3. Discussion

The main goal of this experiment was to investigate whether speeding typicality judgments would undermine the effect of modality mismatch on the efficiency repetition priming. And, indeed, under a short response deadline, the modality mismatch between study and test ceased to moderate the efficacy of repetition priming. This result contrasts with the results obtained by Gallo et al. (2008) as these authors found a reduction of repetition priming in the mismatch condition under time pressure. Note, however, that this reduced effect of repetition priming under time pressure was obtained with recognition judgments while ours was with typicality judgments. In our case, to adopt an overall pre-retrieval orientation to entirely avoid fluency as a cue for typicality would be less functional because conceptual fluency (i.e., the experienced ease that accompany meaning extraction and the abstract categorization of stimuli) is an intrinsic and highly diagnostic characteristic of typical items. This is not the case for an episodic memory

[2] We aimed at collecting data from 160 participants (40 in each condition), however, due to a mistake (we counted the same participant twice) we ended up with one participant less in one of the conditions. We noticed this issue only when we were analyzing the data and, therefore, we decided not to run an extra participant.

Table 5
Mean proportion (and standard deviations) of "yes" responses in the speeded response condition as a function of modality, feature type, and prime type (Experiment 3).

Prime type	Visual modality		Auditory modality	
	Feature type		Feature type	
	Typical	Non-typical	Typical	Non-typical
Primed	0.83 (0.12)	0.25 (0.15)	0.88 (0.12)	0.30 (0.15)
Unprimed	0.81 (0.14)	0.21(0.15)	0.85 (0.14)	0.25 (0.15)

task (as the one of Gallo et al.) in which the study list was composed of arbitrary items. And, in fact, feature typicality always made a difference for typicality judgments both under modality match and mismatch conditions. Thus, in our case, the effects of modality mismatch seem to correspond to a post-retrieval strategy (carried out at test on item-per-item basis) by which, under modality match, features felt as fluent either conceptually and/or perceptively fluent are indiscriminately judged as typical whereas, under modality mismatch, only conceptually fluent features (but not perceptively) are accepted as typical.

Under time pressure, however, participants would only be able to discriminate between conceptual and perceptual fluency and repetition priming would again emerge under modality mismatch.

One surprising result was that, unlike with what happened in our previous experiments, repetition priming effects emerged with typical features (but were apparently weaker than the same effects found with non-typical features). To address this discrepancy and assess the reliability of this effect we performed a meta-analysis.

5. Meta-analysis

The main hypothesis tested in the present set of experiments was that perceptual fluency influences category membership judgments, however such influence is moderated by the degree of typicality of the category features. Namely, fluency should influence the judgments of non-typical but not of typical category features. We found support for this hypothesis in the first two experiments, but not in the third. In the latter experiment, perceptual fluency affected both types of features. Thus, in order to obtain a more precise estimate of the influence of perceptual fluency, we conducted a meta-analysis on three key findings: (a) the interaction between feature type and prime type (contrast codes: 1, −1, 1, −1); (b) the comparison between primed versus not primed non-typical features (contrast codes: 0, 0, 1, −1); and (c) the comparison between primed versus not primed typical features (contrast codes: 1, −1, 0, 0). We used a statistical tool developed by McShane and Böckenholt (2017) for single-paper meta-analyses (SPM).[3] See Appendix B for a brief explanation of this tool.

Thus, we entered in the SPM tool the relevant data (i.e., proportions, standard deviations, covariances, and sample sizes) from each experiment. Namely, on Experiment 1 we selected the all data ($N = 47$). From Experiment 2, we retrieved the data concerning the categorization judgments on the visual modality condition ($N = 41$), because this was the condition where we expected an interaction between feature type and prime type. Finally, from Experiment 3, we retrieved the data of the non-speeded judgments on the visual modality condition because ($N = 40$), again, this was the only condition where we predicted to obtain our interaction. The SPM tool yielded a significant interaction contrast (Estimate = 0.07; 95% CI: 0.02, 0.11; $SE = 0.02$; $z = 3.06$, $p = 0.002$). Consistent with our hypothesis, the priming contrast was significant for the non-typical features (Estimate = 0.05; 95% CI: 0.02, 0.08; $SE = 0.02$; $z = 3.00$, $p = 0.003$) but not for the typical features (Estimate = 0.02; 95% CI: −0.01, 0.05; $SE = 0.01$; $z = 1.22$, $p = 0.222$). Additionally, I^2 was estimated at 63.24% (95% CI: 19.75%–81.28%), suggesting a moderate heterogeneity (see Higgins & Thompson, 2002; Higgins, Thompson, Deeks, & Altman, 2003; Pigott, 2012).

The results of the meta-analysis thus suggest that, taken together, the three experiments provided reliable results in support our main hypothesis that fluency-based malleability is more pronounced for non-typical than to typical category features.

6. General discussion

In the current set of experiments, we investigated whether perceptual fluency (enhanced through repetition priming) can cause category malleability and whether such malleability is total or limited to a subset of category items. In Experiment 1, we tested the hypothesis that repetition priming would influence typicality judgments but only of non-typical category items. The results supported this hypothesis. In Experiment 2, we investigated whether decreasing the perceived relevance of perceptual fluency would affect its influence on judgments of typicality. For this purpose, we manipulated the modality congruence between a study and a test phase (see Westerman et al., 2002). We again found an effect of repetition priming on non-typical but not on typical items, thus replicating our previous results. Importantly, this effect was moderated by modality congruence. Namely, it only occurred when there was a match between study and test modalities. When there was a mismatch, both types of features were immune to repetition priming, supposedly because people did not expect a previous auditory learning episode to be relevant for the visual test.

Finally, in Experiment 3, we explored the role of time pressure in combination with the modality manipulation employed in Experiment 2. Previous research has shown that people are more prone to use perceptual fluency when the task includes a short response deadline. (e.g., Westerman, 2001). However, Gallo et al., 2008 found that the effect of modality mismatch still occurred under time pressure and suggested that modality mismatch results were due to a pre-retrieval orientation (that is, a strategic decision regarding overall test performance, in this case, a decision to ignore fluency as cue). As Gallo et al., 2008 used recognition judgments and we used typicality judgments, it was important to test the effect of speeding typicality responses. Experiment 3 explored this question. Our results showed that the modality mismatch ceased to moderate the effects of repetition priming. Namely, repetition priming influenced judgments of typicality on both modality conditions. Importantly, this effect did not occur when participants were not under time pressure.

Our results are thus incompatible with a pre-retrieval global orientation to avoid fluency as a cue (Gallo et al., 2008). Instead, our results are better explained by the effect of a retrieval response strategy based on indiscriminate fluency when study and test sensory modalities matched. That is, a response strategy that conflates perceptual and conceptual fluency as a cue for responding. We propose that when there was a modality match, for the non-typical features, participants relied only on the perceptual fluency derived from repetition priming whereas for the typical features, participants relied both on the conceptual fluency associated with these features and on the perceptual fluency stemming from priming. When there was a modality mismatch, however, participants relied only on the conceptual fluency associated with typical features. Under time pressure to respond, however, this retrieval response strategy based on indiscriminate fluency was compromised and participants would no longer be able to use perceptual and conceptual fluency differentially and thus participants responded to typical and atypical features identically.

One surprising result of this experiment was the perceptual fluency influence on the both types of features in the non-speeded response condition. Consistent with our main hypothesis, a meta-analysis of the effect of perceptual fluency on the non-typical versus typical features across the three experiments showed a much stronger effect for first

[3] http://www.singlepapermetaanalysis.com.

than the latter. However, it may be possible that repetition priming effects for typical features can emerge but have a smaller magnitude than the same effects for non-typical effects.

6.1. Implications for theorizing on category representation

In our view, the present results have important implications for current research and theorizing on category representation. First of all, our results extend previous work by Oppenheimer and Frank (2008) in that they also show that perceptual fluency can cause category malleability, however its impact is much stronger on non-typical items. As addressed before, Oppenheimer and Frank (2008) mentioned in their paper that they used features with different degrees of typicality, however, they did not report any analysis separating typical from less typical features.

Second, our results resonate with studies showing malleability within category representations is moderated by the degree of typicality or centrality of category items. For example, work by Santos and colleagues showed that only non-typical information (i.e., personality traits with certain social categories) was influenced by a subtle priming manipulation implemented prior to a property verification task (Katz & Braly, 1933) in which participants had to select the traits that best described the individuals of one specific social category (childcare professionals or computer programmers). Further research suggested that this effect is caused by both the subtle activation of the traits and people's poor self-monitoring capabilities such that, if participants are asked to think about the target social category before the priming task, the likelihood that unrelated traits will be picked as good descriptors of a category drops substantially (see Santos et al., 2017).

In our research, why did perceptual fluency have larger effects on non-typical category items? In the case of non-typical items, it is not surprising that people used perceptual fluency as a signal to typicality judgments because these features lacked clear conceptual fluency cue values. Thus, our results suggest that perceptual fluency has a higher probability of influencing typicality judgments when people are somewhat uncertain about the degree of typicality of the target item. That is, they lack other (more diagnostic) informational cues on which to base their judgment. For example, when asked whether the feature *swims* belong to a *bat*, one could think that a bat is a flying animal, however, there are some birds, insects and even fishes that can fly and swim. In situations like this, perceptual fluency has a higher probability of being used to inform one's judgments.

Regarding typical category items, why were they less affected by perceptual fluency? Two related possibilities are, on the one hand, that the degree of judgment accuracy for these items is very high (sometimes near the ceiling) and, on the other hand, that they are already accompanied by a feeling of fluency resulting from their high accessibility level in memory (e.g., Fujihara et al., 1998; Hampton, 1979; Mervis et al., 1976; Reber, Schwarz, & Winkielman, 2004; Uyeda & Mandler, 1980). The combination of these two factors leaves little room for the impact of fluency coming from repetition priming (see Miles & Minda, 2012). This is not to say that perceptual fluency will never influence typical items. As seen in Experiment 3, when for some reason the performance accuracy on typical items decreases, perceptual fluency might play a role in typicality judgments.

In fact, this effect of perceptual fluency on highly accurate judgments is not restricted to typicality judgments. For example, it has been argued that the effect of processing fluency on truth judgments is moderated by stored knowledge. Namely, fluent statements are judged "true" more often than disfluent ones, but only when participants lack knowledge about the statements' veracity, like when sentences are unknown to participants (Unkelbach, 2007). Recent research by Fazio, Brashier, Payne, and Marsh (2015) has, however, challenged this conclusion. In their study, in their study, they found that repetition increased statements' perceived truth, regardless of stored knowledge. They propose that although people might possess the relevant stored knowledge, sometimes they fail to bring their knowledge to bear and/or to appropriately apply it. In these cases, people rely on indiscriminate fluency.

6.2. Implications for theorizing on perceptual fluency

The present findings may also have some implications for the perceptual fluency literature. Namely, in Experiment 2, we showed that visual perceptual fluency ceased to be used as a cue to typicality when the previous learning experience was auditory. These results are thus in line with previous work exploring the role of fluency on recognition judgments. As it has been demonstrated, people possess metacognitive knowledge about the types of learning episodes that give rise to later fluency, and they are capable of using that knowledge to regulate the influence of fluency upon their judgments (e.g., Lloyd et al., 2003; Miller et al., 2008; Westerman et al., 2002, 2003). For example, work by Westerman et al. (2002, 2003) showed that changing the sensory modalities (auditory-visual) and forms (pictures-words) between study and test reduce the validity of fluency as a sign of prior experience with a stimulus. Therefore, our findings also have implications for the literature on fluency effects as they reinforce the idea that the interpretation and translation of fluency into judgments is a very sophisticated process.

Although we had no particular expectations regarding the role of time pressure on typical features, the effect obtained for typical features in this condition is not entirely surprising. Namely, previous research has shown that when judgments are made under sub-optimal conditions, such as when participants are given a response deadline, they have a higher likelihood of relying on fluency cues even when other more diagnostic cues are also available (e.g., Westerman, 2001).

We interpret the absence of a fluency effect in the auditory modality condition as evidence that people are reluctant to attribute the visual fluency caused by repetition priming to a study list that was presented auditorily. In other words, the modality mismatch between study and test compromises the relevance of the perpetual fluency. Thus, despite the fact that the content of the study list (typical features) was highly relevant to the test, participants did not attribute such fluency to the auditory study list when confronted with visually fluent non-typical features. As a result, they were less influenced by fluency than participants who were presented with a visual study list.

Although our hypothesis and results are consistent with the attributional account of fluency (Jacoby et al., 1989; Westerman et al., 2002), there might be some open questions that need further research. For example, repetition priming effects are usually taken to reflect perceptual fluency, but, to differentiate between typical and non-typical features, conceptual fluency is probably involved. Importantly, it has been shown that conceptual fluency (produced by predictive sentence stems) is resistant to modality manipulations like the one employed in the current experiments (see Miller et al., 2008). Thus, if our repetition priming manipulation also induces conceptual fluency, as it has been suggested by some research (Silva, T. Garcia-Marques, & Mello, 2016), why would the responses seem to ignore fluency cues?

Our answer to this question is that under modality match, participants do not need to discriminate between perceptual and conceptual fluency. And, our results suggest they don't. That is, we obtained both typicality (an indication of the use of conceptual fluency) and repetition priming (an indication of the use of perceptual fluency). However, under modality mismatch, participants cease to rely on, in this case, irrelevant perceptual fluency and accept as typical the features with strong conceptual fluency (i.e., typical features). As this selection and discrimination between types of fluency supposedly require time and resources, indiscriminate fluency again affects typicality judgments under time pressure, both under modality match and mismatch, and thus typicality and repetition priming effects are found under these conditions.

6.3. Conclusion

The present work shows that perceptual fluency induced by repetition priming can induce category malleability, but that such malleability is not unlimited. Namely, it is more pronounced for non-typical category information, in situations in which perceptual fluency is perceived as a relevant cue to the judgment task and only when participants possess the resources to discriminate between perceptual and conceptual fluency. These results are thus consistent with a dynamic view of category representations in which these are capable of displaying malleability, but also stability, depending on the context.

Acknowledgments

This research was partially supported by a grant (SFRH/BPD/92271/2013) from the Portuguese Foundation for Science and Technology (FCT), awarded to the first author. We are grateful to Blakeley McShane for the clarifications on the single-paper meta-analysis methodology. We also thank Marianne Lloyd and another (anonymous) reviewer for their constructive comments on the manuscript.

Appendix A

Typical and non-typical features used in both experiments.

Animal names	Typical features	Non-typical features
Bat	caves, blind	swims, crest
Butterfly	cocoon, delicate	scales, horn
Camel	humps, spits	ruminant, hunch
Cat	purrs, independent	grunts, photopic
Crab	claws, sand	hemocyanin, gills
Chicken	cluck, henhouse	bladder, keel
Dog	barks, cats	boletus, appendix
Dolphin	trainable, communicates	wrinkles, electroreceptors
Goat	beard, cheese	metacarpal, tears
Octopus	tentacles, suction cups	autotomy, beak
Pig	bacon, squeals	cartilage, *invenáveis*
Rat	sewers, diseases	vomits, clay
Spider	webs, poison	binocular, tarsus
Tortoise	longevity, slow	teeth, scales
Whale	spouts, breathe	whistles, swallows

Notes. The features were freely translated from the Portuguese. The feature "suction cups", in Portuguese, is only one word ("ventosas"). We could not find a good English translation for the feature *invenáveis*, which basically means that pigs can hardly be killed by poisonous animals or foods.

Appendix B

The single-paper meta-analyses (SPM) is a statistical tool developed by McShane and Böckenholt (2017) specially tailored to meta-analyze a set of experiments. The SPM tool is user-friendly because it requires only basic summary information (e.g., means, standard deviations, and sample sizes).

The SPM tool returns individual estimates for each one of experiments analyzed and one SPM estimate. Each estimate is accompanied by 50% and 95% intervals. The tool also yields an SPM heterogeneity estimate (I^2) that gives the percentage of the variation in the observations (beyond that attributable to the experimental manipulations). Pigott (2012) provides guidance on the typical size of I^2 in behavioral research. This author defines low heterogeneity as I^2 = 25%; medium heterogeneity as I^2 = 50%, and high heterogeneity as I^2 = 75% (see also Higgins & Thompson, 2002).

The authors provide an easy-to-use website that implements the SPM methodology (http://www.singlepa permetaanalysis.com/). This website includes a detailed tutorial that shows how use this tool.

References

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13*, 219–235. http://dx.doi.org/10.1177/1088868309341564.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*, 569–576. http://dx.doi.org/10.1037/0096-3445.136.4.569.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition, 13*(3), 263–308. http://dx.doi.org/10.1016/0010-0277(83)90012-4.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition, 10*, 82–93. http://dx.doi.org/10.3758/BF03197629.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11*, 211–227. http://dx.doi.org/10.3758/BF03196968.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 629–654. http://dx.doi.org/10.1037/0278-7393.11.1-4.629.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.). *Concepts reconsidered: The ecological and intellectual bases of categories* (pp. 101–140). New York: Cambridge University Press.

Barsalou, L. W. (1989). Intra-concept similarity and its implications for inter-concept similarity. In S. Vosniadou, & A. Ortony (Eds.). *Similarity and analogical reasoning.* New York: Cambridge University Press.

Barsalou, L. W. (1993). Structure, flexibility, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. C. Collins, S. E. Gathercole, & M. A. Conway (Eds.). *Theories of memory* (pp. 29–101). London: Lawrence Erlbaum Associates, Inc.

Barsalou, L. W. (2008). Cognitive and neural contributions to understanding the conceptual system. *Current Directions in Psychological Science, 17*, 91–95. http://dx.doi.org/10.1111/j.1467-8721.2008.00555.x.

Barsalou, L. W. (2017). Cognitively plausible theories of concept composition. In Y. Winter, & J. A. Hampton (Eds.). *Compositionality and concepts in linguistics and psychology*London: Springer Publishing (in press).

Barsalou, L. W., & Medin, D. L. (1986). Concepts: Static definitions or context-dependent representations? *Cahiers de Psychologie Cognitive, 6*, 187–202.

Barsalou, L. W., & Sewell, D. R. (1984). Constructing representations of categories from different points of view. *Emory cognition project. Vol. 2* (Atlanta: GA).

Benjamin, A. S., & Craik, F. I. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition, 29*, 691–697. http://dx.doi.org/10.3758/BF03200471.

Brown, A. S., & Marsh, E. J. (2009). Creating illusions of past encounter through brief exposure. *Psychological Science, 20*, 534–538. http://dx.doi.org/10.1111/j.1467-9280.2009.02337.x.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation.

Crocker, J., Fiske, S. T., & Taylor, S. E. (1984). Schematic bases of belief change. In J. R. Eiser (Ed.). *Attitudinal judgment* (pp. 197–226). New York: Springer-Verlag.

De Houwer, J., & Smith, C. T. (2013). Go with your gut! Effects in the affect misattribution procedure become stronger when participants are encouraged to rely on their gut feelings. *Social Psychology, 44*, 299–302. http://dx.doi.org/10.1027/1864-9335/a000115.

Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General, 144*, 993–1002. http://dx.doi.org/10.1037/xge0000098.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition.* Reading, MA: Addison-Wesley.

Fujihara, N., Nageishi, Y., Koyama, S., & Nakajima, Y. (1998). Electrophysiological evidence for the typicality effect of human cognitive categorization. *International Journal of Psychophysiology, 29*, 65–75. http://dx.doi.org/10.1016/S0167-8760(97)00099-8.

Gallo, D. A., Perlmutter, D. H., Moore, C. D., & Schacter, D. L. (2008). Distinctive encoding reduces the Jacoby-Whitehouse illusion. *Memory & Cognition, 36*, 461–466. http://dx.doi.org/10.3758/MC.36.2.461.

Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2004). Reducing false recognition with criterial recollection tests: Distinctiveness heuristic versus criterion shifts. *Journal of Memory and Language, 51*, 473–493. http://dx.doi.org/10.1016/j.jml.2004.06.002.

Garcia-Marques, L., Santos, A. S., & Mackie, D. M. (2006). Stereotypes: Static definitions or dynamic knowledge structures? *Journal of Personality and Social Psychology, 91*, 814–831. http://dx.doi.org/10.1037/0022-3514.91.5.814.

Garcia-Marques, T., Mackie, D. M., Claypool, H., & Garcia-Marques, L. (2010). Is it familiar or positive? Mutual facilitation of response latencies. *Social Cognition, 28*(2), 205–218. http://dx.doi.org/10.1521/soco.2010.28.2.205.

Garcia-Marques, T., Mackie, D. M., Claypool, H. M., & Garcia-Marques, L. (2004). Positivity can cue familiarity. *Personality and Social Psychology Bulletin, 30*, 585–593. http://dx.doi.org/10.1177/0146167203262856.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 441–461. http://dx.doi.org/10.1016/S0022-5371(79)90246-9.

Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*, 1539–1558. http://dx.doi.org/10.1002/sim.1186.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: Br. Med. J. 327*, 557–560. http://dx.doi.org/10.1136/bmj.327.7414.557.

Huber, D. E., Clark, T. F., Curran, T., & Winkielman, P. (2008). Effects of repetition priming on recognition memory: Testing a perceptual fluency-disfluency model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1305–1324. http://dx.doi.org/10.1037/a0013370.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541. http://dx.doi.org/10.1016/0749-596X(91)90025-F.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*, 306–340. http://dx.doi.org/10.1037/0096-3445.110.3.306.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger, & F. I. M. Craik (Eds.). *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). NJ: Erlbaum.: Hill- sdale.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*, 326–338. http://dx.doi.org/10.1037/0022-3514.56.3.326.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118*, 126–135. http://dx.doi.org/10.1037/0096-3445.118.2.126.

Jacoby, L. L., Woloshyn, V., & Kelley, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118*, 115–125. http://dx.doi.org/10.1037/0096-3445.118.2.115.

Katz, D., & Braly, K. W. (1933). Racial stereotypes of 100 college students. *Journal of Abnormal Social Psychology, 28*, 280–290.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32*, 1–24. http://dx.doi.org/10.1006/jmla.1993.1001.

Kurilla, B. P., & Gonsalves, B. D. (2012). An ERP investigation into the strategic regulation of the fluency heuristic during recognition memory. *Brain Research, 1442*, 36–46. http://dx.doi.org/10.1016/j.brainres.2011.12.060.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Larochelle, S., & Pineu, H. (1994). Determinants of response time in the semantic verification task. *Journal of Memory and Language, 33*, 796–823. http://dx.doi.org/10.1006/jmla.1994.1038.

Lloyd, M. E., Hartman, A., Ngo, C. T., Ruser, N., Westerman, D. L., & Miller, J. K. (2015). Not enough familiarity for fluency: Definitional encoding increases familiarity but does not lead to fluency attribution in associative recognition. *Memory & Cognition, 43*, 39–48. http://dx.doi.org/10.3758/s13421-014-0449-3.

Lloyd, M. E., Westerman, D. L., & Miller, J. K. (2003). The fluency heuristic in recognition memory: The effect of repetition. *Journal of Memory and Language, 48*, 603–614. http://dx.doi.org/10.1016/S0749-596X(02)00535-1.

Luka, B. J., & Barsalou, L. W. (2005). Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *Journal of Memory and Language. 52*, 436–459. http://dx.doi.org/10.1016/j.jml.2005.01.013.

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory and Cognition, 6*, 462–472. http://dx.doi.org/10.3758/BF03197480.

McCloskey, M. E., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology, 11*, 1–37. http://dx.doi.org/10.1016/0010-0285(79)90002-1.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37*, 547–559. http://dx.doi.org/10.3758/BF03192726.

McShane, B. B., & Böckenholt, U. (2017). Single-paper meta-analysis: Benefits for study summary, theory testing, and replicability. *Journal of Consumer Research, 43*, 1048–1063. http://dx.doi.org/10.1093/jcr/ucw085.

Medin, D. L., & Smith, E. E. (1984). Concept and concept formation. *Annual Review of Psychology*. http://dx.doi.org/10.1146/ annurev.psych.35.1.113.

Mervis, C. B., Catlin, J., & Rosch, E. (1976). Relationships among goodness of example, category norms, and word frequency. *Bulletin of the Psychonomic Society, 7*, 283–284. http://dx.doi.org/10.3758/BF03337190.

Miles, S. J., & Minda, J. P. (2012). Perceptual fluency can be used as a cue for categorization decisions. *Psychonomic Bulletin & Review, 19*, 737–742. http://dx.doi.org/10.3758/s13423-012-0259-1.

Miller, J. K., Lloyd, M. E., & Westerman, D. L. (2008). When does modality matter? Perceptual versus conceptual fluency-based illusions in recognition memory. *Journal of Memory and Language, 58*, 1080–1094. http://dx.doi.org/10.1016/j.jml.2007.12.006.

Oppenheimer, D. M., & Frank, M. C. (2008). A rose in any other font wouldn't smell as sweet: Fluency effects in categorization. *Cognition, 106*, 1178–1194. http://dx.doi.org/10.1016/j.cognition.2007.05.010.

Palma, T. A., Santos, A. S., & Garcia-Marques, L. (2017). The future is now: The impact of present fluency in judgments about the future. *Memory*. http://dx.doi.org/10.1080/09658211.2017.1335328.

Pigott, T. D. (2012). *Advances in meta-analysis.* New York: Springer.

Ramos, T., Garcia-Marques, L., Santos, A. S., & Carneiro, P. (2015). Stimulus-response bindings in the Go/NoGo task. *Experimental Psychology, 62*, 264–275. http://dx.doi.org/10.1027/1618-3169/a000297.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition, 8*, 338–342. http://dx.doi.org/10.1006/ccog.1999.0386.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Perceptual fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*, 364–382. http://dx.doi.org/10.1207/s15327957pspr0804_3.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9*, 45–48. http://dx.doi.org/10.1111/1467-9280.00008.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12*, 1–20. http://dx.doi.org/10.1016/S0022-5371(73)80056-8.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.). *Cognitive development and the acquisition of language.* New York: Academic Press.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192–233. http://dx.doi.org/10.1037/0096-3445.104.3.192.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573–605. http://dx.doi.org/10.1016/0010-0285(75)90024-9.

Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15*, 346–378. http://dx.doi.org/10.1016/0010-0285(83)90012-9.

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*, 81–104. http://dx.doi.org/10.1111/j.1540-4560.1985.tb01130.x.

Santos, A. S., Garcia-Marques, L., Mackie, D. M., Ferreira, M. B., Payne, B. K., & Moreira, S. (2012). Implicit open-mindedness: Evidence for and limits on stereotype malleability. *Journal of Experimental Social Psychology, 48*, 1257–1266. http://dx.doi.org/10.1016/j.jesp.2012.05.013.

Santos, A. S., Garcia-Marques, L., Mackie, D. M., Palma, T. A., Costa, R. S., & de Almeida, F. (2017). Something in the way you primed me: Belief monitoring when source identification is not possible. *Social Cognition, 35*, 273–298. http://dx.doi.org/10.1521/soco.2017.35.3.273.

Silva, R. R., Garcia-Marques, T., & Mello, J. (2016). The differential effects of fluency due to repetition and fluency due to color contrast on judgments of truth. *Psychological Research, 80*, 821–837. http://dx.doi.org/10.1007/s00426-015-0692-7.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81*, 214–241. http://dx.doi.org/10.1037/h0036351.

Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.). *Cognitive processes in stereotyping and intergroup behavior* (pp. 83–114). Hillsdale, NJ: Erlbaum.

Thapar, A., & Westerman, D. L. (2009). Aging and fluency-based illusions in recognition memory. *Psychology and Aging, 24*, 595–603. http://dx.doi.org/10.1037/a0016575.

Topolinski, S., & Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General, 138*, 39–63. http://dx.doi.org/10.1037/a0014678.

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of perceptual fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 219–230. http://dx.doi.org/10.1037/0278-7393.33.1.219.

Unkelbach, C., & Greifeneder, R. (2013). A general model of fluency effects in judgment and decision making. In C. Unkelbach, & R. Greifeneder (Eds.). *The experience of thinking: How the fluency of mental processes influences cognition and behaviour* (pp. 11–32). New York, NY: Psychology Press.

Uyeda, K. M., & Mandler, F. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods and Instrumentation, 12*, 587–595.

Westerman, D. L. (2001). The role of familiarity in item-recognition, associative-recognition, and plurality recognition on self-paced and speeded tests. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*, 723–732. http://dx.doi.org/10.1037/0278-7393.27.3.723.

Westerman, D. L., Lloyd, M. E., & Miller, J. K. (2002). The attribution of perceptual fluency in recognition memory: The role of expectation. *Journal of Memory and Language, 47*, 607–617. http://dx.doi.org/10.1016/S0749-596X(02)00022-0.

Westerman, D. L., Miller, J. K., & Lloyd, M. E. (2003). Change in perceptual form attenuates the use of the fluency heuristic in recognition. *Memory & Cognition, 31*, 619–629. http://dx.doi.org/10.3758/BF03196102.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning Memory & Cognition, 19*, 1235–1253. http://dx.doi.org/10.1037/0278-7393.19.6.1235.

Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? The unexpected basis of feelings of familiarity. *Acta Psychologica, 98*, 141–166. http://dx.doi.org/10.1016/S0001-6918(97)00040-1.

Whittlesea, B. W. A., & Williams, L. D. (2001a). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1–32. http://dx.doi.org/10.1037/0278-7393.27.1.3.

Whittlesea, B. W. A., & Williams, L. D. (2001b). The discrepancy-attribution hypothesis: II. Expectation, uncertainty, surprise and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 14–33. http://dx.doi.org/10.1037/0278-7393.27.1.14.

Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003). The hedonic marking of perceptual fluency: Implications for evaluative judgment. In J. Musch, & K. C. Klauer (Eds.). *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Mahwah, NJ: Lawrence Erlbaum.

Wolk, D. A., Gold, C. A., Signoff, E. D., & Budson, A. E. (2009). Discrimination and reliance on conceptual fluency cues are inversely related in patients with mild Alzheimer's disease. *Neuropsychologia, 47*, 1865–1872. http://dx.doi.org/10.1016/j.neuropsychologia.2009.02.029.