Dissociation between Categorization and Similarity Judgments

Jean-Pierre Thibaut and **Myriam Dupont**
University of Liège.
Department of Psychology. B 32.
4000 Liège,
Belgium
(jpthibaut, mdupont)@ulg.ac.be

Abstract

A dissociation between categorization and similarity was found by Rips (1989) In one experiment, Rips found that a stimulus half-way between a pizza and a quarter was categorized as a pizza but was rated as more similar to a quarter. Smith & Sloman (1994) discuss these results in terms of the role of necessary and characteristic features. In one experiment, participants had to learn to categorize new stimuli (unknown shapes) built with necessary and characteristic features. We compared two experimental conditions in which we manipulated the association between the characteristic features and the two categories. Contrary to the suggestion made by Smith and Sloman, subjects categorized the stimuli on the basis of a necessary feature. However, their similarity judgments relied on the characteristic features. This resulted, for one of the two experimental conditions, in a perfect dissociation between similarity and categorization. According to Rips, the dissociation indicates that categorization and similarity rating are different processes. On the contrary, we suggest that categorization and similarity are the same processes, but that they sometimes operate on different subsets of features.

Introduction

According to many authors, similarity is a central concept for models of categorization in the sense that categorization is grounded on similarity. Technically, an object X is categorized in a category A instead of a category B if its representation is more similar to the representation of category A than to the representation of category B (see Komatsu, 1992; Thibaut, in press, for reviews).

Contrary to this theoretical proposal, Rips (1989) (see also Rips & Collins, 1993) provided empirical evidence for a dissociation between categorization and similarity. The experimental setup consisted of a comparison between categorization and similarity judgments of a target stimulus X with respect to two categories of stimuli A and B. The rationale was that if the target was categorized in A more often than in B but judged more similar to B than to A, this result would demonstrate the dissociation. In a first experiment, Rips (1989) read his subjects a description of a target object described in terms of a value on a single dimension (e.g., the diameter). This value was chosen halfway between the largest dimensional value of a small category and the smallest value of a large category. To illustrate, a target 3 in. object was chosen to be halfway between subjects' estimate of a US quarter size (1 in.) and their size estimate of the smallest pizza (5 in.). It is important to emphasize that the variance along the critical dimension was different in the two categories. For the category of quarters, the size of the diameter is fixed, whereas the category of pizzas is allowed more variation on the "size" dimension. In the categorization condition, subjects were required to categorize the 3 in. target object in one of the two categories. In the similarity condition, subjects were asked to rate the similarity of the target with respect to the two categories. It was shown that while most (63%) categorization subjects categorized the target in the variable category (e.g., pizza), most similarity subjects (69%) found the target to be more similar to fixed category (e.g., US quarters).

This important result was taken as an evidence that categorization is not based on similarity. In the case of categorization, most subjects seemed to follow a rule. In the preceding example, the quarter diameter size is fixed by the law and cannot be 3 in. Similarity was also estimated by reference to diameter. The diameter of all instances in the fixed category (1 inch) is two inches smaller than the target stimulus. On the other hand, only the smallest pizzas (5-inch diameter) are two inches larger than the target stimulus (for all the other pizzas, the difference between them and the target stimulus, in terms ot the diameter size, is larger than two inches). Consequently, most subjects estimated that the target stimulus was more similar to the fixed category than to the variable category.

Smith and Sloman (1994) tried to replicate Rips' results in 2 experiments. The instructions encouraged subjects to use rule-based categorizations by pointing to the existence of a feature sufficient for categorization in one of the 2 categories. In their second experiment, in which the procedure matched Rips' one very closely, choices in the similarity task (50% of choices in favor of the variable category) differed significantly from the results in the categorization task (67%). However, their results did not replicate one of the results obtained by Rips for similarity. In Rips, results in the similarity judgment task were clearly in favor of the fixed category (69% of choices in favor of this category) but not in Smith and Sloman (50% for the fixed category).

In a second experimental condition (the *rich* condition), Smith and Sloman added a characteristic feature of the fixed category to the original description of the target item. To illustrate, the characteristic feature "that is silver colored" was added to the original description "a circular object with an X-inch diameter". In this *rich* condition, contrary to the preceding condition (called *sparse* condition) there was no difference between categorization and similarity. Participants categorized the target items in the fixed category (77%) and estimated them more similar to this category (74%).

According to Smith and Sloman, those results are incompatible with what they called the *necessary feature* hypothesis according to which subjects categorize objects on the sole basis of a necessary feature whenever there is one available. In contrast, these results seem to corroborate the *characteristic feature* hypothesis which holds that people's categorizations are based on characteristic and necessary features when these two kinds of features are available. Given that, in the categorization task, subjects were strongly encouraged to use a rule to categorize the stimuli, the results also show that subjects do not rely on necessary features for categorization and on characteristic features for similarity judgments

Comparing Rips and Smith and Sloman's results, one is confronted with a discrepancy. As acknowledged by Smith and Sloman, this discrepancy could result from their inclusion of rich descriptions in the list of stimuli. Our experiment explores this discrepancy and manipulates the characteristics of the stimuli used in the preceding experiments. In Rips experiment and in Smith and Sloman's sparse description there is only one dimension used to describe the stimuli. In the pizza-quarter example, it is the size. The size of quarters is fixed and plays the role of a necessary feature for this category. For pizzas, the size is variable and there is a set of values that are characteristic of the category. The two categories are not defined systematically in terms of necessary and characteristic features. As a result, these experiments are not optimal for contrasting the role of characteristic features and necessary features in similarity rating and categorization.

In the following experiment, we manipulated the features constituting the stimuli. We designed stimuli with new features (in the experiment, a set of four appendages called "legs" and connected to another part). There are three types of features : first, a necessary feature, i.e. a feature that can be used as a rule for categorization since it is present in each stimulus of one category and absent in all the stimuli belonging to the second category (in the experiment, the spatial grouping of the four legs : 1-leg-plus-a-group-of-three-legs for category A and 2-legs-plus-2-legs for category B, see Figure 1). Second, a characteristic feature is present in a subset of the stimuli of category A (a mushroom shape) whereas another feature is present in a subset of the stimuli of category B (an angular shape). This characteristic feature was explicitly designed to be very salient. Third, there are features common to both categories that cannot be used to distinguish the two categories (these are, in some sense, distractor features).

We also compared two experimental conditions that differed from one another in terms of the association between the characteristic feature and the two categories. In the first condition (the *restricted* condition in what follows), the characteristic feature of category A (called F2a) was absent in category B and the characteristic feature of category B (F2b) was absent in category A. In the second condition (the *cross-category* condition), F2a was present in 5 out of 8 category A stimuli and in 1 stimulus of category B, F2b was present in 5 stimuli of category B and in one stimulus of category A.

The purpose of the experiment was to compare the similarity judgments and categorizations for "contradictory" stimuli, i.e. stimuli built with the characteristic feature of one category and the necessary feature of the other. Given Smith and Sloman's (1994) results, we would predict that no subject will categorize on the basis of the necessary feature only. Moreover, if categorizations are influenced by the (salient) characteristic feature, some subjects will categorize the "contradictory" stimuli according to the characteristic feature, especially in the first condition (the *restricted* condition). For the similarity judgments on the "contradictory" stimuli, depending on the kind of feature (characteristic or necessary) used by the subjects, the judgment will favor one category or the other. For example, consider a "contradictory" stimulus composed of the characteristic feature of category A and the necessary feature of category B. Subjects will rate it as more similar to category A (category B) if they rely on the characteristic feature (necessary feature) and they will rate it as more similar to category B if they rely on the necessary feature. Dissociations will occur when subjects use the necessary feature for categorization and the characteristic salient feature for similarity (or the reverse, although this is unlikely). If such dissociations should occur, we expect them to be more frequent in the *restricted* condition since, in this condition, the characteristic feature of one category is never displayed in the other category. Following Smith and Sloman's hypothesis about the role of necessary and characteristic features in categorization, we should not expect strong dissociations to occur.

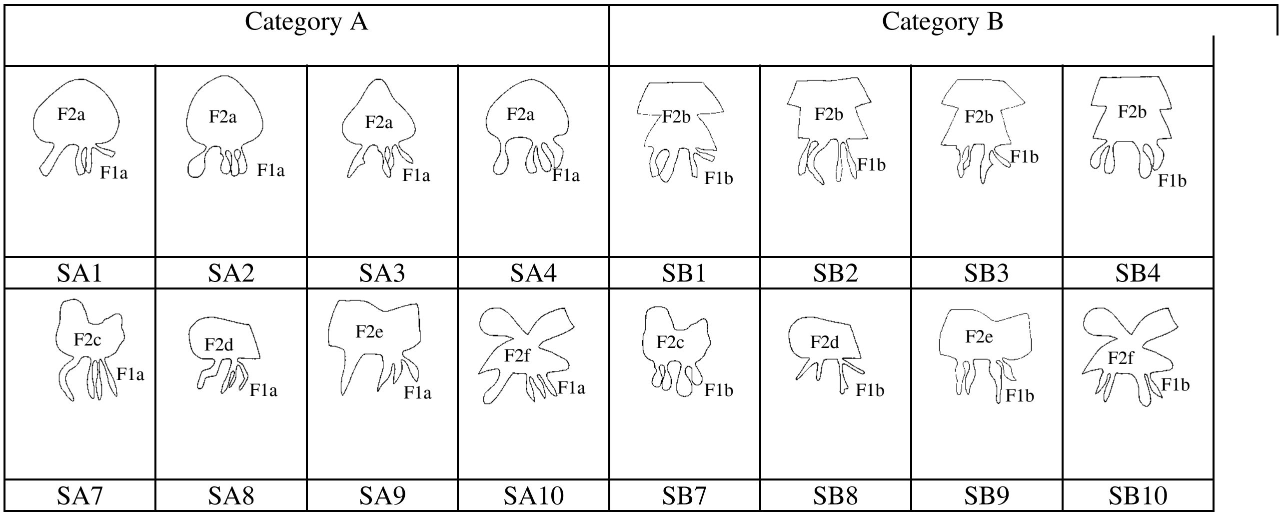

Figure 1a : The 16 stimuli from the *restricted* condition : 8 stimuli from categories A and B used in the learning phase. SA1 to SB4 are congruent stimuli and SA7 to SB10 are neutral stimuli.

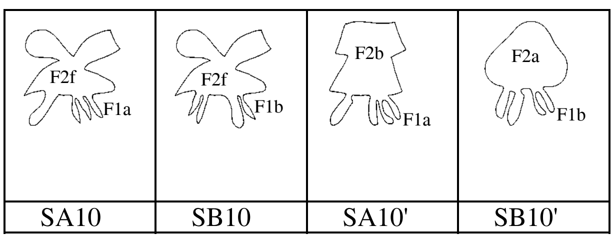

Figure 1b : In the *cross-category* condition, SA10 and SB10 were replaced by the "contradictory" stimuli SA10' and SB10'.

Experiment

Method

Subjects. Twenty-two undergraduate tudents from the University of Liège volunteered for the experiment.

Material. *Learning phase*. Two categories of 10 stimuli that are outlines of unknown shapes were constructed. The stimuli were composed of two parts, the upper part (the different F2 parts in Figure1a) and the lower part (the F1 parts in figure1a). In the *restricted* condition, for six out of the ten stimuli, the upper part, for category A, has a "mushroom" shape that is slightly distorted over the six stimuli (F2a), and an angular shape for stimuli in category B (F2b). These six stimuli were called "congruent" (the first row in Figure 1a displays 4 stimuli of each category). The four remaining stimuli of the two categories, called the "neutral" stimuli, were constructed with four different upper parts (F2c,d, e,f). Since F2c,d,e,f are present in both categories they cannot be used as cues for categorization (the second row of Figure 1a displays the 4 neutral stimuli from the 2 categories). Each lower part is composed of four legs that are spatially grouped as one leg on the left and three legs on the right for Category A (1/3, F1a in Figure 1a), and two sets of two legs in category B (2/2, F1b in Figure 1a). The *cross-category* condition was constructed in the same way except that one of the four "neutral" stimuli from category A and one from category B (SA10 and SB10) were replaced by two new stimuli. One made with the necessary (F1a) feature of category A (i.e., 1-3 legs) and the characteristic feature of category B (F2b), stimulus called SA10', and a stimulus made with one necessary feature from category B (F1b) (2-2 legs) and one characteristic feature of category A (F1a), a stimulus called SB10' (see Figure 1b). These two stimuli were called "contradictory".

Test phase stimuli. Twenty-two new stimuli, eleven per category, were constructed according to the same principles. For each category, there were 2 congruent, 4 neutral and 5 contradictory stimuli.

Procedure.

Each subject was randomly assigned to one of the 2 conditions, i.e. eleven subjects per condition. Subjects were tested individually. They took 20 to 45 minutes to complete the task. The experiment is composed of two phases : a learning phase and of a test phase.

Learning phase. Subjects are told that they would have to learn to sort a set of stimuli into two categories. A first stimulus is presented to the subject who has to guess its category name. Feedback was provided about the accuracy of the answers. The second stimulus, and so forth for the other stimuli. Feedback is provided for each stimulus and the order of presentation of the stimuli is random. Once the entire set has been presented to the subject, it is presented a second time, a third time, etc. The learning phase ends when the subject makes no mistakes during two successive presentations of the set of the stimuli.

Test phase. Subjects were presented with the test stimuli. For each stimulus, each subject was asked to decide which of the two categories the stimulus belonged to and to choose the category the object was more similar to. Half of the subjects did the similarity task first whereas the other half did the categorization task first. A second part of the test phase was a rating task. Subjects had to rate on a scale from 1 to 7 whether the test stimuli were likely to belong to category A or category B. They also had to rate whether the test stimuli were more similar to category A or to category B. The end of the scale corresponding to "1" referred to category A and the end corresponding to "7" referred to category B.

Results

First, we searched for dissociations between categorization and similarity judgments. As predicted, there was no dissociation for congruent stimuli. We analyzed the results obtained for the 10 "contradictory" test-stimuli (i.e. five F1a + F2b stimuli and five F1b + F2a stimuli). We considered that a subject dissociates categorization and similarity when he/she categorized 9 or 10 test-stimuli in one of the 2 categories (A or B) while estimating them more similar to the other category (B or A). Twelve subjects produced such a dissociation. However, a comparison between the *restricted* and the *cross-category* conditions reveals that 10 subjects (out of 11) dissociated in the *restricted* condition and 2 in the *cross-category* condition. A fisher-Exact test reveals that the two conditions differ significantly in the proportion of dissociations obtained ($p < .01$). To summarize, this analysis reveals that twelve subjects dissociated categorization and similarity judgments. However, 10 subjects dissociated in the *restricted* condition while, in the other condition, only 2 subjects made this dissociation.

In order to confirm these analyses, we perform a two-way ANOVA (2 x 2) with Condition (*cross-category* and *restricted*) as a class variable and Task (*categorization* and *similarity*) as a between variable on the ratings for similarity and categorization. Dissociations are obtained for stimuli that get a small score for categorization similarity and a high score for similarity or a small score for similarity and a high score for categorization. In order to perform a single analysis on the 1-3 and 2-2 test-stimuli scores, we recoded the categorization and similarity scores for the 2-2 contradictory test-stimuli (i.e. stimuli that had to be categorized in B when the subject followed the rule 2-2). High scores in categorization, indicating that subjects categorized the 2-2 stimuli in category B, were transformed into small scores (i.e. 7, 6, 5, 4, 3, 2, 1 are transformed into 1, 2, 3, 4, 5, 6, 7 respectively). Small scores for similarity, indicating that subjects judged the 2-2 test-stimuli as more similar to category A, were transformed into high scores (i.e. 1, 2, 3, 4, 5, 6, 7 are transformed into 7, 6, 5, 4, 3, 2, 1 respectively). There was a significant effect of Condition : $F(1,20) = 33.96$, $p < .0001$, of Task : $F(1,20) = 63.51$, $p < .0001$, and a significant interaction Condition x Task : $F(1,20) = 25.77$, $p < .0001$. A posteriori test (Student-Newman-Keuls) revealed a significant difference between *restricted* and *cross-category* conditions for similarity judgments ($p < .05$). There was no difference between these two conditions for categorization (see Figure2). We computed a confidence interval on the similarity and the categorization scores in the two conditions at the level of $p<.05$. The confidence interval are, for the categorization scores _0.81, 1.35_ in the *cross-category* condition, and _0.93, 1.47_ in the *restricted* condition. For the similarity ratings, they are _1.06, 3.10_ in the *cross-category* condition and _4.69, 6.72_ in the *restricted* condition. These results confirm the dissociation in the *restricted* condition. In the *restricted* condition, the hypothesis that the mean is beyond the value 4 (the intermediate value between 1 and 7 on the scale) is rejected for the categorization scores whereas it is accepted for the similarity scores. In other words, the categorization scores for the contradictory stimuli are close to the extremity corresponding to one category whereas the similarity scores are closer to the extremity corresponding to the other category.

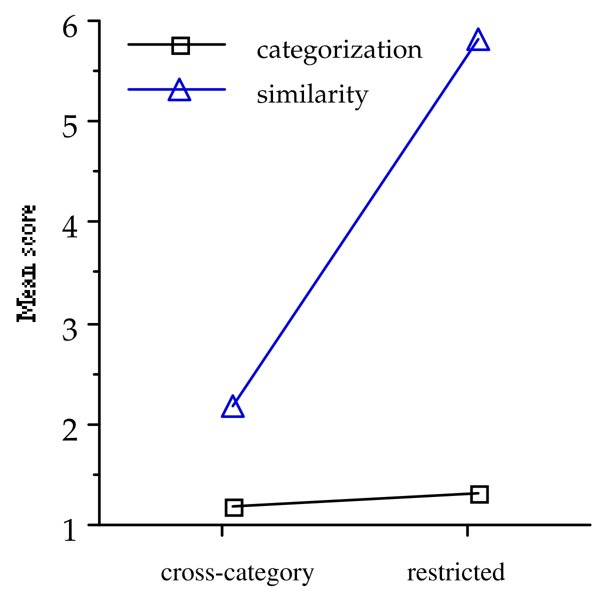

Figure 2. Interaction Task x Condition. Note the dissociation between categorization and similarity in the *restricted* condition.

The second analysis confirms the results obtained in the first one. The dissociation between similarity and categorization appears in the *restricted* condition. A comparison between the *restricted* and the *cross-category* condition indicates that they differ in terms of the similarity scores and not in terms of the categorization scores. Note that we used a within design. This is more persuasive than Rips or Smith and Sloman who used a between design since, in our experiment, the dissociation is obtained "within" a subject.

Discussion

Our first objective was to replicate Rips' dissociation between categorization and similarity judgment. We created two categories of new stimuli and controlled the frequency of association between each feature and each category: there were characteristic, necessary and non distinctive features. We hypothesized that the dissociation would occur for stimuli which are composed of a salient part that is characteristic of one category (neither necessary nor sufficient) and a non salient necessary feature (i.e., the *contradictory* stimuli). The results confirmed this prediction. Twelve subjects produced a dissociation and the majority of them belonged to the *restricted* conditon. The analysis of the ratings confirmed these results.

Contrary to Smith and Sloman's (1994) results (see introduction) who failed to replicate Rips' dissociation completely, our results confirm and expand Rips' results since most of the subjects from the *restricted* condition made a perfect dissociation between similarity and categorization.

Remember that according to Smith and Sloman (1994), categorization is based on characteristic <u>and</u> necessary features whenever these two kinds features are available (the *characteristic* feature hypothesis). They took their results as evidence for the *characteristic* feature hypothesis and against rule-based categorization, i.e. categorization based on a necessary feature. On the contrary, the scores obtained in our two experimental conditions suggest that subjects used the necessary feature to categorize the contradictory test-stimuli. Since no subject categorized the test-stimuli in the other category, it seems that they did not use the characteristic feature as a basis for their categorization. The dissociations obtained in the *restricted* condition indicate that *only* the similarity judgments were influenced by the characteristic features. To summarize, our subjects did categorize on the sole basis of a necessary feature even when a characteristic feature was available.

How can we explain the discrepancy between our results and Smith and Sloman's results? As it is acknowledged by Smith and Sloman, the discrepancy between their results and Rips' could result from their inclusion of rich descriptions in the list of stimuli. Considering the rich descriptions, the feature added is characteristic of the fixed category (i.e., "quarters"); however, this feature is not a plausible property of the stimuli of the variable category ("pizza"). For example, suppose you have to categorize "a circular object with a 3-inch diameter that is silver colored" either as a pizza (the variable category) or as a quarter (the fixed category). This entity is an implausible pizza (most subjects categorized it as a quarter in Smith and Sloman's second experiment). It is possible that, because of the oddness of the descriptions, subjects used the features in their ratings differently that were they would have done with descriptions more representative of the real world. Since the descriptions involve complex representations, there is no easy way to describe how those representations interacted and resulted in the ratings obtained in the experiment. The nature of our stimuli allowed us to avoid this "complexity problem".

We also have to explain the difference between the results obtained in the *restricted* and the *cross-category* conditions as well as the absence of dissociation in the *cross-category* condition. One possibility is that since subjects noticed that each characteristic feature was associated with both categories, they considered that the characteristic features were not good descriptors of the categories and decided to rate similarity (and to categorize) in terms of the necessary feature only. Another possibility -the more likely- is that subjects did not notice that each characteristic feature was closely associated with one of the two categories. During the first trials of the learning phase, they could have used the characteristic feature (i.e., the upper part of the figures) which is very salient, as a basis for their categorization. However, after the presentation of the contradictory stimuli, subjects searched for another classification criterion and may have failed to notice that each characteristic feature was closely associated with one of the two categories.

Why did subjects dissociate similarity and categorization in the *restricted* condition? To analyze this difference, one can analyze the stimuli in terms of their constitutive features. These features compose the feature space used by subjects to categorize or estimate similarity. It can be argued that subjects used a different subset of the feature space to perform the two tasks, because the two tasks have different constraints. Murphy and Medin (1985) suggested that similarity is so unconstrained that, a priori, everything can be similar or dissimilar to everything else in an infinite numbers of ways. Confirming this view, many experiments have demonstrated that similarity judgments can vary according to the context. As a result, Medin, Goldstone, and Gentner (1993) suggested that in order to understand similarity we have to explain the process of fixing the respects for similarity. Fixing the respects has two components: selection of a subset of features and/or weighting the features. For example, the comparison of a dog and a pig in the context of farm animals will not rely on the same space than the comparison of these two animals in the context of food. Categorization also relies on a feature space, which means that when one has to categorize a stimulus he does not use all features available about the stimulus. It could be argued that one difference between categorization and similarity is that categorization is more constrained than similarity. To illustrate, in a given context, one can say that a dog is similar to a cat, but nobody would categorize a dog as a cat. Dissociations between categorization and similarity will happen when subjects do not use the same respects and/or weight the features differently in the 2 tasks. In our experiment, the dissociation apparently arose from the fact that subjects weighted the characteristic and the necessary features differently in both tasks. However, contrary to what several authors have argued (Rips, 1989; Smith & Sloman, 1994), despite their diverging outcome, we think that similarity and categorization rely on the same processes. In our terms, the difference between similarity and categorization amounts to the subset of features selected or to a difference in the weighting of these features (see Thibaut & Schyns, 1995).

To summarize, the present paper provides a new case of dissociation between similarity judgments and categorization. Contrary to preceding papers, the status of the features involved was controlled in the stimuli. The results revealed that subjects categorized on the sole basis of necessary features even when characteristic features were available. The dissociation between similarity and categorization resulted from the use of characteristic features in the similarity task.

Acknowledgements

The author would like to thank Bob French for valuable comments on this manuscript, Rob Goldstone and Philippe Schyns for many stimulating discussions.

References

Komatsu, L.K. (1992). Recent views of conceptual structure. *Psychological Bulletin*, *112*, 500-526.

Medin, D.L., Goldstone, R.L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254-278.

Murphy, G.L., & Medin, D. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289-316.

Rips, L.J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogial reasoning*. Cambridge: Cambridge University Press.

Rips, L.J., & Collins, A. (1993). Categories and resemblance. *Journal of Experimental Psychology, 122*, 468-486.

Schyns, P.G., Goldstone, R. & Thibaut, J.-P. (in press). The development of conceptual features. *Behavioral and Brain Sciences*.

Smith, E.E., & Sloman, S.A. (1994). Similarity-versus rule-based categorization. *Memory & Cognition, 22*, 377-386.

Thibaut, J.P. (1997). Similarité et catégorisation. *L'Année Psychologique*, 97, 701-736.

Thibaut, J.P., & Schyns, P.G. (1995). The development of feature spaces for similarity and categorization. *Psychologica Belgica*, *35*, 167-185.