Cogn Process (2013) 14:401–410
DOI 10.1007/s10339-013-0569-x

RESEARCH REPORT

Psychological distance and reaction time in a Stroop task

Giorgio De Marchis · María del Prado Rivero Expósito · José Manuel Reales Avilés

Received: 26 October 2012 / Accepted: 9 May 2013 / Published online: 1 June 2013

Abstract Several sources of interference may simultaneously affect the onset of the well-known "Stroop effect." Among them is the semantic component, which is reflected in the gradient or *semantic effect*. This effect consists of an increase in the amount of interference as the semantic distance between the word and the color concept decreases. Shepard (Science 237:1317–1323, 1987) relates psychological space, measured through multidimensional scaling, to mean response times. The present investigation aims to study the function relating the semantic gradient with the psychological distance between the word and the color in a Stroop task. After measuring the gradient, we obtained the subjective rating of the degree of dissimilarity of the gradient words with the concept of "color." In our work, we show that the amount of interference in a Stroop task increases when the semantic distance from the word to the color concept decreases, and it does so exponentially. We replicated the study with different stimuli to test the robustness of the results.

Keywords Stroop color-word test · Stimulus generalization · Generalization (cognitive) · Word associations

Introduction

The Stroop effect is an extensively researched phenomenon in the field of cognitive psychology (for a review, see MacLeod 1991). This effect is produced when people are incapable of ignoring the meaning of a word when they are asked to name the color of the ink in which the word is printed. In a typical Stroop task, there are three experimental conditions. In the *incongruent condition*, the participants have to reply *blue* if the stimulus presented is the word RED written in blue ink (i.e., $\mathrm{RED}_{\mathrm{BLUE}}$[1]). That is, in this condition, the word and the ink in which it is written do not coincide. The condition in which the word and the color coincide (i.e., $\mathrm{BLUE}_{\mathrm{BLUE}}$) is called *congruent*. For the third condition, called *control*, different types of stimuli have been used. Thus, Stroop used colored squares in his second experiment and a non-alphanumeric sign in his third experiment because he thought that a sign looked more like letters than a square (Stroop 1935). Other authors have used sets of X, words unrelated to color, or pseudowords (MacLeod 1991).

In the incongruent experimental condition, the mean reaction time (RT) to perform the task increases in comparison with the control condition. The difference between the RTs obtained in the incongruent task and the control task is called the *Stroop interference effect.* Sometimes, the experimental results show (Dalrymple-AIford 1972; Dyer 1973, 1974) a decrease in RT in the congruent condition compared to the control condition (*Stroop facilitation effect*).

Three explanations of the interference effect have been proposed in the literature (Melcher and Gruber 2009): (a) incompatibility of the response may provoke a *response conflict*. The interference appears because there are two mutually incompatible responses (Keele 1972), a correct and an incorrect response that are competing. This is a motor conflict. For example, faced with the stimulus $\mathrm{RED}_{\mathrm{BLUE}}$, the word presented (i.e., RED) elicits a motor

G. De Marchis (✉)
Universidad Complutense de Madrid, Madrid, Spain
e-mail: giorgiodemarchis@ccinf.ucm.es

M. P. Rivero Expósito · J. M. Reales Avilés
Universidad Nacional de Educación a Distancia, Madrid, Spain

[1] Hereafter, the word in uppercase letters refers to the name of the color, and the subindex refers to the color of the ink.

behavior that is incompatible with the task of reading the color (in this case, $_{\mathrm{BLUE}}$). Two responses are activated, which are incompatible with the incongruent condition. For the same reason, in the congruent condition (e.g., $\mathrm{RED}_{\mathrm{RED}}$), the response is facilitated because two possible compatible responses are activated. The word and the color are processed in parallel until the person must initiate the response; (b) other authors (Milham and Banich 2005) have proposed that the attribute that is irrelevant to the task may act like a source of color information, and it may therefore compete with the task-relevant color information. In the above example (i.e., $\mathrm{RED}_{\mathrm{BLUE)}}$, RED is irrelevant to the task, but it is a source of color information that competes with the task (i.e., naming the color of the ink and saying $_{\mathrm{BLUE}}$). This interference effect is thought to occur at an early level of processing, that is, at the attentional level, because the irrelevant information withdraws cognitive resources from the task. A third explanation (c) of the interference is that the semantic incongruence between the word and the color may lead to conflict because in the incongruent condition, the meaning of the word (i.e., RED) activates simultaneous mental representations that are incongruent with the activation of mental representations of the color (i.e., $_{\mathrm{BLUE}}$). The magnitude of the semantic conflict will depend on the conceptual distance between the two representations (Van Veen and Carter 2005), and therefore, interference will occur at the semantic level. The three explanations are not mutually exclusive.

The purpose of the present research is to provide more accurate and quantitative comprehension of the semantic gradient in the Stroop interference. Our aim is to confirm that the semantic component of the Stroop interference depends exponentially on the psychological distance between the word and color concept. Following the lead of Taagepera (2008: 24), we think that it is not enough to test directional predictions; rather, it is necessary to ascertain the functional form of the relationship between the variables of interest. We have used the exponential model inspired by the seminal article of Shepard (1987). In this way, we also step away from the linear relationships frequently used in social sciences, as we do not think they are logically sound (Taagepera 2008: 24) in this area.

The semantic gradient in the Stroop task

It has been found that "The amount of interference decreases with increasing semantic distance" between the meaning of the word and the color concept (Roelofs 2003: 89). Thus, $\mathrm{TOPIC}_{\mathrm{RED}}$ produces less interference than $\mathrm{SKY}_{\mathrm{RED}}$, and the latter produces less interference than $\mathrm{BLUE}_{\mathrm{RED}}$ (Dalrymple-AIford 1972; Klein 1964). This relationship is known as the *semantic gradient* or *semantic effect*. The work of Klein was pioneer in the study of how the meaning of the words can affect the Stroop interference pattern. Klein's goal was to study why words interfere with color naming, and for this purpose, he manipulated the stimuli that contained the colors to be named. This author assumed that the semantic structure of the terms used produced an effect in the RT. In Experiment 1, he manipulated the relation of four words with the four colors used (i.e., red, green, yellow, blue) and used six conditions: (a) nonsense syllables (e.g., evgjc); (b) infrequent words (i.e., abjure); (c) words not associated with the color names either in meaning or response category (i.e., friend); (d) words perceptively or semantically related to the colors presented (i.e., blood with red); (e) words of colors that did not form a part of the possible responses (i.e., gray); and (f) a standard condition with the names of the colors presented (i.e., $\mathrm{RED}_{\mathrm{GREEN}}$) in incongruent color-word combinations. Klein found an ordinal relation between these conditions when showing that the interference increased with the increase of the strength of the semantic relationship between the word and the color.

Other authors have replicated and studied this effect in more detail. For example, Dalrymple-Alford and Azkoul (1972) confirmed that words related to the colors presented (i.e., $\mathrm{BLOOD}_{\mathrm{GREEN}}$, when red was presented in the experiment) interfered more than words unrelated to colors (i.e., $\mathrm{JOY}_{\mathrm{GREEN}}$); moreover, words related to the color (i.e., $\mathrm{BLOOD}_{\mathrm{RED}}$) provoked faster responses (i.e., facilitation) than unrelated words (i.e., $\mathrm{JOY}_{\mathrm{RED}}$).

The experimental results (Klein 1964; Proctor 1978) show that interference decreases if the word refers to a color that is not present in the experiment. This difference in the degree of interference is known as the *response set effect*. For example, BLUE written in red (i.e., $\mathrm{BLUE}_{\mathrm{RED}}$) would provoke less interference than $\mathrm{GREEN}_{\mathrm{RED}}$ when no stimulus in the experiment was presented in blue ink. Proctor showed that for words of colors that did not belong to the response set, the amount of interference depends on the strength of the association of the words with the color concept and with word frequency. The response set effect, therefore, reflects the selective activation of the memory of the words of colors belonging to that response set. Fox and Shor (1976) and Fox et al. (1971) found a similar semantic gradient when they used numbers and spatial positions, thus revealing that the effect found by Klein can also occur in relations to other than verbal ones. Perhaps one of the clearest demonstrations of the effect of meaning on interference was performed by Alperson (1967). This author trained a series of participants to pair different terms with colors. When the participants had to perform a Stroop task in which the stimuli to be ignored were the terms previously paired with colors, the semantic relationship had a significant effect on the RTs. In a similar vein, Pritchat

(1968) asked a series of experimental participants to associate nonsense syllables with colors. Subsequently, these syllables were presented in a Stroop task, obtaining usual interference pattern in the semantic gradient. Stirling (1979) performed a similar experiment. A group of participants learned to emit certain letters in response to colored forms. Later, when the participants had to perform a Stroop task with these letters (i.e., name the color without paying attention to the letters), interference was found. In a more recent investigation, Risko et al. (2006) studied in depth the membership of the set of colors presented. These investigators used words related to colors in the set (i.e., *snow* for *white*), words related to colors that were not members of the set (i.e., *earth* with brown when *brown* did not appear as a color to be named), and words unrelated to colors (i.e., *seat*). The results showed that the words related to colors in the set interfered more than the words related to colors that were not present in the stimulus set, and the latter interfered more than the words unrelated to colors.

All these experimental results reveal the consistency and generality of the semantic gradient in the Stroop effect.

Structures of color in the semantic memory

As mentioned, at least part of the Stroop interference effect seemed to be due to the relationship between the words presented and the color concept. Specifically, the semantic effect was a consequence of the stronger or weaker semantic relationship of any word with the color concept. However, till now, research in the literature on the semantic gradient has not measured the psychological distance between the words in the color dimension, and therefore, the relationship found for the effect of the semantic gradient have been simply ordinal relations. The measurement of the relationship of each word with the color dimension would allow specifying the function that relates the amount of interference to the psychological distance. Therefore, our initial goal is to establish the psychological distance of the stimulus compound from the color concept. For this purpose, we used the multidimensional scaling technique (MDS) to qualitatively measure the relationship between two elements and the underlying dimensions in psychological space (Ekman 1963; Kruskal 1964a, b).

Color and MDS

Color has three basic psychological properties: *hue* (also known as *shade* or tone), which ranges from red to violet, *saturation* or hue purity, and *brightness* or amount of white in a color. These psychological properties correspond approximately to the physical properties of wavelength, homogeneity, and intensity of the light waves. The hue of a color is determined by its wavelength. Thus, when using the word "color" colloquially, we usually mean hue. Ekman (1954) used MDS to study the main dimensions of color vision, specifically the dimension of hue. He presented colors in combinations of 2 and asked 31 participants to judge the similarity of each one of the 91 combinations of colors presented, on a scale of 0–4. He specified that the comparison was qualitative so that the participants would value hue and not brightness or saturation. When comparing the results of MDS and the spectrum of visible light, Ekman found five factors that corresponded to red, green, blue, yellow, and violet, and he proposed that human beings use five principal sensations to compare colors with each other, and this discrimination is a function of the wavelength (Ekman 1956).

Universal law of generalization

In a seminal article, Shepard (1987) proposed the universal law of generalization. This author formulated the law from twelve gradients in which the distances in psychological space, measured by means of MDS, matched the empirical generalization values. The data used proceeded from various investigators who had measured both visual and auditive stimuli in human beings and in animals. Shepard proposed that the two variables, generalization function $g(d)$ and normalized distance d in psychological space, are exponentially related to each other in different spheres. For the one-dimensional case, he suggested that the distribution of the probability density that explains the relationship is an Erlang density function with scale parameter μ (Eq. 1).

$$g(d) = \exp\left(-2\frac{d}{\mu}\right) \tag{1}$$

Taking the above into account, we need to further our knowledge about the relationships between the semantic gradient in the Stroop task and the universal law of generalization. Accordingly, our hypothesis is that, as it depends on the psychological distance between the words and the color concept, the semantic effect should match the same type of exponential equation, because it represents a kind of generalization.

In order to test this, we conducted an investigation the first goal of which was to analyze the psychological distance between each color word and the color concept itself. If the semantic gradient is affected by the semantic relationship between the *target stimulus* (i.e., the colored ink) and the *distracter stimulus* (i.e., the word), then the closer a word is to the color concept, the more interference there will be in the response, and therefore, the mean RT will increase according to the exponential function.

To our knowledge, there are no previous works that study the semantic gradient of the Stroop effect in Spanish. Therefore, first, we shall replicate the results found in the non-Spanish literature and verify that the words we are using present the semantic gradient. As there are many psycholinguistic variables that can affect the RT when words are used—from the use of written accent (Gutiérrez-Palma and Palma-Reyes 2008) to the frequency of the words (Langlois 1974)—we have made an effort to control all these variables in our stimuli set.

Second, we shall verify the assumption that the semantic relation between the color concept and the words we used in the study of the semantic gradient is one-dimensional. For this purpose, we shall use the MDS technique. The advantage of using the MDS technique is that it allows us to measure the psychological distance between the color words on an interval scale, and this, in turn, allows us to find the mathematical function that presents the best fit between the positions of the color word in the MDS and the RT means in the incongruent Stroop condition. Following the law proposed by Shepard, the mathematical function that we fit to the data was the Erlang function.

Experiment 1

Method

Participants

A total of 21 students (4 men and 17 women), aged between 19 and 23 years ($M = 19.9$, SD $= 1.34$), from the Complutense University of Madrid took part in this experiment and received academic credits for their participation. They all had normal or corrected-to-normal vision.

Materials and stimuli

We selected 16 words, 4 for each of the following conditions: (a) words unrelated to colors or neutral words (i.e., rhythm, center, trail, signature); (b) words related to colors present in the response set (i.e., fire, field, sky, cacao); (c) words of colors not present in the response set (i.e., gray, black, tan, lilac); and (d) words of colors in the response set (i.e., red, green, blue, brown). The color words of the response set (d) were presented both in the congruent (i.e., RED_{RED}) and the incongruent conditions (i.e., $\text{RED}_{\text{GREEN}}$). In the condition of words related to colors (b), the stimuli could also be classified as congruent (i.e., FIRE_{RED}) and incongruent (i.e., $\text{FIRE}_{\text{GREEN}}$). The remaining conditions (a and c) did not allow this classification. The participants' task was to name the color of the ink in which the word was printed.

The mean length of the distracter words in the Stroop task ranged between 4 and 6 letters ($M = 4.88$, SD $= 0.62$), with a frequency range between 1.25 and 153.21 per million ($M = 55.17$, SD $= 47.55$). The frequency values were obtained from the psycholinguistic database B-Pal (Davis and Perea 2005). The frequency of the words did not present significant differences among the four conditions, $F(3, 12) = 0.23$, MSe $= 621.47$, $p > 0.05$. The words of the diverse conditions were randomly presented both in the incongruent and in the congruent condition. The total percentage of trials presented in the incongruent and congruent conditions of words in the response set was 60 and 40 %, respectively. As accented words can provoke some reading delay because they require more effort (Protopapas 2006), it is advisable to use as few as possible. Nevertheless, some investigators have retained the accent when it is required by grammar (Gutiérrez-Palma and Palma-Reyes 2008). In our case, we only used one word that has an accent in Spanish, and it was retained (i.e., "marrón" [brown]).

Procedure

The stimuli were presented individually and randomly on a 17" CRT monitor with a 100 Hz refresh rate. The experiment was programmed in Inquisit (Inquisit 2.0.61004.7 [Computer software] 2008). The words were projected in lowercase and boldface (Courier new, 17), and the participants were seated 60 cm from the screen, which was against a white background.

Upon arrival at the laboratory, the participants sat down in front of the computer. After reading the instructions on the screen, which emphasized the need to be precise and fast, trying to not commit any errors, they performed three experimental blocks. After completing each block, they received feedback about their performance to prevent a drop in the precision of the responses (Kole et al. 2008). The first

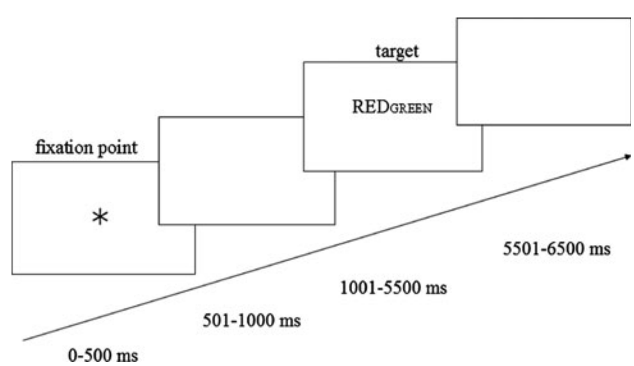

Fig. 1 Outline of the presentation of the trials

block, considered practice for the gradient effect in Stroop, was different from the next two (blocks two and three), in that, it was shorter (20 trials vs. 168 in the remaining two blocks), and the participants were informed each time they made a mistake. The participants' task was to name out loud the color in which each one of the words was written. RT was measured by means of a vocal key that recognized spoken words. Thus, the computer classified the response as erroneous or correct. Figure 1 presents an outline of the sequence and the presentation times of each one of the trials.

Results and discussion

Responses with an RT higher than 4.500 ms were excluded from the analysis (0.67 % of the total). There were no responses with an RT lower than 200 ms. For the subsequent analyses, neither the errors (0.91 % of all the trials) nor the congruent trials were taken into account. Table 1 presents the mean RTs and standard deviations for each condition. In Table 1, a gradient in the RTs of the experimental conditions was observed, in the sense that the RTs were arranged as follows: neutral words < words related to colors < words of colors not in the set < words of colors in the set. This arrangement reproduces the one found by Klein (1964).

We conducted repeated measures ANOVA with the manipulated gradient factor with four levels (neutral words, words related to colors, words of colors not in the set, words of colors in the set). As the gradient factor is within-subject, we used the Greenhouse-Geisser correction for the degrees of freedom. The results showed that the main factor gradient was highly significant, $F(2.04, 40.88) = 65.19$, $\text{MSe} = 1{,}027.02$, $p < 0.001$, $\eta^2 = 0.76$. We applied a posteriori comparative analysis with Bonferroni's test. The differences were significant in all the comparisons ($ps \leq 0.005$) (see Fig. 2).

The results obtained reveal the existence of a semantic gradient in the mean RT for the diverse experimental conditions, thus replicating previous results (Klein 1964).

Experiment 2

In Experiment 1, we replicated the existence of the semantic gradient in Spanish. In the order relation of the semantic gradient, we are implicitly assuming the existence of a unique dimension in which the diverse elements are located with regard to the color dimension; that is, the words used to find the gradient are unidimensionally related to the color concept. The goal of this experiment was to scale the words used in the first experiment along this perceptual/semantic dimension. We scaled the words by imposing a one-dimensional solution to subjects' ratings in an MDS analysis. This means that MDS was not used to establish the number of dimensions but to scale the words unidimensionally.

Table 1 Mean reaction times (in ms) and their standard deviations (SD) for each condition

Condition	*M*	SD
Neutral words	621.94	162.68
Words related to colors	654.53	178.53
Words of colors not in the set	678.87	182.96
Words of colors in the set	733.36	223.34

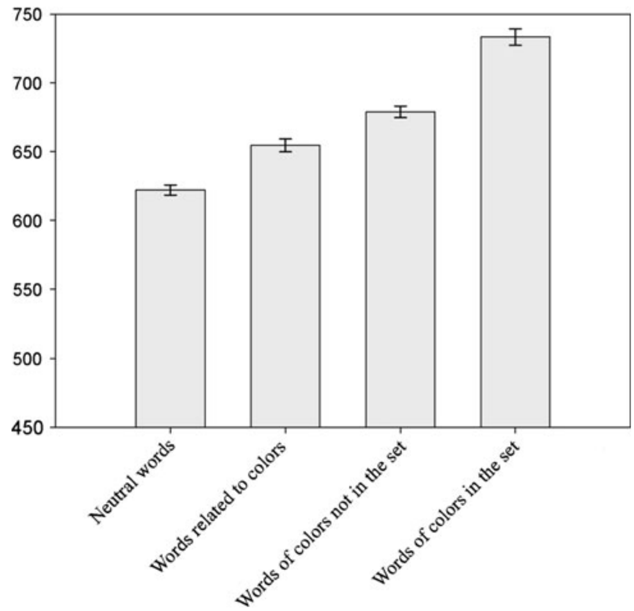

Fig. 2 Histogram of the conditions and their mean RTs. *Error bars* represent standard errors of the mean

Method

Participants

Participants in this study were 54 students (19 men and 35 women), aged between 20 and 47 years ($M = 22.22$, $SD = 4.45$), from the Complutense University of Madrid. They received academic credits for their participation. The design was between-subject to obtain responses about psychological values from a relatively large number of participants.

Materials and stimuli

We used a paper-and-pencil questionnaire to obtain the data. The same 16 terms used in the previous experiment were presented. This number of terms is sufficient to achieve stable perceptual maps of up to three dimensions

(Hair et al. 1999), even though we only used one dimension.

Procedure

Before the experiment, the participants were informed that their responses to the questionnaire were anonymous, and the responses were neither correct nor incorrect, because we only wished to know their subjective impressions. They were requested in the instructions to appraise the degree of dissimilarity (or difference) of each term with the reference concept "color," on a range of 0–100. The value 0 meant that the word had a maximum relationship with the concept of "color," whereas the value 100 meant that the word had no relationship. The words were randomly presented in two inverse orders. Lastly, participants were reminded that it was essential to appraise all the words without skipping any. This classification method allowed us to consider the data as originally metric (Hair et al. 1999).

Results and discussion

The mean score given to each one of the stimuli by the participants was the dependent variable. The dissimilarity matrix of the Euclidean distances among the variables was calculated. With this new matrix, we conducted an aggregate MDS analysis with the PROXSCAL algorithm (Garson 2012), forcing the program to accept that the data represented a single dimension. The results indicated that this single dimension reached a fit index or stress of 0.13, a Tucker's congruence coefficient of 0.99, and a dispersion accounted for (D.A.F.) of 0.98.

The results obtained through unidimensional scaling by MDS showed that participants effectively referred to a single dimension when performing the requested assessment task in compliance with instructions. We will call this dimension "Color."

Table 2 shows the values of the perceptual distance of each one of the words to the color concept. The terms are placed on the dimension as a function of three large blocks that are coherent with Klein's (1964) semantic gradient: (a) color words (red, green, blue, brown, gray, black, tan, lilac); (b) words related to color but not color names (cacao, field, fire, sky); and (c) words unrelated to the concept of color (center, trail, signature, rhythm).

Unidimensional scaling by MDS reproduced the order underlying the semantic gradient.

The gradient does not foresee any order of the color words, but only of the conditions. An analysis of the results of the scaling indicated that the basic colors[2] are perceived as "more color," which is coherent with the findings of Ekman (1954) and with the idea of Proctor (1978) that the sources of interference in the semantic gradient are the strength of the association and word frequency. In order to discard the possibility that the position of the stimuli on the MDS could be related to their linguistic frequency, we performed a correlational analysis between the MDS scores and linguistic frequency according to the B-Pal dictionary (Davis and Perea 2005). The results showed that this relation did not reach significance ($r = 0.15$, $p > 0.05$).

Table 2 Values of the dimension "color" through scaling, calculated with MDS

Words	Values
Red	−0.63
Green	−0.63
Blue	−0.62
Brown	−0.61
Black	−0.58
Tan	−0.52
Lilac	−0.51
Gray	−0.43
Sky	−0.16
Cacao	0.05
Fire	0.12
Field	0.25
Rhythm	0.86
Signature	1.10
Center	1.14
Trail	1.19

Functional relation between the MDS distance and mean RT

It has been shown that the semantic gradient is compatible with an order relation; however, we wished to establish the functional relation between the position of the stimuli according to the MDS and the mean RTs. For this purpose, the results obtained in the previous two experiments were combined to investigate the concrete form adopted by their functional relation. Going back to the seminal work of Shepard (1987), who proposed the exponential function as the metric distance in psychological space for any set of stimuli, and the data were fitted to an exponential function (Eq. 2). The equation we used to fit the mean RTs to the

[2] The photoreceptors of color called "cones" can be divided into three types—blues, greens, and reds—and, therefore, we can consider them to be the basic colors of sight. The activation or inhibition of these different types of cells depends on spectral opponency, which can be roughly classified into two types: blue–yellow contrast and red–green contrast (Mather 2006). These colors correspond to those found by Ekman (1956).

positions of the stimuli in the MDS is more general than the one proposed by Shepard (Eq. 1).

$$g(x) = a + b^{cx} \quad (2)$$

We considered the argument of the function 2 (x) to be the position of the stimuli in the MDS found in Experiment 2, whereas the value of the function $g(x)$ was the mean RT in the conditions of Experiment 1. The values a, b, and c in Eq. 2 were the parameters fitted to the data. Data were also fitted to the linear and quadratic functions, to compare the fit of various models.

Analysis of the results

We fitted the data with the NonLinearFit function of Mathematica. The equation that best fit the data to the proposed exponential function was:

$$g(x) = 637.42 + 13.32^{2.81x} \quad (3)$$

Figure 3 presents the data as well as the exponential function with least-squares fitting. We used the statistical index R^2 as a measure of goodness of fit. As the postulated relation is nonlinear, we calculated R^2 as the quotient between the difference of the total corrected sum of squares and the residual sum of squares, divided by the total corrected sum of squares. The model was highly significant, $F(3, 13) = 8{,}970.16$, MSe $= 269.52$, $p < 0.001$. The fit index reached a very satisfactory value ($R^2 = 0.87$). This value of fit for the exponential equation was better than the fit of the other models that were compared and that were also significant: the linear equation, $F(1, 14) = 36.34$, MSe $= 552.14$, $p < 0.001$, $R^2 = 0.72$, and the quadratic equation, $F(3, 13) = 5{,}411.54$, MSe $= 446.61$, $p < 0.001$, $R^2 = 0.79$.

In order to better visualize the results graphically, we performed a linear transformation of the results of the psychological dimension (Hair et al. 1999), multiplying

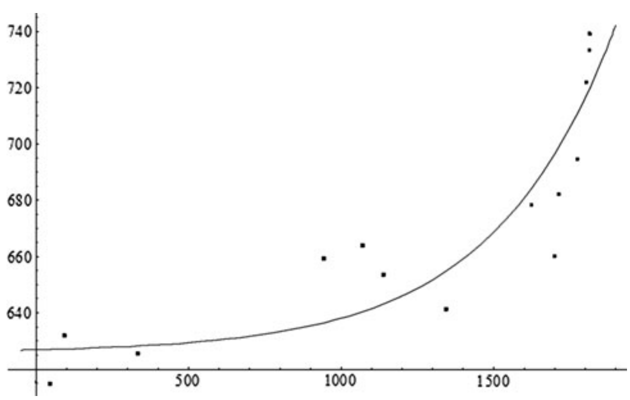

Fig. 3 Fit of the RTs to the psychological space values after the linear transformation. The abscissa axis corresponds to the MDS results, and the ordinate axis corresponds to the mean RT

each item by 1,000 and subsequently adding 1,189.36, so that the lowest value of the MDS would coincide with the value 0 for the abscissa axis (see Fig. 3).

We can, therefore, propose that the semantic gradient follows the universal law of generalization, so that the mean RTs in a Stroop task are exponentially related to the distance between the stimulus and the color concept.

Experiment 3

In order to further validate these results, we ran a third experiment. Our goal was to evaluate the robustness of the results with a different set of stimuli also related to the color dimension. We must take into account that the set of color words is very limited in number. Therefore, the stimuli used in this experiment are not as prototypical color words as the words used in the first experiment. We tried to control as much as possible all the psycholinguistic variables of the various sets of words, as in our previous experiment.

Method

Participants

Eleven Complutense University undergraduates participated in exchange for course credits. One participant was excluded for not speaking Spanish as his first language. The rest of participants reported normal or corrected-to-normal vision and spoke Spanish as their first language. Therefore, the data in this study come from 10 students (2 men and 8 women) aged between 19 and 26 years ($M = 20.70$, SD $= 2.83$).

Apparatus

We used the same devices and procedure as in the previous experiments.

Stimuli

We used four sets of stimuli of three words in each (i.e., words of colors in the set, words of colors not in the set, words semantically related to colors, and words not related to colors). The full set of twelve words was different from the words used in previous experiments. The words of colors in the set included the Spanish words for "white," "yellow," and "purple" ("Blanco," "Amarillo," and "Morado," respectively). These were also the target colors. The words of colors not in the set included the Spanish words for "orange," "pink," and "garnet" ("Naranja", "Rosa," and "Granate," respectively). The words semantically related to colors included the Spanish words for "snow," "lemon," and "eggplant" ("Nieve", "Limón," and "Berenjena," respectively). Finally, the words not related to colors included the Spanish words for "shop window," "guard," and "breeze" ("Vitrina", "Guardia," and "Brisa," respectively). The screen background was black because the color white appeared in the set of target colors. The words were randomly presented both in the incongruent and in the congruent condition for all sets of stimuli. The total percentage of trials presented in the incongruent and congruent conditions of words in the response set was 64.7 and 35.3 %, respectively.

The mean length of the distracter words in the Stroop task ranged between 6 and 8 letters ($M = 6.66$, $SD = 1.15$), with a frequency range between 5.18 and 116.25 per million ($M = 48.81$, $SD = 59.24$). The frequency of the words did not present significant differences among the four conditions, $F(3, 8) = 0.55$, $MSe = 1{,}227.67$, $p > 0.500$.

Procedure

The procedure was similar to the one used in Experiment 1, but this time a within-subjects design was used. Participants responded to the MDS questionnaire before doing the Stroop task. Afterward, the participants were given two blocks to perform the Stroop task: first a learning block as in Experiment 1 and two experimental blocks of 100 trials each.

Results and discussion

Responses with RTs higher than 4.500 ms (0.3 %) and lower than 200 ms (0.04 %) were excluded from further analysis. For subsequent analyses, errors (0.27 % of all trials) and congruent trials were excluded. Table 3 presents the mean RTs and the SDs for each condition. As in Table 1, a gradient is observed in Table 3; in that, mean RTs matched the expected ordinal relation: words of colors in the set > words of colors not in the set > words related to colors > neutral words. The arrangement of the means reproduced the results from Experiment 1 and the results found by Klein (1964).

Table 3 Mean reaction times (in ms) and their standard deviations (SD) for each condition

Condition	*M*	SD
Neutral words	600.56	160.77
Words related to colors	628.21	176.70
Words of colors not in the set	652.80	197.45
Words of colors in the set	708.10	271.72

We conducted repeated measures ANOVA with the manipulated gradient factor with four levels (i.e., neutral words, words related to colors, words of colors not in the set, words of colors in the set). As the gradient factor was manipulated within-subjects, we used the Greenhouse-Geisser correction for the degrees of freedom. The results showed that the main effect of Gradient was highly significant, $F(2.53, 22.75) = 26.51$, $MSe = 951.62$, $p < 0.001$, $\eta^2 = 0.75$. We applied a posteriori analysis with the LSD test. The differences were significant in all the comparisons ($ps \leq 0.05$), but only between words related to colors and words of colors in the set ($p = 0.17$).

The analysis of the distance data with MDS PROXSCAL algorithm was performed, forcing the program to accept that the data represented a single dimension, as in the previous experiment. The results indicated that this single dimension reached a fit index or Stress-I of 0.14, a Tucker's congruence coefficient of 0.99, and a dispersion accounted for (D.A.F.) of 0.98. Table 4 shows the results.

The mean RT for each word and the corresponding MDS values were fitted to the exponential, the quadratic, and the linear functions. The equation that best fit the data was the exponential function. The parameters appear in Eq. 4.

$$g(x) = 610.83 + 69.01^{1.87x} \quad (4)$$

Figure 4 presents the data as well as the exponential function with least-squares fitting. The model was highly significant, $F(3, 9) = 877.07$, $MSe = 1{,}919.57$, $p < 0.001$. We used the statistical index R^2 as a measure of goodness of fit. The R^2 fit index reached a value of 0.54. This value of fit for the exponential equation was better than the fit of the other models that were tested and that were also significant: the linear equation, $F(1, 10) = 5.40$, $MSe = 2{,}439$, $p < 0.05$, $R^2 = 0.35$, and the quadratic equation $F(3, 9) = 720.01$, $MSe = 2{,}336.54$, $p < 0.001$, $R^2 = 0.44$.

Table 4 Values of the dimension "color" computed through MDS

Words	Values
Yellow	−0.598
Orange	−0.598
Garnet	−0.532
White	−0.498
Purple	−0.485
Pink	−0.458
Lemon	−0.270
Eggplant	−0.043
Snow	0.246
Breeze	0.870
Guard	1.146
Shop window	1.219

In order to better visualize the results, we performed a linear transformation of the results of the psychological dimension (Hair et al. 1999), changing the sign of the MDS values (see Fig. 4).

The results of the third experiment showed the same patterns observed in the previous experiment with some qualifications. That is, the conditions roughly followed the gradient effect, the color words were located on a single dimension, and the exponential function fit better than linear or quadratic functions. So, we think that the gradient effect and the fit to the exponential function are both quite robust effects. We observed two main differences in the results with respect to the previous experiments. The color words not in the set did not differ significantly from the words semantically related to colors, although the means followed the expected pattern. Another difference was that the fit of the exponential function achieved a significant but lower R^2 in comparison with previous experiments. The differences could be explained by the lower color prototypicality of the words used in the third experiment, when compared with the words used in the previous experiments. That is, the color words red, green, blue, and brown are more typical colors than white, yellow, and purple.

General conclusions

In the experiments presented, we studied the semantic gradient of the Stroop effect. The results of non-Spanish literature were replicated to verify that the words used presented the semantic gradient. By means of the MDS technique, we measured the psychological distance between the color words on an interval scale. This relationship was shown to be one-dimensional. Lastly, we sought the mathematical function that best fit the data of the psychological distance to the mean RTs in the incongruent Stroop condition. In order to perform this fit, we used the exponential function, according to the law proposed by Shepard, and also the linear and the quadratic functions.

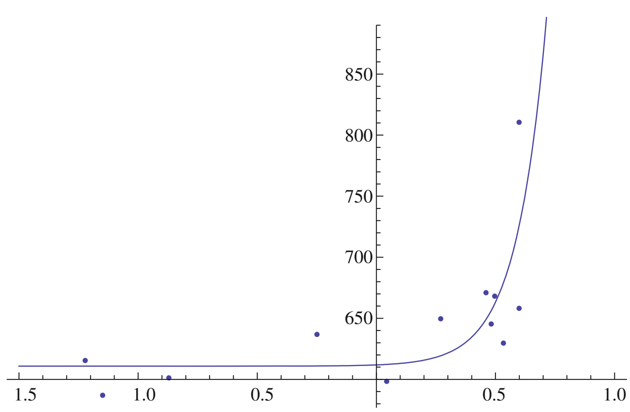

Fig. 4 Fit of the RTs to the psychological space values in Experiment 3 after the linear transformation

We found a semantic gradient in Spanish with significant differences in all the experimental conditions. In the first experiment and in all but one condition in the third experiment, where the color words not in the set did not differ significantly from the words semantically related to colors, although the means followed the expected pattern. This result suggests that the semantic gradient is not a language-specific effect.

The semantic gradient shown in the first experiment arranged the conditions from lowest to highest, but it did not indicate the distance between conditions. Implicitly, it was considered that the semantic gradient is measured on an ordinal scale; however, in our study, we proposed the possibility of measuring these distances by means of MDS, which is a tool to measure the psychological dimensions underlying a set of stimuli. Therefore, we considered it adequate to transform an ordinal variable into a ratio variable, so that we could subsequently specify the quantitative relation between psychological distance and mean RT. It would not be possible to establish this relationship if one of the scales was ordinal.

Our results with MDS, both in the first and in the third experiment, confirm that the psychological distance between each concept and the color concept is one-dimensional, and they provided the scores for each stimulus on that scale. The scores of stimuli place the words in the same order that has been found traditionally.

The psychological dimension (MDS) and the temporal dimension—RTs—fit an exponential curve better than a linear or a quadratic curve. The goodness of fit to the exponential curve allows us to state that the semantic effect follows an exponential function. We have shown that the amount of interference in a Stroop task increases when the semantic distance from the color concept decreases. To our knowledge, this is the first time that the functional relation between the distance of a word to the color concept and the mean RT has been shown. The fact that the relationship we have tested is exponential due to the fact that we used Shepard's first universal law of generalization as the theoretical framework.

We have proposed a more generic equation than the one suggested by Shepard, with specific values, whose validity and range will have to be confirmed in future studies. We suggest that this equation can be applied to the relation between semantic memory and RTs, as indicators of the strength of the association. Our results support the semantic conflict theory of Van Veen and Carter (2005), adding that the conflict depends exponentially on the conceptual distance.

The present investigation provides the first confirmation of the existence of the gradient using Spanish terms and also the first demonstration that the gradient fits an exponential function, and not a linear or a quadratic one. The present outcomes add to the large body of empirical results associated with the Stroop effect, thus providing additional perspectives on the evolving theories of Stroop interference. For example, the existing theoretical models (Cohen et al. 1990; Logan 1985; Melara and Algom 2003) should be able to reproduce this exponential relation to fit the data. Future research should focus on the implications of our results for the theories of the Stroop effect. It is also necessary to expand the number of terms used to obtain a better fit to the exponential curve and to determine the importance of the diverse components of the Stroop effect. Furthermore, if our results are replicated, it would be necessary to find a theory that explains both the specific relationship found between the position of words in the color dimension and mean RT, and the meaning of the parameters of the exponential function. Finally, these results should be related to the conceptual network theories for the semantic memory as we suspect that this specific relationship we have found could be explained by some attributes of the semantic network underlying the gradient effect.

References

Alperson BL (1967) The effect of semantic relatedness and practice on the color-word test. Michigan State University, East Lansing

Cohen JD, Dunbar K, McClelland JL (1990) On the control of automatic processes: a parallel distributed processing account of the Stroop effect. Psychol Rev 97:332–361

Dalrymple-AIford EC (1972) Associative facilitation and interference in the Stroop color-word task. Percept Psychophys 11:274–276

Dalrymple-Alford EC, Azkoul J (1972) The locus of interference in the Stroop and related tasks. Percept Psychophys 11:385–388

Davis CJ, Perea M (2005) BuscaPalabras: a program for deriving orthographic and phonological neighbourhood statistics and other psycholinguistic indices in Spanish. Behavior Research Methods 37(4):665–671. doi:10.3758/BF03192738

Dyer EN (1973) Same and different judgments for word-color pairs with "irrelevant" words or colors: evidence for word-code comparisons. J Exp Psychol 98:102–108

Dyer EN (1974) Stroop interference with long preexposures of the word: comparison of pure and mixed preexposure sequences. Bull Psychon Soc 3:8–10

Ekman G (1954) Dimensions of color vision. J Psychol 38:467

Ekman G (1956) Discrimination of hue as a function of wave length. Acta Psychol 12:15–18

Ekman G (1963) A direct method for multidimensional ratio scaling. Psychometrika 28:33–41

Fox LA, Shor RE (1976) Semantic gradients and interference with sorting according to color, spatial position, and numerosity. Bull Psychon Soc 7(2):187–189

Fox LA, Shor RE, Steinman RJ (1971) Semantic gradients and interference in naming color, spatial direction, and numerosity. J Exp Psychol 91(1):59–65

Garson GD (2012) Multidimensional scaling. Retrieved 10 Feb 2012, from http://faculty.chass.ncsu.edu/garson/PA765/mds.htm

Gutiérrez-Palma N, Palma-Reyes A (2008) On the use of lexical stress in reading Spanish. Read Writ Interdiscip J 21:645–650. doi:10.1007/s11145-007-9082-x

Hair JFJ, Anderson RE, Tatham RL, Black WC (1999) Análisis multivariante, 5th edn. Pearson Educación, Madrid

Inquisit 2.0.61004.7 [Computer software] (2008). Millisecond Software, Seattle, WA

Keele S (1972) Attention demands of memory retrieval. J Exp Psychol 93:245–248

Klein GS (1964) Semantic power measured through the interference of words with color-naming. Am J Psychol 77:576–588

Kole JA, Healy AF, Bourne LE Jr (2008) Cognitive complications moderate the speed-accuracy tradeoff in data entry: a cognitive antidote to inhibition. Appl Cogn Psychol 22:917–937. doi:10.1002/acp.1401

Kruskal JB (1964a) Multidimensional scaling by optimizing goodness to fit to a nonmetric hypothesis. Psychometrika 29:1–27

Kruskal JB (1964b) Nonmetric Multidimensional scaling: a numerical method. Psychometrika 29:115–129

Langlois J (1974) Frequency of occurrence as a factor in interference on the Stroop word-color test. Percept Mot Skills 38:986

Logan GD (1985) Skill and automaticity: relations, implications, and future directions. Can J Psychol 39:367–386

MacLeod CM (1991) Half a century of research on the Stroop effect: an integrative review. Psychol Bull 109(2):163–203

Mather G (2006) Foundations of perception. Taylor & Francis Inc., New York

Melara RD, Algom D (2003) Driven by information: a tectonic theory of Stroop effects. Psychol Rev 110(3):422–471. doi:10.1037/0033-295X.110.3.422

Melcher T, Gruber O (2009) Decomposing interference during Stroop performance into different conflict factors: an event-related fMRI study. Cortex 45:189–200. doi:10.1016/j.cortex.2007.06.004

Milham MP, Banich MT (2005) Anterior cingulate cortex: an fMRI analysis of conflict specificity and functional differentiation. Hum Brain Mapp 25:328–335

Pritchat D (1968) An Investigation into some of underlying associative verbal processes of Stroop colour effect. Q J Exp Psychol 20:351–359

Proctor RW (1978) Sources of color-word interference in the Stroop color-naming task. Percept Psychophys 23:413–419

Protopapas A (2006) On the use and usefulness of stress diacritics in reading Greek. Read Writ Interdiscip J 19:171–178. doi:10.1111/j.1467-9817.2006.00316.x

Risko EF, Schmidt JR, Besner D (2006) Filling a gap in the semantic gradient: color associates and response set effects in the Stroop task. Psychon Bull Rev 13(2):310–316

Roelofs A (2003) Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. Psychol Rev 110(1):88–125. doi:10.1037/0033-295X.110.1.88

Shepard RN (1987) Toward a universal law of generalization for psychological science. Science 237:1317–1323. doi:10.1126/science.3629243

Stirling N (1979) Stroop interference—input and an output phenomenon. Q J Exp Psychol 31(FEB):121–132

Stroop JR (1935) Studies of interference in serial verbal reactions. J Exp Psychol 18:643–662

Taagepera R (2008) Making social sciences more scientific. The need of predictive models. Oxford University Press, New York

Van Veen V, Carter CS (2005) Separating semantic conflict and response conflict in the Stroop task: a functional MRI study. NeuroImage 27(3):497–504. doi:10.1016/j.neuroimage.2005.04.042