Journal of the International Neuropsychological Society (2002), **8**, 819–827.

DOI: 10.1017.S1355617702860106

Stroop effect in Spanish–English bilinguals

MÓNICA ROSSELLI,[1] ALFREDO ARDILA,[2] MIRTHA N. SANTISI,[1]
MARÍA DEL ROSARIO ARECCO,[1] JUDY SALVATIERRA,[1]
ALEJANDRA CONDE,[1] AND BONIE LENIS[1]

[1]Department of Psychology, Florida Atlantic University, Davie, Florida
[2]Communication Sciences and Disorders, Florida International University, Miami, Florida

(RECEIVED February 5, 2001; REVISED November 1, 2001; ACCEPTED November 2, 2001)

Abstract

The aim of this study was to analyze the performance of Spanish–English bilinguals on the Golden Stroop Test. The effects of bilingualism, participants' age, age of acquisition of the second language, and proficiency in each language were analyzed. Participants consisted of 71 Spanish–English bilinguals, 40 English monolinguals, and 11 Spanish monolinguals from South Florida. Proficiency in Spanish and English was established using a self-report questionnaire and the English and Spanish versions of the Boston Naming Test. In bilinguals, the Golden Stroop Test was administered in English and in Spanish. Overall, performance was slower in bilinguals than in monolinguals. No significant differences were observed in color reading but bilinguals performed worse in the naming color condition. Even though bilinguals were 5% to 10% slower in the color–word condition, one-way ANOVAs revealed no statistically significant differences between groups. Within the bilingual group, the Stroop Test scores were similar in both English and Spanish. Age of acquisition of the second language did not predict the Stroop Test performance. Repeated measures ANOVA demonstrated a significant interaction between Language Proficiency × Language (in which the test was administered) in some of the ST conditions. In balanced bilinguals, the language used in the ST did not matter, but in unbalanced subjects, the best-spoken language showed better results. In addition, our results support the presence of both between- and within-language interference in Spanish–English bilinguals. Different conceptualization models of the structure of bilingual memory are disclosed. (*JINS*, 2002, *8*, 819–827.)

Keywords: Bilingualism, Spanish–English bilinguals, Neuropsychology, Stroop Test, Language interference

INTRODUCTION

One of the best-known procedures to study response interference and inhibition is the Stroop Test (ST; Stroop, 1935). In clinical neuropsychology, the ST has frequently been used as an executive functioning task (Lezak, 1995; Spreen & Strauss, 1998). The ST requires participants to name/read colors as fast as they can under three conditions. In one control condition, participants are presented a list of color words printed in black ink and requested to read them. In a second control condition, participants are presented with a list of *XXXX*s printed in different colors and are required to name the color of the ink. In the last condition (color–word) participants are presented with a list of color words printed in ink of an incongruent color and are required to name the color of the ink while ignoring the written word. In this last condition the subjects need to inhibit the automatic response of reading the word (e.g., when presented with the word *RED* written in green they are supposed to answer "green"). Ink color naming is faster when the two dimensions of the stimulus are congruent (e.g., *RED* written in red) than when they are incongruent (e.g., *RED* written in green). That is, the irrelevant, to-be-ignored dimension interferes with processing the relevant dimension. This pattern of response is called the Stroop effect.

The ST has been used to study the effect of simultaneously possessing of two different lexical systems in bilingual participants (e.g., Altarriba & Mathis, 1997; Brauer, 1998; Dyer, 1971; Fang et al., 1981; Gerhand et al., 1995; Heij et al., 1990, 1996; Lee & Chan, 2000; Magiste, 1984; Tzelgov et al., 1990). In the bilingual ST the Stroop affect is studied in two language conditions. In the first condition

Reprint requests to: Mónica Rosselli, Ph.D., Florida Atlantic University, Department of Psychology, 2912 College Avenue, Davie, FL 33314.

there is congruency between the language in which words are written (e.g., English) and the language in which answers are requested (e.g., English). In the second language condition the language of the written words (e.g., English) and the language of the requested answers are incongruent (e.g., Spanish). The Stroop interference seen in the first language condition is called intralingual or within-language interference and the Stroop interference seen in the second language condition is called interlingual or between-languages interference. Research using the bilingual ST has demonstrated that both within-language and between-language interference can be found, but within language interference is usually higher (Magiste, 1984). The difference is attributed to the fact that bilinguals store different languages in different mental dictionaries (Brauer, 1998). When only one language is involved (*within-language* condition) more interference is expected because the bilingual is using only one dictionary while in the between language condition the interference has to go from one dictionary to another. Research has also demonstrated that the degree of interference using the ST in bilinguals depend upon the level of language proficiency and the similarity between both languages (Brauer, 1998). For example, ST performance has been found to decrease in proficient bilinguals of two similar languages (German and English). Furthermore, subjects do better in reducing within language interference in their native language, but not in a low-proficient language (Tzelgov et al., 1990). Finally, research has shown conflicting results when comparing different writing systems. While several authors (e.g., Biederman & Tsao, 1981; Chen & Juola, 1982; Fang et al., 1981; Henderson, 1984; Seinderberg, 1985) have found a greater Stroop effect in logographic (e.g., Japanese, Chinese) than in sonographic writing systems (e.g., English), but Lee et al. (2000) failed to find significant differences when comparing ST in Chinese and English in Chinese–English bilinguals.

The patterns of within-language *versus* between-language interference, using the ST, allow the drawing of conclusions about how bilinguals access and store lexical information. Researchers of bilingual memory have proposed two hypotheses: (1) the word association model assumes that there is a direct link between a bilingual's first language (L1) and his or her second language (L2); (2) the concept mediation model assumes that there is no direct link between languages at the lexical level but that both lexicons are connected to a common semantic representation (Kroll & Stewart, 1994). Studies that show larger within- than between-language interference effects are consistent with the concept mediation model. Studies that demonstrate the influence of proficiency on the size of the interference effects support the word association model (Magiste, 1984). Altarriba and Mathis (1997) found that both conceptual and lexical links are formed for second language words even after a single learning session.

Two experiments have analyzed the ST effect in Spanish–English bilinguals. Dyer (1971) asked 16 Spanish–English bilinguals and 11 English monolinguals to name colors in both languages with color stimuli that were either Spanish color names, English color names, or control *X*s. Greater interference was found in bilinguals than for monolinguals. Color naming was the slowest when the naming language and the language of the color names were different suggesting more between language interference. Altarriba and Mathis (1997) found that novice and expert Spanish–English bilinguals demonstrated the Stroop effect both within and between languages. Thus, interference can be observed in low and high proficient bilinguals in both intra and interlingual conditions. There are still inconsistencies on the role of language proficiency in the ST conditions and little is known about the applicability of ST to bilinguals in clinical settings.

The Stroop Test has a lengthy history as an experimental measure in psychology and only recently has it been adapted for clinical neuropsychology use (Mitrushina et al., 1999). The ST has traditionally been viewed as a measure of executive functioning. Even though six different scores are obtained on the ST (time and errors for the three conditions), the truly informative score from the clinical point of view is the executive functioning score, which refers to the color–word condition (Lezak, 1995; Spreen & Strauss, 1998). This Stroop color–word condition has been found to measure cognitive processes, such as selective attention and flexibility (Glasser & Glasser, 1989). In the compilation of the ST data sets done by Mitrushina et al. (1999) there does not appear to be any data about the effect of bilingualism on ST scores. It may be more important to know if the color–word scores (time and errors) are impacted by bilingualism, and if impacted, what bilingualism variables (e.g., proficiency, age, age of acquisition of the second language, etc.) have to be taken into consideration.

The aims of this study were (1) to analyze within and between language interference using the ST among English and Spanish monolinguals, and Spanish–English bilinguals, and (2) to analyze the impact of language proficiency and age of acquisition of the second language on the ST scores. We predicted, as stated by Magiste (1984), that language proficiency would be responsible for the pattern of within- and between-language interference. The higher the proficiency in a language the higher the level of interference expected in that given language. Previous studies (Altarriba & Mathis, 1997; Dyer, 1971) have usually relied on self-reports to establish language proficiency. In the present study, an objective method to measure level of language proficiency was used.

METHODS

Research Participants

Participants were college students, their family members, and friends. There were 122 right-handed participants: 71 Spanish–English bilinguals, 40 English monolinguals, and 11 Spanish monolinguals. None of the participants had neurological or psychiatric antecedents. All participants had normal scores on the Mini Mental Status Exam (MMSE). No significant age effect was observed among groups (see Table 1). A Bonferroni *post-hoc* comparisons revealed no significant differences ($p = .13$) between the mean age of bilinguals ($M = 31.98$, $SD = 13.14$) and English monolinguals ($M = 35.90$, $SD = 13.80$) and no significant differences ($p = 0.43$) between the mean age of bilinguals and Spanish monolinguals ($M = 40.91$, $SD = 15.17$). Spanish was the first language (L1) spoken by 90.1% of the bilingual participants whereas 7% claimed English as a first language, and 2.8% reported both languages as the "first language." The average age of exposure to the second language (L2) was 11.40 ($SD = 12.06$), and the average number of years using L2 was 18.61 ($SD = 12.06$). Participants who spoke another language, besides Spanish and English, were excluded from the sample. Table 1 presents the general characteristics of the sample.

Characteristics of bilingualism were recorded. As children, 61 (85.9%) bilingual participants spoke Spanish at home, 2 (2.8%) spoke English, and 8 (11.3%) spoke both languages. Forty-three (60.6%) of the bilingual participants preferred speaking Spanish, 17 (23.9%) preferred English and 11 (15.5%) had no special preference for either language. Thirty-six (50.7%) of the bilingual participants reported Spanish as the most spoken language, 25 (35.2%) reported English, and 10 (14.1%) reported speaking both languages equally. Forty-six (64.8%) of the bilingual participants reported Spanish as the best-spoken language, 20 (28.2%) reported English, and 5 (7%) answered both Spanish and English as "the best-spoken language." All bilingual participants were living in the United States and had at least some formal education in English.

Instruments

In addition to the demographic history and bilingualism questionnaire, two different tests were administered: (1) the Boston Naming Test (BNT; Kaplan et al., 1983) was used to assess level of proficiency; (2) the stimuli of the Golden Stroop Color and Word Test (Golden, 1978; Mitrushina et al., 1999) were administered, but the time in seconds to complete all stimuli (instead of number of responses within 45 s in the three conditions) was used as a score. Errors were also scored. A between (interlanguage) language condition for the color–word ST was specifically designed for this study to further test the bilingual participants. In this interlanguage condition the language used in the stimulus (e.g., English or Spanish) and the response language (e.g., Spanish or English) were different. Time and errors for this new condition were also recorded.

Parallel Spanish versions of the BNT (*Test de Vocabulario de Boston*; Kaplan et al., 1996) and of the Golden Stroop Test (Rey, unpublished) were used. The parallel BNT in Spanish has the same number of items that the English version but 12 items were modified because of linguistic (e.g., "toothbrush" in Spanish is *cepillo de dientes*—three words; "noose" does not have a direct and evident name in Spanish, etc.) or cultural reasons (e.g., "pretzel" does not exist in Spanish-speaking countries). The scoring system of both BNT versions is the same (Kaplan et al., 1983, 1996).

Procedure

An informed consent in accordance with APA Ethical Guidelines for research with human subjects was given to all participants in the study. To determine eligibility, all participants were interviewed. The interview consisted of a demographic history and a specific questionnaire assessing language proficiency (Rosselli et al., 2000). Bilingual participants were asked to rate themselves on how well they understood, spoke, wrote, and read Spanish and English using the following scale: 1 (*not at all*); 2 (*limited*); 3 (*relatively well*); 4 (*quite well*); and 5 (*very well*). The interviewers were also proficient bilinguals who were able to corroborate participants' understanding and expression scores in both languages while doing the interview. Only participants who scored *quite well* or *very well* in all four linguistic abilities (speaking, understanding, reading and writing) in both languages were selected.

Bilinguals were grouped into (1) *balanced bilinguals* (*high-proficient* and *low proficient*), and (2) *unbalanced bilinguals* (*English dominant* and *Spanish dominant*). Scores of the BNT in Spanish and English were used as selection criteria. The following procedure was used to assign participants to groups. Participants with a difference between the BNT–Spanish score and BNT–English score equal to or lower than 8 points (1 *SD* according to the whole group

Table 1. Demographic characteristics of the three groups

Variable	English–Spanish bilinguals ($n = 71$)	English monolinguals ($n = 40$)	Spanish monolinguals ($n = 11$)	F	p
Age	$M = 31.98$ (13.14)	$M = 35.90$ (13.80)	$M = 40.91$ (15.17)	2.609	.078
Education (years)	$M = 14.92$ (2.35)	$M = 15.35$ (2.45)	$M = 14.25$ (3.49)	.704	.497
MMSE	$M = 29.53$ (.734)	$M = 29.70$ (.608)	$M = 29.40$ (.516)	1.129	.327
Gender M:F	45% (32):54% (39)	32.5% (13):67% (27)	27.3% (3):72.7% (8)		

Note. Standard deviations in parentheses.

scores) were considered *Balanced Bilinguals* ($n = 25$), and participants with a difference between the two BNT scores greater than 8 points were regarded as *Unbalanced Bilinguals* ($n = 46$). It was noted, however, that some of the participants in the Balanced group had very high scores in the BNT in both languages, while in others the performance was low in both language versions of the test. Thus, we decided to further subdivide the balanced group into two subgroups using the BNT scores. Bilingual participants with a BNT score (both Spanish and English) of 50 (group *MDN*) or above were considered *high-proficient–balanced bilinguals*. Those bilinguals in the balanced group with a score in both versions of the BNT below the group median were considered *low-proficient–balanced bilinguals*. The unbalanced bilingual group was further subdivided into *English-dominant* and *Spanish-dominant* depending on their dominant language performance on the BNT. English dominant unbalanced bilinguals had a higher BNT–English score (at least 8 points) when compared to the BNT–Spanish score. The reverse was true for the Spanish-dominant unbalanced bilinguals. In consequence, four different groups were distinguished: *high-proficient–balanced bilinguals*, *low-proficient–balanced bilinguals*, *unbalanced English-dominant bilinguals*, and *unbalanced Spanish-dominant bilinguals*. Both variables (balance and proficiency) were separately analyzed. An overall significant correlation was observed between self-reported language proficiency (using the bilingualism questionnaire) and the Boston Naming Test scores ($r = .56$, $p < .01$).

In the bilingual group, the order of administration of English and Spanish tests (BNT and Stroop test) was randomized to avoid language order effects. The order of administration of each component within the Stroop Test was always the same (*reading of words*, *naming colors*, *color–word*, and for bilinguals, the *color–word between-language condition*).

Statistical Analyses

Three one-way analyses of variance were done to compare the scores in the three ST conditions for the following groups: (1) Spanish–English bilinguals with the English and Spanish monolinguals; (2) balanced bilinguals to the unbalanced English- and Spanish-dominant bilingual groups; and (3) high-proficient balanced bilinguals, low-proficient balanced bilinguals, unbalanced English-dominant bilinguals, and unbalanced Spanish-dominant bilinguals. The scores from the later four groups were also compared in the inter-language (test language and test response in different languages) condition. Bonferroni *post-hoc* tests corrections were used. To control for multiple comparisons and Type I error, 2×3 repeated measures procedures were used to analyze the effects of the test language and its interactions with the group variables over all four conditions of the ST for the bilingual groups. Language (Spanish or English) was the within-subjects factor, and type of bilingualism (balanced bilinguals, unbalanced bilinguals–English dominant and unbalanced bilinguals–Spanish dominant) or proficiency (high-proficient balanced bilinguals, low-proficient balanced bilinguals, unbalanced bilinguals–English dominant, and unbalanced bilinguals Spanish–dominant) was the between-subjects factor. Finally, regression analyses were performed to find the predictive value of age of acquisition of L2, and participant's age on the time in the color–word ST condition.

RESULTS

Comparisons of the Spanish–English bilinguals with both monolingual groups in the three conditions of the ST showed no significant differences except for the English color naming condition in which the bilinguals' performance was significantly slower than the English monolinguals (Table 2).

Balanced bilinguals were compared with unbalanced bilinguals in the different conditions of the ST using a one-way ANOVA. In addition, a 2×3 repeated measures ANOVA was used to analyze the effects of the test language and its interactions with the group variable over all conditions of the ST. Language (Spanish or English) was the within-subjects factor, and type of bilingualism (balanced bilinguals, unbalanced bilinguals–English dominant and unbalanced bilinguals–Spanish dominant) was the between-subjects factor (Table 3). Only one statistically significant difference was found among the bilingual groups. The unbalanced Spanish-dominant bilinguals were significantly slower than the unbalanced English-dominant bilinguals (*M* difference $= 14.58$, $p = .027$) and the balanced bilinguals (*M* difference $= 12.21$, $p = .024$) in the English color-naming condition of the ST. Group effect interacted with the language in which the test was presented: The unbalanced Spanish-dominant bilinguals were slower in the English conditions of the test but not in the Spanish conditions. The only ST condition that was significantly affected by the test language was the between-language color–word condition. More interference was observed when the words were written in English and named in Spanish than when they were written in Spanish but named in English. This language effect was moderated by the group. Only balanced bilinguals (*M* difference $= 11.04$, $p = .0250$) and unbalanced bilinguals English dominant (*M* difference $= 16.63$, $p = .002$) presented these interference effects. In addition, there was a significant Group $\times$ Language interaction for the ST reading and color–word conditions. Unbalanced Spanish bilinguals tended to be slower in the English versions of the ST test.

The effect of language proficiency on the time in seconds for the four conditions of the ST is presented in Table 4. The proficiency variable had a significant effect in the English naming condition only. This group effect was moderated by the test language. The unbalanced bilingual Spanish-dominant group spent more time than the other groups in the ST-naming color condition in English only. No significant differences were found between the high-proficient, low-proficient, and unbalanced bilinguals in the other conditions of the Stroop Test. However there was a significant

Table 2. Comparison of Spanish–English bilinguals, English and Spanish monolinguals in the three conditions of the Stroop Test

	Spanish–English bilinguals (n = 71)		English monolinguals (n = 40)		Spanish monolinguals (n = 11)			
Stimulus	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*F*	*p*
Spanish Stroop								
Reading								
Time (sec.)	46.89	(10.01)	—		45.73	(5.39)	.069	.933
Errors	0.09	(.42)	—		.00	(.00)	.297	.744
Naming								
Time (sec.)	68.76	(16.14)	—		63.56	(12.18)	.519	.597
Errors	0.28	(.72)	—		0.09	(.26)	.367	.694
Color–word								
Time (sec.)	112.85	(30.18)	—		97.91	(27.44)	1.177	.313
Errors	0.76	(1.20)	—		0.64	(.93)	.053	.948
English Stroop								
Reading								
Time (sec.)	47.20	(14.34)	43.68	(8.59)	—		.992	.374
Errors	0.42	(.20)	0.15	(.700)	—		.730	.484
Naming								
Time (sec.)	72.07	(17.94)	61.98	(12.53)	—		4.914	.009*
Errors	0.23	(.54)	0.20	(.56)	—		.027	.973
Color–word								
Time (sec.)	114.24	(32.22)	108.40	(30.17)	—		.436	.648
Errors	0.59	(1.69)	0.68	(1.07)	—		.039	.961

*Significant, $p < .01$.

language effect in the between-language color–word condition and a trend, but no significant interaction, with proficiency. Most bilingual groups were slower in the between language condition in which the letters were in English and the reading in Spanish. The interaction of this language effect with proficiency was not significant. In the color–word (intralanguage) condition, there was a significant interaction between Language × Proficiency Group, without significant main effects. Low-proficient balanced bilinguals and unbalanced Spanish-dominant bilinguals were slower than the other groups in the English color–word. The unbalanced English-dominant bilinguals, on the other hand, were slower in the Spanish color–word.

Regression analyses were performed to find the predictive value of age of acquisition of L2 and participant's age on the color–word ST condition. Only time in seconds was used as the dependent measure. Age of acquisition of L2 did not predict the ST color–word interference scores either in English (coefficient $B = .16$, $t = 046$, $p < .650$) or in Spanish (coefficient $B = .22$, $t = .59$, $p < .550$) The participant's age, however, was significantly related to scores on the Spanish Stroop Color naming ($r = .25$, $p < .005$) and color–word scores ($r = .26$, $p < .005$). Age was related to the English Stroop color-naming ($r = .39$, $p < .001$), reading ($r = .30$, $p < .001$), and color–word ($r = .53$, $p < .001$) scores.

DISCUSSION

Bilinguals' overall performance was slower on the ST when compared to monolinguals. Significant differences, however, were found only in the color naming condition. Bilingualism significantly affected word-retrieval ability (naming condition) but did not have a statistically significant effect on the response interference condition (color–word). Performance was only about 5% to 10% slower in bilinguals than in monolinguals in the color–word condition. Greater interference for bilinguals than for monolinguals has been previously reported particularly when the interfering and naming languages differed (Dyer, 1971). In our study, the comparisons between bilinguals and monolinguals were done in the within language condition only (naming language and the language of the color–word were the same). In the within language condition the level of interference between these two groups was very similar. One possible reason for the lack of significant differences between our bilingual and monolingual groups is the small size of the Spanish monolingual sample. If the Spanish monolingual sample were larger, statistically significant differences might have been found eventually. However, the size of the monolingual sample in Dyer's study was also small ($n = 12$).

Another potential confounding variable when comparing the bilinguals' and monolinguals' ST performance is that

Table 3. Comparison of balanced bilinguals, unbalanced bilinguals, English- and Spanish-dominant in three conditions of the Stroop test

	Balanced bilinguals ($n = 25$)		Unbalanced bilinguals, English dominant ($n = 15$)		Unbalanced bilinguals, Spanish dominant ($n = 31$)			
Stimulus	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*F*	*p*
Spanish								
Reading								
Time	46.04	(10.15)	46.53	(9.03)	47.74	(10.57)	.207	.813
Errors	0.24	(0.66)	0.00	(0.00)	0.01	(0.17)	2.31	.107
Naming								
Time	68.84	(17.11)	72.47	(15.42)	66.90	(15.89)	.594	.555
Errors	0.32	(0.69)	0.40	(1.06)	0.19	(0.54)	.462	.632
Color-Word								
Time	105.60	(25.72)	120.00	(41.58)	115.23	(26.72)	1.247	.294
Errors	0.88	(1.17)	0.93	(1.58)	0.58	(1.02)	.620	.541
BL color–word*								
Time	113.04	(32.16)	98.72	(13.34)	101.23	(26.92)	1.552	.220
Errors	0.41	(0.77)	0.45	(0.68)	0.26	(0.60)	.408	.667
English								
Reading								
Time	44.96	(9.01)	42.27	(7.53)	51.39	(18.82)	2.631	.079
Errors	0.01	(0.28)	0.01	(0.25)	0.00	(0.00)	1.225	.300
Naming								
Time	67.24	(14.07)	64.87	(12.65)	79.45	(20.41)	5.324	.007†
Errors	0.20	(0.50)	0.01	(0.26)	0.32	(0.65)	1.185	.312
Color–word								
Time	109.96	(31.96)	103.27	(30.10)	123.00	(32.05)	2.321	.106
Errors	0.88	(.02)	0.33	(0.62)	0.48	(0.96)	.598	.553
BL color–word**								
Time	102.00	(35.37)	82.09	(22.93)	106.46	(29.49)	2.442	.096
Errors	0.37	(0.37)	0.27	(0.90)	0.42	(0.80)	.115	.895

*Between-language condition: Stimuli were in English and the names of the color–words were reported in Spanish.
**Between-language condition: Stimuli were in Spanish and the names of the color–words were reported in English.
†Significant, $p < .05$.
Note. Language effects (vertical analyses). reading: ($F = 0.187$, $p = .667$); naming: ($F = 2.55$, $p = .615$); color–word: ($F = .267$, $p = .607$); BL color–word: ($F = 4.37$, $p = .041$).
Interaction effects (Group × Language). reading: ($F = 3.16$, $p = .048$); naming: ($F = 7.95$, $p = .01$); color–word: ($F = 5.61$, $p = .06$); BL color–word: ($F = 3.77$, $p = .029$).

the bilingual individuals were administered the ST twice (once in English and another time in Spanish), thus providing the bilingual group with additional practice on the test. This may affect their scores relative to the monolingual groups, which received only one administration of the ST. Familiarity and practice has shown to reduce the magnitude of the C–W interference but does not alter patterns of interference (Dyer, 1971). Furthermore, having to switch languages from one administration to the next may negatively impact performance. According to Grosjean (1998) a bilingual's state of activation with respect to his or her two languages can influence test performance. This state of activation is controlled by factors such as who the listener is. When the listener is bilingual, both languages are activated and mixing of the two languages will take place. As a consequence, the scores of our bilingual sample may be influenced by the language state of activation, and it can be conjectured that if the ST were administered in different languages, though on different days, performance may have been different.

Our results demonstrated that, depending on the test language, language proficiency among bilinguals influenced the ST performance. Participants that were Spanish-dominant bilinguals were significantly slower in all the ST conditions in English, and the English-dominant bilinguals were slower in the ST Spanish conditions. It has been reported that unbalanced (i.e., individuals mastering both languages at different levels) Spanish–English bilinguals retain fewer words in their L2 when compared to balanced bilinguals and to English monolinguals. These differences disappear when

Table 4. Means and standard deviations of the Stroop task (in seconds) performance in high-proficient–balanced, low-proficient–balanced, and unbalanced bilingual groups

	High-proficient–balanced bilinguals (n = 10)		Low-proficient–balanced bilinguals (n = 15)		Unbalanced bilinguals–English dominant (n = 15)		Unbalanced bilinguals–Spanish dominant (n = 31)			
Stimulus	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*F*	*p*
Age exposure to L2	8.90	(6.54)	8.93	(4.08)	6.60	(9.57)	15.74	(12.53)	3.66	.017†
BN Spanish	51.50	(3.17)	43.93	(5.75)	35.87	(5.90)	53.77	(3.07)	34.62	.000†
BN English	53.10	(2.60)	43.13	(5.82)	51.07	(2.89)	38.48	(5.71)	60.71	.000†
Spanish										
Reading	46.90	(8.16)	45.47	(11.53)	46.53	(9.04)	47.74	(10.57)	0.17	.912
Naming	66.40	(17.70)	70.47	(17.13)	72.47	(15.42)	66.90	(15.89)	.516	.672
Color–word	101.70	(26.06)	108.20	(26.07)	120.0	(41.58)	115.23	(26.72)	.456	.714
BL color–word* time	112.60	(37.17)	113.35	(29.56)	98.72	(13.34)	101.23	(26.92)	1.02	.391
English										
Reading	45.40	(9.38)	44.66	(9.08)	42.276	(7.53)	51.39	(18.82)	1.73	.168
Naming	64.50	(13.59)	69.07	(14.54)	64.87	(12.65)	79.45	(20.41)	3.66	.017†
Color–word time	100.40	(31.02)	116.33	(32.00)	103.27	(30.10)	123.00	(32.05)	2.07	.112
BL color–word** time	103.70	(34.69)	100.78	(37.10)	82.09	(22.93)	106.46	(29.49)	1.62	0.195

*Between-language condition: Stimuli were in English and the name of the color–words were reported in Spanish.
**Between-language condition: Stimuli were in Spanish and the names of the color–words were reported in English.
†Significant $p < .05$.
Note. Language effects (vertical analyses). Reading: ($F = .28$, $p = .596$); naming: ($F = 0.03$, $p = .859$); color–word: ($F = .03$, $p = .863$); BL color–word: ($F = 5.25$, $p = .026$).
Interaction effects (Group × Language). Reading: ($F = 2.09$, $p = .11$); naming: ($F = 5.22$, $p = .03$); color–word: ($F = 4.05$, $p = .01$); BL color–word: ($F = 2.51$, $p = .067$).

the bilinguals are tested in their dominant language (Harris, 1995). The majority (85.9%) of our bilingual subjects spoke Spanish as the L1, and since they were not exposed to L2 until a mean age of 11.0 years, L1 was presumably more established than L2. La Heu et al. (1996) found that unbalanced Dutch–English bilinguals spend more time translating ST color words than reading them in either language. So, it might be possible that our bilinguals used translation in the color naming condition increasing the time of performance.

The balanced bilinguals in our study performed similarly to the unbalanced English-dominant bilinguals in most of the English ST conditions and to the unbalanced Spanish-dominant group in the Spanish ST. When testing balanced bilinguals in this within language ST condition, the language of the test does not make a difference. Equal performance in balanced Chinese–English bilinguals (Lee & Chan, 2000) and balanced Swedish–German bilinguals (Magiste, 1984) has been reported elsewhere.

Bilinguals who have mastered their languages to different degrees (dominant in English or in Spanish) had more interaction in within-language (intralanguage) conditions than in the between-language (interlanguage) conditions. Our results confirm previous studies with Spanish–English (Dyer, 1971) and English–Turkish bilinguals (Kijak, 1982). However, an important point to stress is that this ratio of interference (more within- than between-language) did not apply to the balanced bilingual group. In the balanced bilinguals the time difference between the two conditions was either minimal or higher in the between language condition in which the inhibiting language was English. Our results support recent studies in which the between as well as the within language interference is determined by the bilingual's language proficiency (MacLeod, 1991; Magiste, 1984). Moreover, this pattern of ST interference can change with experience.

Magiste (1984; 1985) observed a changing pattern in the ST interference with the development of the individuals' second language. Initially, her subjects were German dominant and showed more interference when responding in German regardless of whether the words were Swedish or German. Gradually as their experience with Swedish increased, they reached a point of equivalence; then, as Swedish increased, they showed more interference when naming in Swedish. These results suggest that the interference between two languages follow a dynamic process in accordance with language experience of L2.

Cognitive psychologists have proposed two models to understand the lexical access and store of words in bilingual participants (Kroll & Stewart, 1994). The word association model suggests that bilinguals' access to words is through translation and, therefore, a clear link between the words of L1 and L2 is proposed. The second model is the concept mediation model, which assumes lexical independence of L1 and L2. If the former is true, more interference is expected in the between language ST conditions, while more interference in the within-language ST conditions would support the concept mediation model. Our results showed that the type of interference varies among bilingual groups. The unbalanced bilingual groups presented more interference in the within- than in the between-language conditions, and the balanced bilingual group presented interference levels that were similarly small in both the within- and between-language conditions. Our results do not support the concept mediation model or the association model for the access of lexical memory in bilinguals but suggest more independence of the lexical dictionaries in the unbalanced groups (Brauer, 1998).

Low-proficient balanced bilinguals deserve some special consideration. They were bilinguals with less than an 8-point difference between the Spanish and English BNT scores but a mean BNT (Spanish and English combined) score below the group median of 50 points. According to the self-report of the participants in this group they considered themselves proficient in both languages. But according to both the English and Spanish BNT norms (Kaplan et al., 1996; Mitrushina et al., 1999) their scores were low in both languages. It is noteworthy that these bilingual subjects reported active use of both languages in everyday life. People displaying this type of bilingualism have been called semilinguals by Paradis (1998) and they are quite prevalent in a bicultural society such South Florida's, in which two languages may be required for everyday communication. These subjects have a low level of language proficiency in both languages ability in either language if (and only if) monolingual norms are used. However a bilingual can be considered not just as the speaker of two different languages, but also as the speaker of one extended language (Grosjean, 1989). It is obviously unfair and conceptually inappropriate to use monolingual norms to evaluate language ability in this type of bilingual. Further study of neuropsychological differences of this group is required.

Although in our study the compared groups did not differ significantly in age, age was a predictor of the color–word subtest among all bilinguals. Ivnik et al. (1996) found a strong influence of age on all the Golden ST subtests. There is consistent evidence of a significant decline in performance with advancing age. Few studies have examined changes in the Stroop interference due to normal aging. In a developmental study, Cornalli et al. (1962) administered the Stroop Test to 235 subjects ages 7 to 80. Older adults and children had longer response latencies than young adults and middle-aged individuals in the color–word subtest but were nearly as fast when reading or naming the colors. Panek et al. (1984) administered the Stroop to 50 young adults (18–23) and 31 healthy older adults (61–85). Older adults were slower in all three conditions of the Stroop test and were particularly slower in the color–word subtest. The interaction between advanced age and bilingualism requires further investigation.

Our results did not support the association between age of acquisition of L2 and ST scores. There is very little research about the influence of age of learning L2 in neuropsychological test performance. Harris (1995) found that unbalanced Spanish–English bilinguals performed worst in the California Verbal Learning Test than balanced bilinguals. Most of the bilinguals in the unbalanced group had learned English during adulthood. Rosselli (2000) found that elderly Spanish–English bilinguals that had learned English (L2) during early adulthood and who maintained use of both languages did not suffer a linguistic decline in either language.

The present findings suggest using the within-language ST in L1 when testing bilinguals and that using the ST in L2 should affect color naming time by about 10 to 15% and by about 5 to 10% in the color–word condition. As a consequence, a minor score correction is suggested. The present study does not support drawing conclusions about the color reading score. Differences in time in color reading between Spanish and English were small, and for most of our subjects Spanish was L1. All our bilingual subjects had attended at least some school in English, and reading ability was presumably variable in both Spanish and English. Additionally, we did not test reading speed in either language.

Nonetheless, it seems that the bilingual population is quite heterogeneous, and our findings may be specific to bilingual subgroups: Spanish–English bilinguals. We demonstrated that language proficiency or mastery of one or two languages might influence the ST performance. In balanced bilinguals it does not matter what language is used in the ST, but in unbalanced subjects, the best-spoken language may be preferred. In addition, our results support the presence of both, between- and within-language interference in Spanish–English bilinguals.

ACKNOWLEDGMENTS

Our sincere gratitude goes out to Dr. Erika Hoff for her valuable suggestions and editorial support.

REFERENCES

Altarriba, J. & Mathis, K.M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory and Language*, *36*, 550–568.

Biederman, I. & Tsao, Y.C. (1981). On processing Chinese ideographs and English words: Some implications from the Stroop-test. *Cognitive Psychology*, *11*, 125–132

Brauer, M. (1998). Stroop interference in bilinguals: The role of similarity between the two languages. In A.F. Healy & L.E. Baurne, Jr. (Eds.). *Foreign language learning* (pp. 317–337). Mahwah, NJ: Erlbaum.

Chen, H.C. & Juola, J.F. (1986). Dimensions of lexical coding in Chinese and English. *Memory and Cognition*, *10*, 216–224.

Cornalli, P., Wapner, S., & Werner, H. (1962). Interference effects of the Stroop Color–Word test in childhood, adulthood and aging. *Journal of Genetic Psychology*, *100*, 47–53.

Dyer, F.N. (1971). Color-naming interference in monolinguals and bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *10*, 297–302.

Fang, S.P., Tzeng, O.J.L., & Alva, L. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. *Memory and Cognition*, *9*, 609–617.

Gerhand, S.J., Deregowski, J.B., & McAllister, H. (1995). Stroop phenomenon as a measure of cognitive functioning of bilingual (Gaelic–English) subjects. *British Journal of Psychology*, *86*, 89–92

Glasser, W. & Glasser, M. (1989). Context effects in the Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, *118*, 13–42.

Golden, J.C. (1978). *Stroop Color and Word Test*. Chicago: Stoelting Co.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*, 131–149.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*, 3–15.

Harris, J.G., Cullum, C.M., & Puente, A.E. (1995). Effects of bilingualism on verbal learning and memory in Hispanic adults. *Journal of the International Neuropsychological Society*, *1*, 10–16.

Heij, W.L., de Bruyn, E., Elens, E., Hartsuiker, R., Helaha, D., & van Schelven, L. (1990). Orthographic facilitation and categorical interference in a word-translation variant of the Stroop task. *Journal of Canadian Psychology*, *44*, 76–83.

Heij, W.L., Hooglander, A., Kerling, R., & van der Velden, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory and Language*, *35*, 648–665

Henderson, L. (Ed.). (1984). *Orthographies and reading: Perspectives from cognitive psychology, neuropsychology and linguistics*. Hillsdales, NJ: Erlbaum.

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., & Petersen, R.C. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE, Token, WRAT–R Reading, AMNART, STROOP, TMT, and JLO. *Clinical Neuropsychologist*, *10*, 262–278.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test* (2nd ed.). Boston: Lea & Febiger.

Kaplan, E., Goodglass, H., & Weintraub, S. (1996). *Test de Vocabulario de Boston* [Boston Vocabulary Test]. In H. Goodglass & E. Kaplan, *Evaluación de las afasias y de trastornos relacionados* (pp. 185–190). Madrid: Editorial Médica Panamericana.

Kijak, H.A. (1982). Interlingual interference in naming color words. *Journal of Cross-Cultural Psychology*, *13*, 125–135.

Kroll, J. & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.

La Heu, W., Hooglander, A., Kerling, R., & vander Velden, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory and Language*, *35*, 648–665.

Lee, T.M.C. & Chan, C.C.H. (2000). Stroop interference in Chinese and English. *Journal of Clinical and Experimental Neuropsychology*, *22*, 465–471.

Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

MacLeod, C.M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.

Magiste, E. (1984). Stroop tasks and dichotic translation: The development of interference patterns in bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *10*, 304–315.

Magiste, E. (1985). Development of intra- and inter lingual interference in bilinguals. *Journal of Psycholinguistic Research*, *14*, 137–154.

Mitrushina, M.N., Boone, K.B., & D'Elia, L.F. (1999). *Handbook of normative data for neuropsychological assessment*. New York: Oxford University Press.

Paradis, M. (1998). Aphasia in bilinguals: How atypical is it? In P. Coppens, Y. Lebrun, & A. Basso (Eds.), *Aphasia in atypical populations* (pp. 35–66). Mahwah, NJ: Lawrence Erlbaum Associates.

Panek, P., Rush, M., & Slade, L. (1984). Locus of the age–Stroop interference relationship. *Journal of Genetic Psychology*, *145*, 209–216.

Rey, G. (2000). *Spanish Stroop Test*. Unpublished manuscript.

Rosselli, M., Ardila, A., Araujo, K., Weekes, V.A., Caracciolo, V., Pradilla, M., & Ostrosky, F. (2000). Verbal fluency and repetition skills in healthy older Spanish–English bilinguals. *Applied Neuropsychology*, *7*, 17–24.

Seinderberg, M.S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19*, 1–30.

Spreen, O. & Strauss E. (1998). *A compendium of neuropsychological tests* (2nd ed.). New York: Oxford University Press.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 760–771.