Bilingualism: Language and Cognition 15 (3), 2012, 663–676 © Cambridge University Press 2012 doi:10.1017/S1366728912000077

The effects of first- and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task*

CHI-SHING TSE
The Chinese University of Hong Kong
JEANETTE ALTARRIBA
University at Albany, State University of New York

(Received: May 25, 2011; final revision received: October 1, 2011; accepted: January 12, 2012; first published online 20 March 2012)

By administering a Stroop task to college-student bilinguals varied in self-rated first- (L1) and second-language (L2) proficiency, the current study examined the effects of L1 and L2 proficiencies on selective attention performance. We conducted ex-Gaussian analyses to capture the modal and positive-tail components of participants' reaction time distributions. Both L1 and L2 proficiencies were associated with a shift of reaction time distributions in incongruent trials, relative to congruent trials, and the tail size of reaction time distributions regardless of trial types. This suggests that bilinguals' L1 and L2 proficiencies could affect their Stroop performance via modulating their conflict resolution and goal maintenance abilities.

Keywords: attentional control, bilingualism, conflict resolution, goal maintenance, response time distribution

While abilities to speak, write, read, and comprehend two languages are clearly language skills, bilinguals have been reported to outperform monolinguals in tasks demanding attentional control (e.g., Bialystok, 2010; Costa, Hernández & Sebastián-Gallés, 2008; see Bialystok, 2009, for a review). This bilingual advantage could be explained as follows: when processing languages, bilinguals' first and second language (L1 and L2) systems create a conflict for selection, so they need to continuously monitor attentional resources to the target language (GOAL MAINTENANCE) and inhibit unwanted language to avoid confusion in language processing (CONFLICT RESOLUTION) (e.g., Abutalebi & Green, 2007; Bialystok, 2009; Costa, La Heij & Navarrete, 2006). The experiences with the ongoing demand of coordinating two languages lead bilinguals to practice more on coordinating attentional resources, thereby enhancing their performance in selective attention tasks.

* The work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project no. CUHK455010). We thank David Green and Gigi Luk for their comments on earlier versions of this manuscript and Rita German, Jacalynn Romeyn, and Dana Schiffman for their help with data collection.

Bilinguals' language selection processes are analogous to the processes involved in a selective attention task. We adapted Balota and Faust's (2001) attentional control framework (Figure 1; see also Kane & Engle, 2003, for a similar idea) that explains performance in selective attention tasks involving the processing of multiple stimuli in multiple dimensions. Consider, for example, an incongruent trial in the Stroop task, the prototypical measure of attentional control (MacLeod, 1992).

When individuals read aloud the name of the ink color of a color name (e.g., RED) that is printed in a nonmatching color (e.g., green), multiple pathways that represent ink color and color name are engaged and compete for output. Participants must maintain the task goals (i.e., responding to the ink color instead of the color name) and resolve the conflict by suppressing the interference from prepotent, task-irrelevant color-name pathways and accessing the subordinate, task-relevant ink-color pathway. No such suppression need occur for congruent trials where the color and the word match (RED in red). Hence, the effectiveness of attentional control depends on the ability to regulate relative strength of task-appropriate and task-inappropriate pathways (conflict resolution) and to maintain the representation of task demands during the task (goal maintenance). The

Address for correspondence:
Chi-Shing Tse, Department of Educational Psychology, The Chinese University of Hong Kong, New Territories, Hong Kong, China
cstse@cuhk.edu.hk

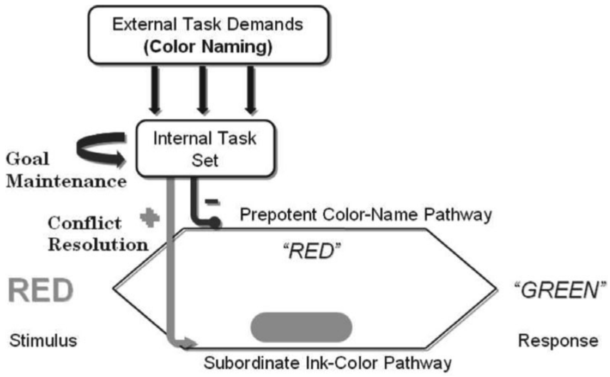

Figure 1. Attentional control framework (in the context of the Stroop task). Modified from Balota & Faust (2001, p. 71).

Stroop effect (i.e., reaction time (RT)/error difference in incongruent vs. congruent trials) reflects how well one can resolve the conflict between color and word pathways. A larger Stroop effect occurs when participants produce slower and/or less accurate naming responses for the color in the incongruent vs. congruent trials.

Past evidence reflects the roles of conflict resolution and goal maintenance in a bilingual advantage in selective attention tasks. For example, bilinguals show smaller Stroop effects than monolinguals, despite the absence of their difference in naming speed (Bialystok, Craik & Luk, 2008, see also Carlson & Meltzoff, 2008). Using the Attention Network Test (ANT), Costa et al. (2008) found that relative to monolingual young adults, their bilingual peers showed faster overall RTs, larger benefit from alerting cues, smaller conflict costs, and smaller switch costs. The bilinguals' superiority in resolving conflict among different sources of information may also enhance their memory discrimination. Bialystok and Feng (2009) found that bilinguals outperformed monolinguals in a release from proactive interference memory task after taking into account their difference in vocabulary size. Relative to monolingual children, bilingual children showed less build-up of proactive interference; that is, fewer intrusion errors from words in the same category appeared on previous lists. While these findings clearly indicate a bilingual superiority in conflict resolution, others have demonstrated the role of goal maintenance in bilinguals' superior performance. Using a Simon task, Martin-Rhee and Bialystok (2008) had children remember the color cue that corresponded to left/right response keys and respond to the identity of the color cue, while ignoring the location of the cue (left or right) that is irrelevant to the correct response. In congruent trials, the location and identity of the cue converged on the same response. In incongruent trials, they diverged on different responses. If only conflict resolution played a role in a bilingual advantage, bilinguals would have performed better than monolinguals only in incongruent trials, where a conflict occurred between two responses, but not in congruent trials, where such conflicts did not occur. Contrary to this, Martin-Rhee and Bialystok (2008) reported that bilinguals responded faster than monolinguals in BOTH incongruent and congruent trials that were intermixed together in the task. To perform this task, participants need to assess and adjust their behavior for each trial to make the proper response. Costa, Hernández, Costa-Faidella and Sebastián-Gallés (2009) reported a bilingual advantage (relative to monolinguals) in tasks requiring high monitoring (e.g., congruent and incongruent trials are in equal proportion) but not in tasks requiring low monitoring (e.g., most of the trials were either congruent or incongruent). Overall, these results suggest that relative to monolinguals, bilinguals are more flexible in monitoring and detecting potential conflicting information, so they respond faster than monolinguals in selective attention tasks, whether or not the trials involved conflict resolution.

Apart from the monolingual vs. bilingual comparison, other studies have examined the relationship between L2 proficiency and attention control within bilinguals. For instance, Carlson and Meltzoff (2008) had balanced and unbalanced bilingual children perform several selective attention tasks. After controlling for their age, parents' education, and group differences in overall RTs, balanced bilinguals did better than unbalanced bilinguals in these tasks, showing their superior conflict resolution abilities. Luo, Luk and Bialystok (2010) found that bilinguals with more vocabulary knowledge showed better performance in letter fluency than those with less vocabulary knowledge. Tao, Marzecova, Taft, Asanowicz and Wodniecka (2011) tested bilingual college-age participants in the ANT and found that those with more balanced L1 and L2 proficiency and usage demonstrated better conflict resolution abilities. Luk, de Sa and Bialystok (2011) reported that a later onset age of active bilingualism was associated with a smaller L2 vocabulary size and lower conflict resolution abilities, as indicated by a stronger interference effect in a flanker task. Similarly, by investigating bilingual children in an L2 immersion program, Bialystok and Barac (2012) observed that the interference effect in a flanker task was stronger when they had less balanced L1/L2 proficiencies and/or received fewer years of bilingual education. These findings showed that the benefit of bilinguals' experience of coordinating two language systems in resolving the conflict among different sources of information depends on their L1/L2 proficiency, as quantified by various indices.

Despite the above-mentioned evidence for bilinguals' superiority in conflict resolution and goal maintenance, several questions have been left unanswered. First, it is unclear to what extent bilinguals' L1 and L2 proficiencies could separately modulate selective attention

https://doi.org/10.1017/S1366728912000077 Published online by Cambridge University Press

performance. Most of the studies (see Carlson & Meltzoff, 2008; Luk et al., 2011, for exceptions) collapsed bilinguals differing in L1/L2 proficiency into one group, and did not, at least explicitly, control for L1 proficiency. Second, while previous experiments often reported the positive effect of bilingualism in children (e.g., Bialystok, 2010) and older adults (e.g., Bialystok, Craik and Ryan, 2006), few studies (e.g., Costa et al., 2008, 2009) have targeted young adults who typically have better attentional control than other age groups. Currently, we recruited young-adult bilinguals (age range: 17–26 years) with varying levels of L1/L2 proficiency to test the effects of L1/L2 proficiencies on selective attention in a Stroop color-naming task. A wide range of proficiencies allows us to examine the unique contributions of L1 and L2 abilities via fitting participants' performance in multiple regression models.

Apart from central tendency measures (e.g., means), we modeled the RT data at the distributional level. The analysis of mean RT *per se* may not be sufficient and can sometimes even be misleading, because failing to take the shape of the RT distribution into account may obscure more subtle aspects of performance (e.g., Balota & Yap, 2011; Heathcote, Popiel & Mewhort, 1991; Spieler, Balota & Faust, 1996; Tse, Balota, Yap, Duchek & McCabe, 2010). Analyses of RT distributional characteristics can be done by fitting individual raw RTs to a theoretical ex-Gaussian distribution that closely approximates the typical positively-skewed empirical distribution. This distribution is defined by a three-parameter function (see Figure 2): μ and σ reflect the mean and standard deviation of the Gaussian component, respectively, and τ reflects the exponential component of the RT distribution. The ex-Gaussian analyses can be used as a descriptive model for capturing the influence of a variable on RT distributions, with the parameters having a direct relation to the mean of a distribution. The algebraic sum of μ and τ is constrained to approximate closely the empirical distribution, so the difference in mean RT between two conditions can be partitioned to be distributional shifting and/or a change in the tail of the RT distribution. As is shown in Figure 3, a variable could shift the RT distribution, as reflected by an increase in μ (panels A and B), produce a larger proportion of slow RTs that flattens the tail of the RT distribution, as reflected by an increase in τ (panels A and C), or have no effect on overall mean RT, but have opposing effects on the components of the RT distributions (μ and τ, panels A and D) (see Heathcote et al.'s (1991) congruent vs. neutral trials, i.e., XXX printed in red, in a Stroop task, as an example). A longer RT due to a group difference or a manipulation could be attributed to a shift of the RT distribution, an increase in the tail size of the RT distribution, or both.

The μ and τ estimated on the basis of participants' RT distribution in the selective attention task have implications for the efficiency of their attentional control

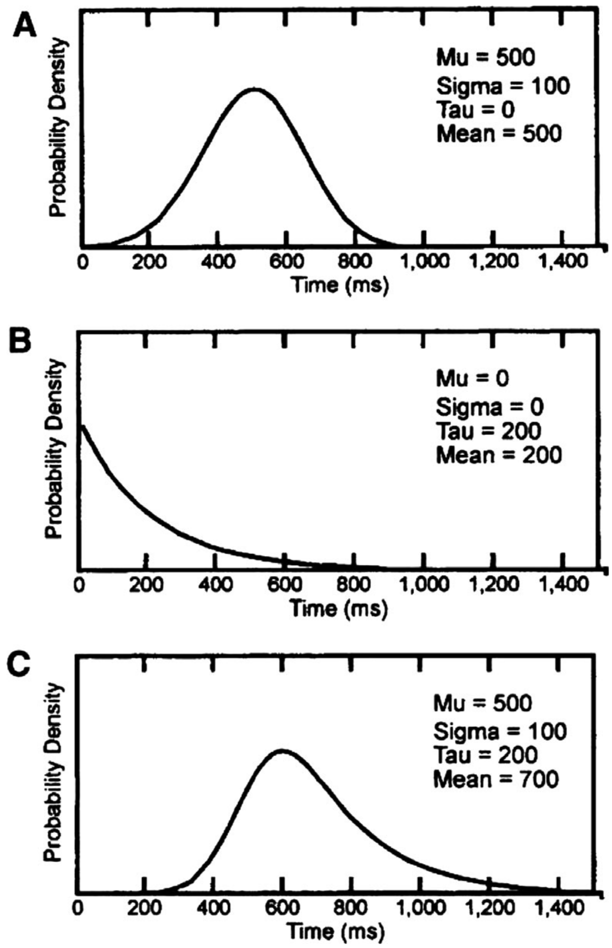

Figure 2. The Gaussian and exponential distributions and their convolution for an ex-Gaussian distribution in panels A, B, and C, respectively. Adapted from Figure 2 in Balota et al. (2008).

system. Those who have higher difficulty in suppressing irrelevant information and resolving response conflict are likely to produce slower responses. A shift of RT distribution (i.e., difference in μ) between two conditions may reflect impairment in conflict resolution. For instance, in Spieler, Balota and Faust's (2000) global-local task, participants attend to a spatial dimension of a stimulus in the face of a congruent or incongruent dimension that is scaled at a different size (e.g., "an H made up of smaller Es" in an incongruent trial vs. "an H made up of smaller Hs" in a congruent trial). They found that the interference effect in RT (difference in incongruent vs. congruent trials) could be attributed to a larger μ (but not a change in σ or in τ) in incongruent vs. congruent trials. Incongruent trials require a certain degree of conflict resolution, as reflected by the fixed amount of RT (i.e., μ) increase for incongruent vs. congruent trials across the RT distribution. Similarly, a difference in μ, but not in σ or τ, was reported to contribute

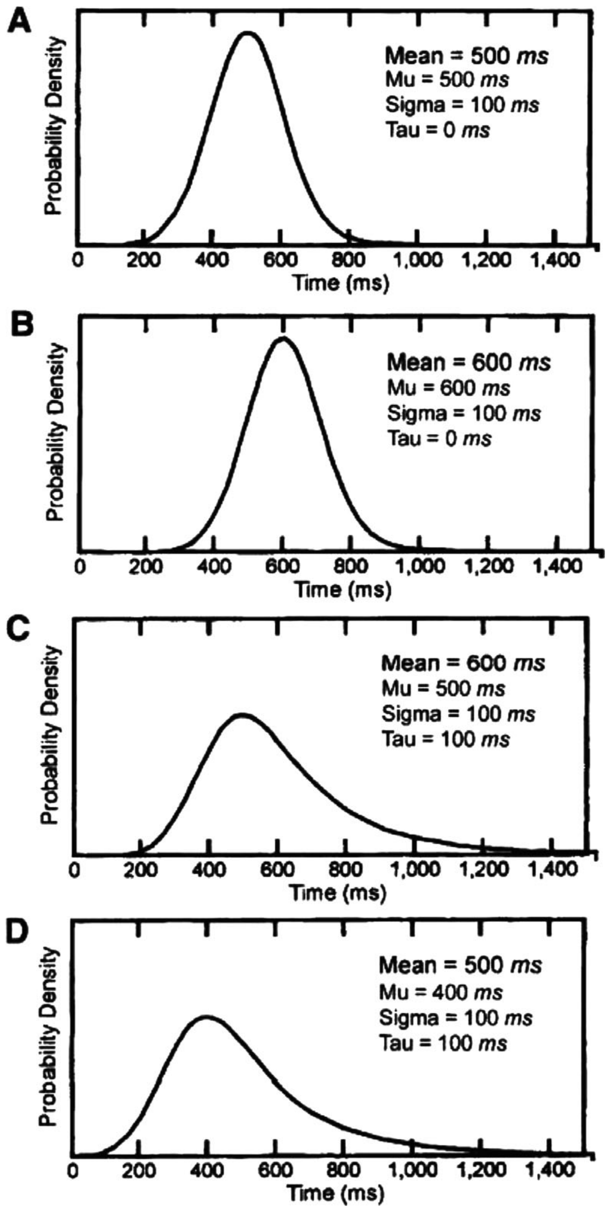

Figure 3. Possible changes in the RT distribution and their influences on mean RT and parameter estimates of ex-Gaussian analyses. Adapted from Figure 3 in Balota et al. (2008).

to the RT delay when participants named a picture while ignoring a simultaneous superimposed distracting word (e.g., Piai, Roelofs & Schriefers, 2011). Following these interpretations, we tested individual differences in conflict resolution abilities by comparing participants' μ difference in incongruent and congruent trials (i.e., Stroop effect in μ). In contrast, a lengthening of the tail of an RT distribution (i.e., a change in τ) is due to an increase in the proportion of slow RTs on the trials in which the task goal is momentarily lost but then corrected before an overt error is produced (see periodic goal neglect, de Jong, Berendsen & Cools, 1999; a fluctuation of goal maintenance, Kane & Engle, 2003). Individuals who fail to maintain the task goal may be more likely to experience losses of control over time and produce a larger tail of the RT distribution (i.e., larger τ). For example, in an antisaccade task, participants were first cued by a flashing cue appearing on the left or right of the screen and then instructed to shift their attention to the same side (prosaccade trial) or the opposite side (antisaccade trial) of the screen. Unsworth, Spillers, Brewer and McMillan (2011) found that antisaccade cost (i.e., RT difference between antisaccade and prosaccade trials) was due to changes in both μ and τ, so it was associated with participants' conflict resolution and goal maintenance abilities. In the current study, we tested individual differences in goal maintenance abilities by comparing participants' τ.

Previous studies showed that individual differences in attentional control abilities can be localized to specific ex-Gaussian parameters (e.g., Leth-Steensen, King Elbaz & Douglas, 2000; McAuley, Yap, Christ & White, 2006; West, Murphy, Armilio, Craik & Stuss, 2002). Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann (2007) estimated ex-Gaussian parameters for young adults' performance in eight choice RT tasks and showed that τ, but not μ or σ, was strongly related to working memory capacity, providing evidence for the relationship between τ and attentional control. Leth-Steensen et al. (2000) reported a disproportional increase in the tail of the RT distribution (i.e., τ) in a choice RT task for children with attention-deficit/hyperactivity disorder. Tse et al. (2010) showed that relative to μ and σ, τ in the Stroop task is a better discriminator for healthy older adults vs. those who are in the earliest stages of dementia, suggesting that the overall slower RT in older adults with early-stage dementia can be attributed to their failure in maintaining the task goal during the task.

Most of the studies (e.g., Spieler et al., 1996, 2000; Tse et al., 2010) that examined selective attention performance at the level of RT distribution did not test the effect of L1/L2 proficiencies and they did not even report if their participants knew more than one language. In the current study, we conducted RT distributional analyses to examine the effect of language proficiencies on bilinguals' selective attention performance. By estimating ex-Gaussian parameters for participants' RT distributions, we tested if bilinguals' L1 and L2 proficiencies could affect goal maintenance (as reflected by τ) and/or conflict resolution (as reflected by an increase of μ in incongruent vs. congruent trials) in the attentional control system. This would not be revealed in participants' mean RT data as a slower RT could be due to a change in μ (a shift of the RT distribution), in τ (an increase in the tail size of the RT distribution), or both. The Stroop

https://doi.org/10.1017/S1366728912000077 Published online by Cambridge University Press

task can test participants' conflict resolution and goal maintenance abilities (e.g., Kane & Engle, 2003). A shift due to the color vs. word congruency in the RT distribution (i.e., the Stroop effect in μ–μ difference in incongruent vs. congruent trials) can be attributed to the fact that incongruent trials demand conflict resolution in the form of the inhibition of a pre-potent word response and generation of a correct color response. This adds about a constant amount of time to incongruent trials, making them on average slower than congruent trials. The task also requires the participants to maintain the task goal throughout the task. Transient failures of goal maintenance lead to very slow correct RTs as indexed by a lengthening tail of the RT distribution. Thus, RT distributional analyses could shed light on the interaction between language and cognition via the study of bilingualism and attentional control.

The predictions of the findings were straightforward. After controlling for demographic variables like socioeconomic status, bilinguals with high L2 proficiencies would show faster overall RTs in congruent and incongruent trials and smaller Stroop effects than those with lower L2 proficiencies. If a bilingual advantage stems from the ability to manage attention between TWO language systems, we expected that L1 proficiency would show similar effects as would L2 proficiency. The distributional analyses would show if L1/L2 proficiency modulated conflict resolution and/or goal maintenance. Given that τ reflects how well bilinguals stay attuned to the task, if the effects of language proficiencies occur in goal maintenance, τ would be negatively correlated with L1/L2 proficiency. Moreover, as participants' difficulty with maintaining their task goals was likely at the list-wide level, rather than in a specific condition, we expected that the τ vs. L1/L2 proficiency correlation would occur in both congruent and incongruent trials. Finally, given that the change in μ in incongruent vs. congruent trials reflects the degree to which bilinguals can resolve a conflict, if L1/L2 proficiency has an effect on their conflict resolution abilities, the Stroop effect in μ would also be negatively correlated with their L1/L2 proficiency.

Method

Participants

One hundred and ten bilingual undergraduates, with English as their L2, participated in exchange for partial fulfillment of a course requirement. Due to high error rates (>30%) and based on multivariate outlier analyses via Mahalanobis distances, data from six and four participants, respectively, were discarded. Hence, the following analyses were based on the remaining 100 participants' data. Informed consents were obtained from participants at the start of the study.

Materials and procedures

PC-compatible computers with E-Prime were used to display stimuli and collect RT and error data. Participants were individually tested in a quiet cubicle. All stimuli were presented on a black background, one at a time, at the center of the screen. The task involved four color names in English (red, blue, green, and yellow) and participants were instructed to respond in their L2 – English. There were three blocks. The first two were "baseline" blocks in which participants read aloud each of the four color-name words printed in black six times in one block and each of four ink colors appearing as color patches six times in the other block. Participants' RTs in these two blocks were treated as covariates in multiple regression analyses to control for bilinguals' word production abilities that could affect their verbal responses in the Stroop task. The remaining one was the color naming block, in which each of the four color names appeared 18 times in its corresponding color (72 congruent trials) and each of the four color names appeared six times in each of the three nonmatching colors (72 incongruent trials). To increase the sensitivity of the task in detecting the role of monitoring (see Costa et al., 2009), congruent and incongruent trials were in equal proportion and randomly intermixed. Participants read aloud the ink color. Prior to actual trials, participants were given eight practice trials. Each trial began with a fixation "+++" for 700 ms, followed by a 50 ms blank-screen inter-stimulus interval. The stimulus then appeared and remained until a response. The program recorded participants' RT (i.e., time interval between the onset of the presentation of the stimulus and their vocal responses as captured by the microphone) and the researcher coded correct responses, non-intrusion errors (e.g., stutter), or intrusions (e.g., naming the word). Due to sparse non-intrusion errors, we combined them with intrusion errors in the following analyses. The pattern of findings remained the same when only intrusion errors were considered.

After the Stroop task, we had participants separately rate, from 0%–100%, how closely their L1 and L2 proficiencies in four language skills (comprehension, speaking, reading, and writing) were relative to native speakers. By comparing language proficiency with those of native speakers we could define a common baseline for participants who varied in proficiency. The relationship between language proficiency measures was investigated using principal components analysis. The L1 and L2 proficiency scores were entered separately in two factor analyses with varimax rotation techniques with the eigenvalue limits being set over 1 but without specifying the number of factors to be extracted. In either analysis, there was only one significant component, which accounted for 67% and 81% of the variance in L1 and L2 ratings across four activities, with all activities

https://doi.org/10.1017/S1366728912000077 Published online by Cambridge University Press

being loaded positively (.71/.90, .66/.89, .93/.88, and .94/.92, for L1/L2 comprehension, speaking, reading, and writing, respectively). Thus, we computed factor scores to represent L1 and L2 proficiencies. (As the factor analyses with this set of parameters only yielded one factor for L1 (and L2) proficiency, the analyses that allow an oblique rotation (e.g., direct quartimin) yielded identical results. When we entered all L1 and L2 proficiency measures together, we obtained two factor scores, each of which represented the proficiency of one language. For simplicity, we only reported analyses based on factor scores yielded when L1 and L2 proficiency scores were analyzed separately.) The participants ranked the relative fluency for the languages that they have known. Those who ranked English as the first were regarded as English dominant ($N = 44$), whereas those who did not were regarded as English non-dominant ($N = 56$). They also provided their onset age of active bilingualism; that is, the age at which they considered that they had actively begun using their L2. We also sought participants' age, sex, handedness, and socioeconomic status and treated them as covariates in the following analyses. Following Morton and Harper's (2007) procedure, we computed socioeconomic status based on participants' parents' highest level of education and their total annual income. Each parent received a score (1–4) based on their level of academic achievement (junior high or less = 1; postgraduate = 4). Families received a score (1–5) based on their total income (e.g., <US$20,000 = 1; US$20,000–US$60,000 = 2; etc.). Parent education scores were averaged and combined with income scores to create a composite socioeconomic status score ranging from 2 to 9. After participants completed the questionnaire, we indexed their L2 proficiency by using their raw scores on a 40-item computerized vocabulary test (Cronbach's alpha = .87; Shipley, 1940). Despite its age, this test is still widely used (e.g., Yap, Tse & Balota, 2009). Due to the diversity of participants' L1 (Albanian, Cantonese, Creole, French, Fujian, German, Hebrew, Hindi, Italian, Japanese, Korean, Malay, Mandarin, Nepali, Polish, Portuguese, Romanian, Russian, Somali, Spanish, Turkish, Ukrainian, and Urdu), it is extremely difficult to measure their L1 proficiencies and compare them across different tests. All participants were debriefed at the end of the entire study.

Results

The level of significance was set at .05. Table 1 presents participants' demographic information and Shipley vocabulary raw scores. The self-rated L2 proficiencies were lower than the self-rated L1 proficiencies in all four language activities (all ts > 2.15, ps $< .05$). The standard deviations of these self-rated proficiencies indicated that participants' L1/L2 proficiencies were quite diverse. Table 2 presents cell means and regression results of participants' performance in the Stroop task. Prior to the analyses, we performed the following preliminary data treatment. We first excluded RTs for incorrect responses. For correct-response trials, we excluded those that were shorter than 200 ms and >3 *SD* above and below each participant's overall mean (~2.1%). The non-negative correlations between participants' overall RTs and errors (+.12, *ns*) showed that the interpretation of our findings was not clouded by a speed–accuracy tradeoff. The Stroop effects (RT/error differences in incongruent vs. congruent trials) were significant (105 ms/7.0%, both ts > 12.00, ps $< .001$). We examined individual raw RT distribution by estimating participants' ex-Gaussian parameters using a quantile maximum likelihood estimation procedure in QMPE 2.18. This procedure provides unbiased parameter estimates and is reported to be more effective than continuous maximum likelihood estimation for small samples (e.g., Cousineau, Brown & Heathcote, 2004). All fits successfully converged within 250 iterations. We yielded ex-Gaussian parameters for overall RTs and RTs of congruent and incongruent trials as there were sufficient numbers of observations for congruent (~70) and incongruent trials (~67). We did not estimate ex-Gaussian parameters for RTs in the baseline blocks as there were insufficient observations (~22), and these data were only used to control for participants' baseline naming speed. To check if ex-Gaussian parameters provide a good description for the RT data at the distributional level,

Table 1. *Mean statistics for participants' demographic information and Shipley vocabulary test scores.*

	Mean (*SD*)	
Age	19.48 (1.56)	
Proportion of sex (M:F)	28:72	
Proportion of handedness (left:right)	7:93	
Socioeconomic status	5.28 (1.87)	
Shipley Vocabulary Raw Scores	24.86 (4.63)	
	L1	L2
L1/L2 proficiency in speaking	85.96 (22.45)	70.96 (25.12)
L1/L2 proficiency in comprehension	92.31 (16.13)	75.26 (22.74)
L1/L2 proficiency in reading	78.46 (30.20)	69.51 (28.83)
L1/L2 proficiency in writing	73.81 (30.65)	61.70 (32.87)

Note: See Method section for the details of the scales and definitions of L1/L2 proficiencies in four language activities and socioeconomic status.

https://doi.org/10.1017/S1366728912000077 Published online by Cambridge University Press

Table 2. *Mean statistics and findings of the full model of multiple regression analyses for participants' performance in the Stroop task.*

		Mean (*SD*)	*F*(9,90)	*MSE*	R^2
	RT	703 (97)	30.60*	2564	.75
	Error	5.56 (4.17)	0.52	18.21	.05
Overall	μ	600 (84)	12.16*	3503	.55
	σ	123 (42)	1.74	1640	.15
	τ	103 (54)	4.16*	2294	.29
	RT	105* (47)	2.37*	1943	.19
	Error	7.01* (5.84)	1.27	33.35	.11
Stroop Effect	μ	115* (80)	2.20*	5780	.18
	σ	41* (44)	0.56	1998	.05
	τ	–10 (60)	1.99*	3284	.17
	RT	651 (88)	27.06*	2312	.73
	Error	2.05 (3.14)	1.08	9.77	.10
Congruent	μ	560 (72)	10.42*	2754	.51
	σ	98 (37)	1.29	1318	.11
	τ	91 (55)	3.64*	2401	.27
	RT	756 (111)	25.52*	3787	.72
	Error	9.06 (6.49)	0.68	43.33	.06
Incongruent	μ	675 (104)	13.49*	5076	.57
	σ	139 (51)	1.38	2532	.12
	τ	81 (58)	3.43*	2776	.26

* $p < .05$ (two-tailed)
Note: The mean RT/error in word- and color-naming baseline blocks were 498 ms/1.62% (*SD* = 68/2.90) and 578 ms/3.90% (*SD* = 82/6.81), respectively.

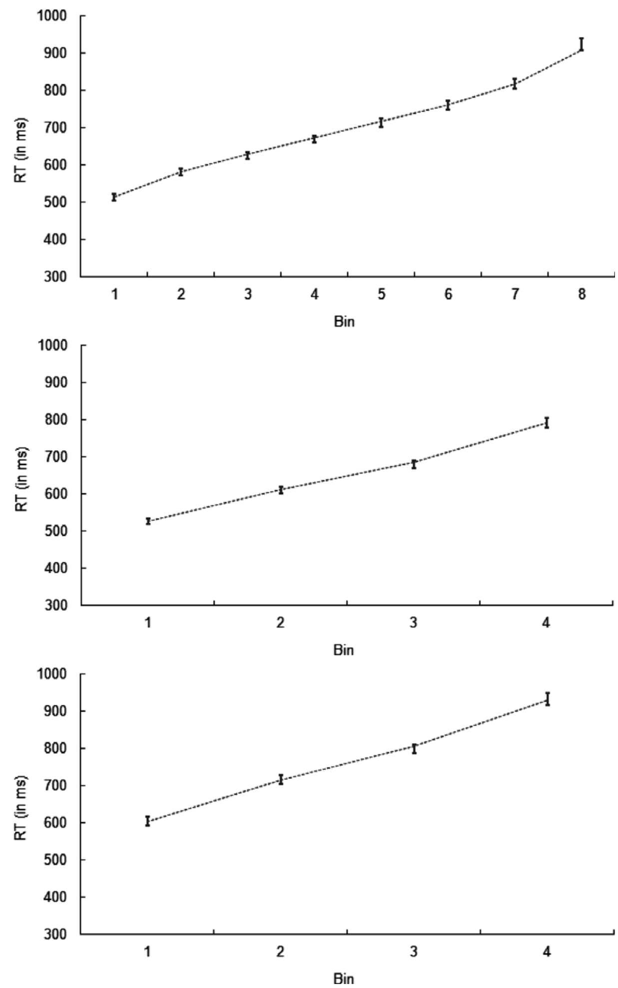

Figure 4. RT (in ms) as a function of vincentiles for participants (N = 100). Empirical vincentiles are represented by error bars while fitted ex-Gaussian vincentiles are represented by lines. The top panel depicts the vincentile plot for RTs in all trials. The middle panel depicts the vincentile plot for RTs in congruent trials. The bottom panel depicts the vincentile plot for RTs in incongruent trials.

we plotted mean vincentiles for the data to provide a graphic complement to the ex-Gaussian fits. Vincentizing averages RT distributions across participants (e.g., Balota, Yap, Cortese & Watson, 2008) to produce a typical participant's RT distribution (see Figure 4 – we used four bins, rather than eight bins in vincentile plots for congruent and incongruent trials due to the fewer observations in congruent and incongruent trials). The empirical vincentiles are represented by data points and standard error bars, and the vincentiles of the respective best-fitting ex-Gaussian distribution are represented by lines. The theoretical vincentiles were computed by line search on the numerical integral of the fitted ex-Gaussian distribution (see also Yap et al., 2009, for a similar data treatment). Each vincentile represents a different range of RTs for each participant. Presenting the data in this way allows one to visually assess the goodness of fit between empirical and theoretical vincentiles. As shown in Figure 4, RT data were clearly fitted very well by the ex-Gaussian distribution, and in most cases the divergence between the empirical vincentiles and theoretical ex-Gaussian vincentiles was smaller than one standard error. Hence, empirical RT distributions can be accurately captured by ex-Gaussian parameters for overall RTs, as well as for RTs of incongruent and congruent trials. The analyses of ex-Gaussian parameters showed that the Stroop effect in RT was due to a congruent vs. incongruent difference in μ (675 vs. 560, $t(99) = 14.39$, $d = 2.05$), but not in τ (81 vs. 91, $t(99) = 1.72$, $d = .24$) (see Table 2).

We performed multiple regression analyses to examine whether L1/L2 proficiency could predict participants' Stroop performance, after partialing out other extraneous variables. We entered participants' age, sex, handedness, and socioeconomic status in the first step, their baseline word- and color-naming RTs (for controlling participants' baseline naming speed) in the second step, the L1 and L2 proficiency factor scores in the third step, and their interaction term in the fourth step. These analyses allowed for an examination of the unique variance on each dependent measure that is accounted for by the main effects of L1 and L2 proficiencies and their interaction. Table 2 shows the significances of the full models of multiple regression analyses. There was no multicollinearity problem in the analyses, as supported by a low variance inflation ratio (<1.28). As verified by observing the residual scatterplots between the predicted values of dependent variables (i.e., Stroop performance) and errors of prediction, the assumptions of normality, homoscedasticity, and linearity for multiple regression analyses were generally met for all dependent variables that yielded significant full models. Given their nonsignificant full regression models, we did not discuss the regression findings of σ and error rate, although we still reported their statistics in Table 2 for the sake of completeness.

For overall RTs, the main effects of L1 and L2 proficiencies were significant (*beta* = –.40, $t(90) = 6.53$ and *beta* = –.34, $t(90) = 6.10$, respectively), but the L1 × L2 proficiency interaction was not (*beta* = –.04, $t(90) = .61$). Bilinguals with high L1 (or L2) proficiency responded faster than those with low L1 (or L2) proficiency. Analyses on ex-Gaussian parameters that were estimated by RTs collapsed across congruent and incongruent trials revealed significant main effects of L1 and L2 proficiencies for μ (*beta* = –.24, $t(90) = 2.87$ and *beta* = –.23, $t(90) = 3.11$, respectively) and τ (*beta* = –.34, $t(90) = 3.30$ and *beta* = –.25, $t(90) = 2.66$, respectively), suggesting that the benefit of L1/L2 proficiency in overall RTs was due to a reduction in both μ and τ (see Figure 5). None of the L1 × L2 proficiency interactions was significant for μ or τ (*beta* = –.01, $t(90) = .07$ and *beta* = –.07, $t(90) = .65$, respectively).

For the Stroop effect in RT, the main effects of language proficiency were significant for L1 (*beta* = –.31, $t(90) = 2.85$) and marginally so for L2 (*beta* = –.19, $t(90) = 1.93$, $p = .06$), but the L1 × L2 proficiency interaction was not (*beta* = –.02, $t(90) = .23$), indicating that the Stroop effect in RT decreased as a function of L1 and L2 proficiencies (see Figure 5). As shown in the ex-Gaussian analyses, the Stroop effect could be attributed to the change in μ, rather than in τ, in incongruent trials, relative to congruent trials. We computed the Stroop effect in μ for each individual participant by subtracting his/her μ estimated for congruent trials from his/her μ estimated for incongruent trials. Similar procedures were used to yield the Stroop effect in τ. The main effects of language proficiency on the Stroop effect in μ were significant for L1 (*beta* = –.27, $t(90) = 2.48$) and marginally so for L2 (*beta* = –.17, $t(90) = 1.70$, $p = .09$), but the L1 × L2 proficiency interaction was not (*beta* = –.04, $t(90) = .39$). None of the effects on the Stroop effect in τ was significant (all *beta*s < |.12|, *t*s < 1.19). The absence of a significant effect for the analyses on the Stroop effect may not be due to a lack of reliability in difference scores because we still obtained similar findings when we fit the regression model for RT, μ, and τ in incongruent trials, after controlling for RT, μ, and τ in congruent trials and other extraneous variables. These analyses yielded significant main effects of L1 and L2 proficiencies for RT (*beta* = –.16, $t(89) = 2.99$ and *beta* = –.11, $t(89) = 2.20$, respectively) and μ (*beta* = –.29, $t(89) = 3.62$ and *beta* = –.23, $t(89) = 3.09$, respectively), but not for τ (*beta* = –.07, $t(89) = .69$ and *beta* = –.05, $t(89) = .57$, respectively). None of the L1 × L2 proficiency interactions was significant in these measures (all *beta*s < |.04|, *t*s < .48).

When the RT, μ, and τ in congruent and incongruent trials were separately fit into the regression models, we obtained significant main effects of L1 proficiency for all measures in congruent trials (RT: *beta* = –.35, $t(90) = 5.57$; μ: *beta* = –.17, $t(90) = 2.03$; τ: *beta* = –.35, $t(90) = 3.33$) and for RT and μ in incongruent trials (RT: *beta* = –.41, $t(90) = 6.39$; μ: *beta* = –.33, $t(90) = 4.14$), but not for τ in incongruent trials (*beta* = –.20, $t(90) = 1.90$, $p = .06$). The main effects of L2 proficiency were significant for all measures in congruent trials (RT: *beta* = –.32, $t(90) = 5.54$; μ: *beta* = –.22, $t(90) = 2.82$; τ: *beta* = –.24, $t(90) = 2.47$) and for RT and μ in incongruent trials (RT: *beta* = –.34, $t(90) = 5.72$; μ: *beta* = –.28, $t(90) = 3.89$), but not for τ in incongruent trials (*beta* = –.14, $t(90) = 1.45$, $p = .15$). None of the L1 × L2 proficiency interactions was significant, all *beta*s < |.04|, *t*s < .59. Despite being nonsignificant at the conventional level ($p < .05$, two-tailed), the beta direction for τ in incongruent trials was consistent with those for τ in overall and congruent trials (see Figure 6).

Finally, analyses based on other measures of participants' L2 proficiency replicated the above results yielded from the factor scores of participants' self-rated L2 proficiency. First, we observed similar patterns for the effect of L2 proficiency that was quantified by mean-centered Shipley Vocabulary raw scores, except that the main effect of L2 proficiency on τ in incongruent trials was marginally significant, rather than nonsignificant (*beta* = –.18, $t(90) = 1.94$, $p = .06$). Second, after controlling for age, sex, handedness, socioeconomic status, and L1 proficiency factor score, the effect of English dominance (i.e., whether bilinguals ranked English as their most fluent language) revealed by ANOVAs mirrored those yielded in regression analyses based on the L2 proficiency factor scores. Third, regression analyses based on bilinguals' onset age of active bilingualism showed that the later their onset age of active bilingualism, the poorer their performance in the Stroop task. This echoes Luk et al.'s finding (2011) that attentional control abilities, as reflected

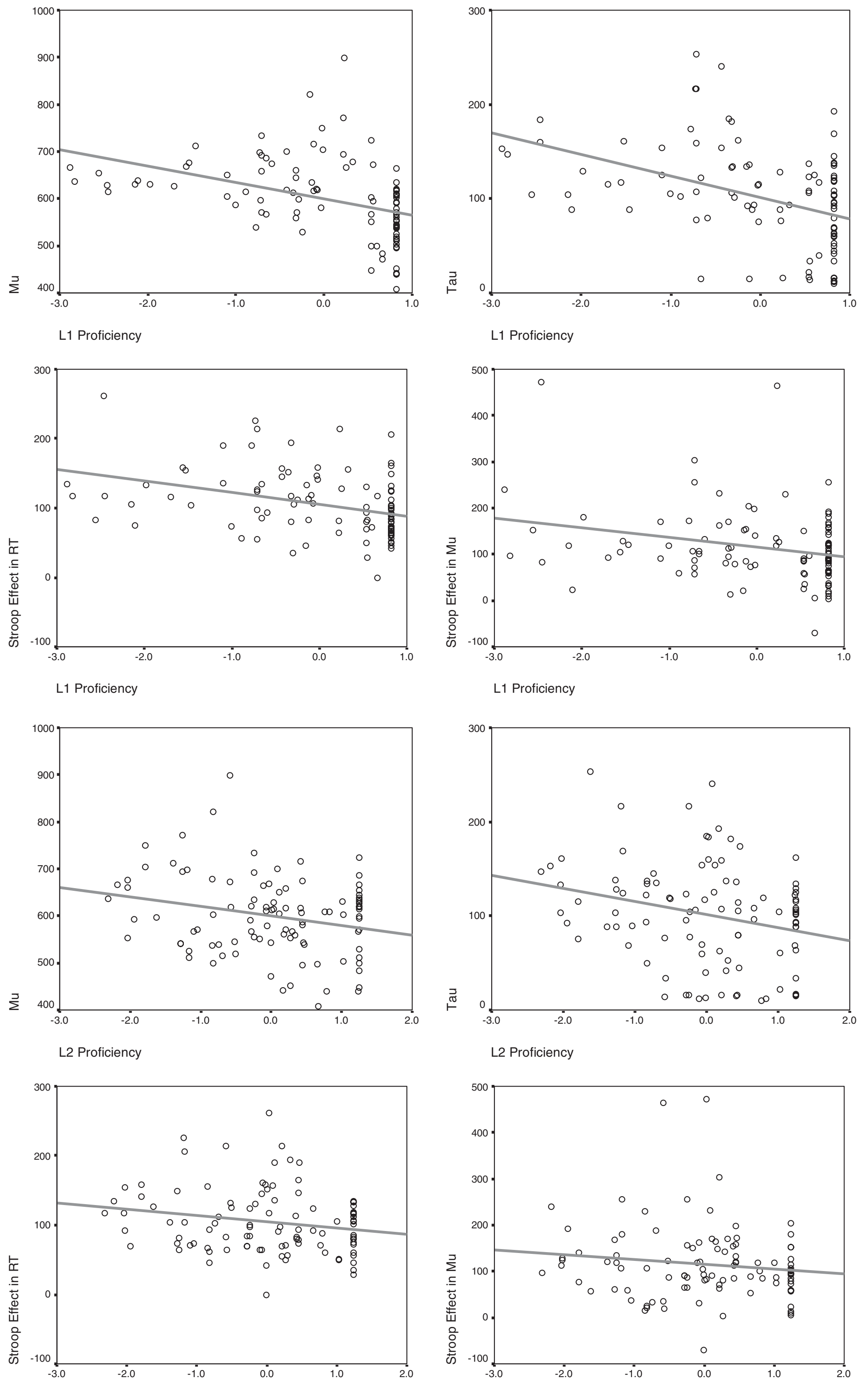

Figure 5. Scatterplots for μ (mu) and τ (tau) (based on RTs collapsed across all trials) and Stroop effect in RT and μ vs. L1 or L2 proficiency.

https://doi.org/10.1017/S1366728912000077 Published online by Cambridge University Press

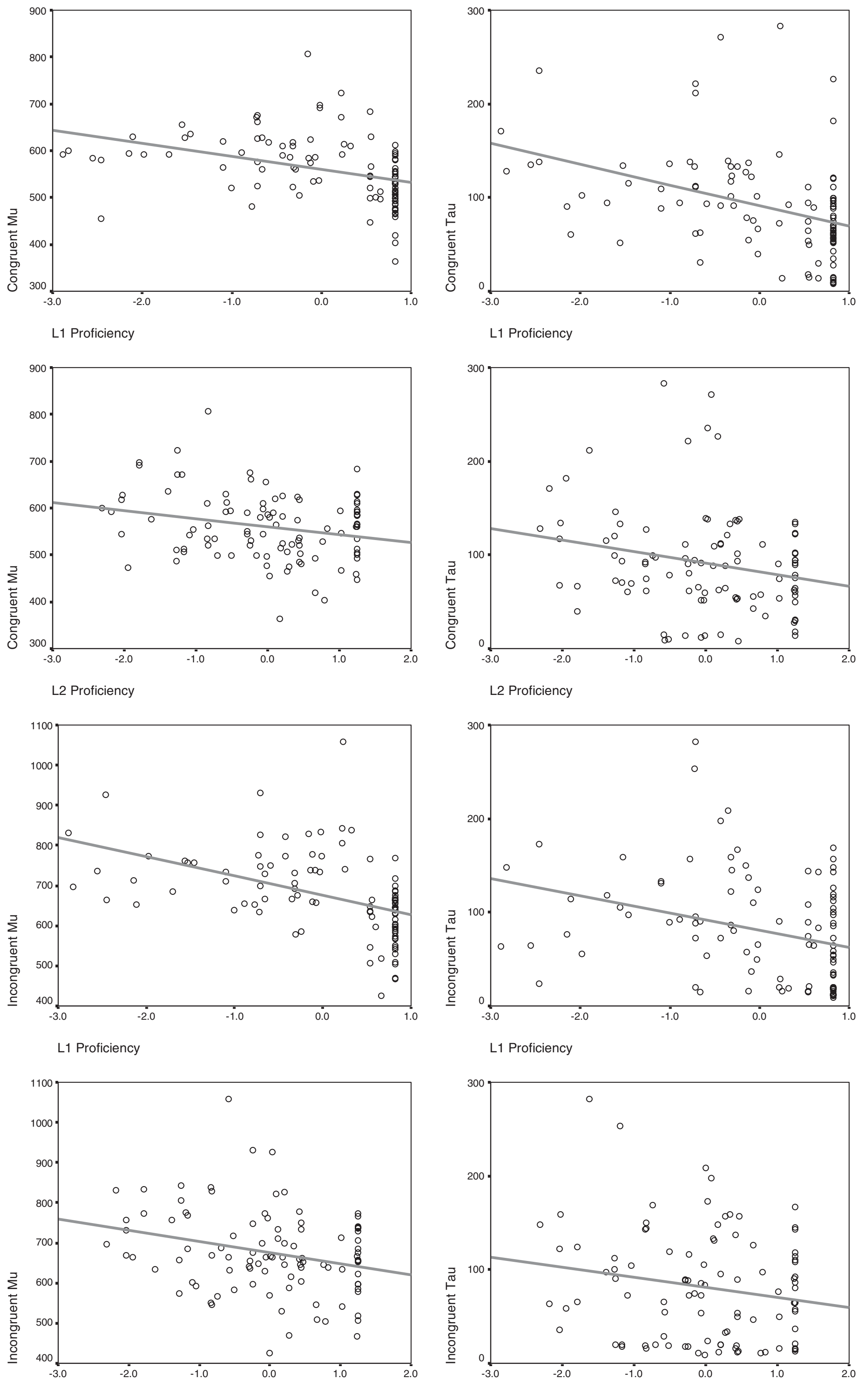

Figure 6. Scatterplots for μ (mu) and τ (tau) in congruent and incongruent trials vs. L1 or L2 proficiency factor scores.

https://doi.org/10.1017/S1366728912000077 Published online by Cambridge University Press

by their interference effect in a flanker task, were negatively correlated with bilinguals' onset age of active bilingualism. We also replicated the negative correlation between bilinguals' L2 vocabulary size, as currently indicated by Shipley Vocabulary raw scores, and their onset age of active bilingualism. In short, all these findings converged to the conclusions drawn from self-rated L2 proficiency factor scores. Due to their similarity with those reported above, we did not elaborate on the analyses based on these alternative L2 proficiency measures, although they can be obtained from the first author by request.

Discussion

In the current study, we investigated the influences of bilinguals' L1 and L2 proficiencies on performance in a Stroop task by analyzing their RT data at the mean and distributional levels. We obtained four major findings whether L2 proficiency was defined by participants' self-ratings, vocabulary scores, English dominance, or onset ages of active bilingualism. First, bilinguals responded faster as their L2 proficiency increased. This is consistent with previous studies that defined full vs. partial bilinguals only by L2 proficiency (e.g., Carlson & Meltzoff, 2008). We obtained similar findings for bilinguals' self-rated L1 proficiency (see Figure 5), demonstrating the positive influences of L1 and L2 proficiencies on overall RTs. Second, the extent to which bilinguals slowed down their RTs in incongruent trials, relative to congruent trials (i.e., Stroop effect) was associated with their L1/L2 proficiency, indicating the effect of bilinguals' L1/L2 proficiency on conflict resolution. Third, by partitioning bilinguals' overall RTs and Stroop effect into ex-Gaussian parameters, we found that L1 and L2 proficiencies modulated τ in overall RT that reflects goal maintenance and the Stroop effect in μ that reflects conflict resolution (see Figure 5). Fourth, regardless of RTs in congruent or incongruent trials, both μ and τ dropped as a function of L1 and L2 proficiencies (see Figure 6). Bilinguals' L1/L2 proficiency could affect their goal maintenance abilities since the L1/L2 proficiency modulated the size of τ even when bilinguals did not need to resolve any conflict between the color and word pathways in congruent trials. Overall, the findings suggest that both L1 and L2 proficiencies affect selective attention performance in the Stroop task via affecting both conflict resolution and goal maintenance, as conceptualized in Balota and Faust's (2001) attentional control framework (see Figure 1). To our knowledge, the current study is one of the first few to examine the relationship between bilinguals' L1/L2 proficiencies and attentional control at the level of RT distributions.[1]

It is noteworthy that there is no one-to-one mapping between ex-Gaussian parameters and underlying cognitive processes (Heathcote et al., 1991; Spieler et al., 2000); for example, the τ in the Stroop task may not reflect the same meaning as the one in the free recall task. We regard fitting the ex-Gaussian function as a way of describing the RT distribution and investigating how L1/L2 proficiencies affect various components of RT distributions. Given that RT distributional analyses should be coupled with specific computational models of a given task performance, the current analyses should be regarded as a descriptive account of the effect of L1/L2 proficiencies on the RT distributions in a selective attention task. Even though there are alternative ways to capture empirical RT distributions (see e.g., Van Zandt, 2002, for a review), we fit our data to ex-Gaussian functions to make contact with previous Stroop RT distribution studies that also used the ex-Gaussian function (e.g., Spieler et al., 1996; Tse et al., 2010). More importantly, as demonstrated in Vincentile plots (see Figure 4), our participants' empirical RT distributions were perfectly fitted with theoretical ex-Gaussian models.

One could argue that those affected more by L1/L2 proficiency could have used a more stringent response criterion. In making a decision, bilinguals with lower L1/L2 proficiency might be biased toward awaiting the results of further processing to distinguish among the word and color pathways, rather than producing any available "well-formed" response to compensate for a reduced reliability of their attentional control system, thereby making slower responses more frequently. This would predict that those who made fewer errors in their responses should yield larger τ. After partialing out age, sex, handedness, and socioeconomic status, we found null or even positive correlations between τ and error rates across bilinguals (overall: +.29, congruent: +.28, and incongruent: –.06). Those who showed a larger tail in

[1] Costa et al. (2009) identified the RT distributions for monolingual vs. bilingual performance, but they did not quantify the individual components of participants' RT distributions. When the current article was under review, we found another paper, Calabria, Hernandez, Martin and Costa (2011), which analyzed monolinguals' and bilinguals' performance in a flanker task via fitting them into ex-Gaussian distributions. Consistent with the current study, relative to monolinguals, bilinguals showed overall faster RTs, which could be captured by μ and τ. However, they found that the bilingual vs. monolingual difference in the flanker effect lay within τ, rather than μ, whereas we showed that the change of the Stroop effect as a function of bilinguals' L1/L2 proficiency lay within μ, rather than τ. Given the procedural differences between Calabria et al. and the current study, such as the sample population (monolingual and bilinguals vs. bilinguals with varied L1/L2 proficiency), the task (the flanker task that required key-press responses vs. the Stroop task that required vocal responses), and the task demand (unbalanced vs. balanced in the proportion of conflict vs. non-conflict trials), it is not clear which factor has contributed to the discrepancies in the findings between the two studies. Future research should include more than one attention task within the same experiment in order to test the generalizability of Calabria et al.'s and the current findings.

their RT distribution tended to be *less* accurate in their responses. Hence, the inter-individual differences in τ could not be entirely attributed to participants' speed–accuracy tradeoffs.

The result that the Stroop effect decreased as a function of L2 proficiency within bilinguals was compatible with most of the previous studies that focused on the monolingual vs. bilingual comparison (e.g., Bialystok et al., 2008) as long as monolinguals can be regarded as bilinguals with NO knowledge in L2 (i.e., the least L2-proficient bilinguals). The unique contribution of L1 proficiency on the reduction of the Stroop effect suggests that, at least within bilinguals, the ability to resolve conflicting information also depends on L1 proficiency. These findings were conceptually consistent with the monolingual findings in Spieler et al. (2000). Similar to their interference effect in their global–local task, the Stroop effect could be attributed to an increase in μ, rather than τ, in incongruent trials, relative to congruent trials. The analyses on the component of RT distribution for the Stroop effect showed that the Stroop effect in μ was associated with bilinguals' L1/L2 proficiency, similar to the results that we obtained for the Stroop effect in RTs. The positive influence of L1/L2 proficiency on goal maintenance, as reflected by τ in overall RTs, echoed the enhanced monitoring abilities reported for bilinguals (vs. monolinguals) in previous studies using different tasks (e.g., Attention Network Test (ANT) in Costa et al., 2009). Communication among bilinguals taps into the monitoring of the language to be used during a conversation. This process is not the same as those involved in conflict resolution: bilinguals should first decide the suitable language and then conduct lexicalization in that language without interference from the other. Hence, goal maintenance may be sensitive to L1/L2 proficiency even without the need to perform conflict resolution in bilingual language processing. This can be indirectly shown by the facilitating effect of L1/L2 proficiency on goal maintenance (as reflected by τ) even in the congruent trials that did not involve conflict resolution. Given that the proportion of congruent and incongruent trials were identical (50:50) in our task, bilinguals needed to monitor the requirements of each trial constantly and consider whether they would have to ignore the information provided by the word pathway, such that they could take specific actions (i.e., reading aloud the ink color). Costa et al. (2009) found a larger bilingual vs. monolingual difference in overall RTs when the selective attention task demanded higher monitoring – equal proportion of two types of trials – than when it demanded lower monitoring – unbalanced proportion of two types of trials. Future studies may test whether the L1/L2 proficiency associated with τ would be weakened when incongruent and congruent trials were no longer balanced in proportion, in order to clarify the relationship between bilinguals' L1/L2 proficiency and goal maintenance.

The modulation of bilinguals' L1 and L2 proficiencies on goal maintenance and conflict resolution in selective attention performance is compatible with the view that bilinguals need to decide which language should be prioritized for production by selecting the suitable language schema (e.g., Green, 1998) before proceeding to subsequent lexical processing. This checking process is analogous to the maintenance of task goal in the selective attention task. After this decision is made, bilinguals need to resolve the lexical competition between languages, which is analogous to the conflict resolution mechanism in the selective attention task. We showed that bilinguals with higher L1/L2 proficiency, who have more practice with these two processes than those with lower L1/L2 proficiency, possessed stronger abilities in conflict resolution and goal maintenance and in turn outperformed them in the selective attention task.

While in congruent trials no conflict resolution was presumably necessary in the access to word pathway vs. color pathway in the Stroop color-naming task, one could argue that conflict resolution could still occur during color-name word retrieval. For instance, when participants read aloud "blue" to a stimulus *blue* printed in blue color, they might still be interfered with by other words that share semantic or phonological features (e.g., *black* and *blew*, respectively). This conflict is likely to be stronger in bilinguals because (i) relative to monolinguals, they do not use the word label in either language as often (see weaker links hypothesis in Gollan, Montoya, Cera & Sandoval, 2008); and (ii) they might need to resolve the conflict among multiple linguistic labels for the same word (e.g., suppressing the L1 label *bleu* in order to produce the L2 label *blue*). Those with high L1 and L2 proficiencies would be more able to resolve this cross-language conflict in word retrieval than those who are proficient in one language or not proficient in either language. This could explain the change of μ as a function of L1 and L2 proficiencies in congruent trials. After separating from the influence of goal maintenance (as indicated by τ), those with higher L1/L2 proficiencies were more able to resolve this word-retrieval conflict and hence respond faster in congruent trials across the RT bins. Even though we did control for bilinguals' baseline naming time (for color patches and color names) in our analyses, it is not clear whether this procedure could take the word-retrieval factor into account. To clarify this issue, it is necessary to replicate the current study with a task that involves nonverbal materials (e.g., arrow) yet does not require vocal responses (e.g., key-press responses). We are now running an experiment involving the Simon task with arrow stimuli to test this notion and to track down bilinguals' developmental trajectories for their conflict resolution and goal maintenance abilities.

Before concluding this study, it is important to note that self-rated L1/L2 proficiency is not the only factor that influences attentional control. Given that bilingualism is a multidimensional experience, bilinguals' L1/L2 proficiency could be influenced by various factors, such as the frequency of usage for the two languages and the frequency of code switching between them. As summarized in the Results section, we did obtain consistent patterns whether bilinguals' L2 proficiency was defined by their self-ratings, Shipley Vocabulary raw scores, self-reported English dominance, or onset age of active bilingualism. However, other factors (e.g., the extent to which bilinguals switch codes), which we did not measure in the current research, could have also contributed to the superior performance of our more proficient bilinguals. Future studies should compare different indices of bilingualism and pinpoint those that are more closely related to bilinguals' attentional control abilities.

In conclusion, we are the first to show that bilinguals' both L1 and L2 proficiencies were associated with (i) the tail size (τ) of the RT distribution in a Stroop task, whether we considered the RTs in all trials or in congruent trials, and (ii) the Stroop effect (in RT/μ), which reflects the degree to which bilinguals could resolve the conflict in incongruent trials, relative to congruent trials. This supports the roles of goal maintenance and conflict resolution in the effect of bilingualism on selective attention (Bialystok, 2010; Costa et al., 2008). By analyzing the data at the level of RT distributions in the study of bilingualism, we further our understanding of the interaction between language and cognition (see also Calabria, Hernández, Martin, & Costa, 2011). More research should be done within bilingual populations by taking into account both of their L1 and L2 proficiencies to shed light on how the processes involved in coordinating between two language systems could enhance one's attentional control system.

References

Abutalebi, J., & Green, D. W. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20,* 242–275.

Balota, D. A., & Faust, M. E. (2001). Attention in dementia of the Alzheimer's type. In F. Boller, & S. Cappa (eds.), *The handbook of neuropsychology: Aging and dementia* (2nd edn.), pp. 51–80. New York: Elsevier.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20,* 160–166.

Balota, D. A., Yap, M. J., Cortese, M. I., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59,* 495–523.

Bialystok, E. (2009). Bilingualism: The good, the bad and the indifferent. *Bilingualism: Language and Cognition, 12,* 3–11.

Bialystok, E. (2010). Global–local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. *Developmental Psychology, 46,* 93–105.

Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition, 122,* 67–73.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34,* 859–873.

Bialystok, E., Craik, F. I. M., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32,* 1341–1354.

Bialystok, E., & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and Language, 109,* 93–100.

Calabria, M., Hernández, M., Martin, C. D., & Costa, A. (2011). When the tail counts: The advantage of bilingualism through the ex-Gaussian distribution analysis. *Frontiers in Psychology, 2 (250),* 8 pages. doi: 10.3389/fpsyg.2011.00250.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11,* 282–298.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113,* 135–149.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106,* 59–86.

Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition, 9,* 137–152.

Cousineau, D., Brown, S., & Heathcote, A. (2004). Fitting distributions using maximum likelihood: Methods and packages. *Behavior Research Methods, Instruments, & Computers, 36,* 742–756.

de Jong, R. D., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica, 101,* 379–394.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58,* 787–814.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67–81.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109,* 340–347.

Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to

Stroop interference. *Journal of Experimental Psychology: General, 132,* 47–70.

Leth-Steensen, C., King Elbaz, Z., & Douglas, V. I. (2000). Mean response time, variability, and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica, 104,* 167–190.

Luk, G., de Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition, 14,* 588–595.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114,* 29–41.

MacLeod, C. M. (1992). The Stroop task: The "gold standard" of attentional measures. *Journal of Experimental Psychology: General, 121,* 12–14.

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition, 11,* 1–13.

McAuley, T., Yap, M. J., Christ, S. E., & White, D. A. (2006). Revisiting inhibitory control across the life span: Insights from the ex-Gaussian distribution. *Developmental Neuropsychology, 29,* 447–458.

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10,* 719–726.

Piai, V., Roelofs, A. P. A., & Schriefers, H. J. (2011). Semantic interference in immediate and delayed naming and reading: Attention and task decisions. *Journal of Memory and Language, 64,* 404–423.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136,* 414–429.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *Journal of Psychology, 9,* 371–377.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 461–479.

Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 506–526.

Tao, L., Marzecova, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: The role of age of acquisition. *Frontiers in Psychology, 2,* 123.

Tse, C.-S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of healthy aging and early-stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology, 24,* 300–315.

Unsworth, N., Spillers, G. J., Brewer, G. A., & McMillan, B. (2011). Attentional control and the antisaccade task: A response time distribution analysis. *Acta Psychologica, 137,* 90–100.

Van Zandt, T. (2002). Analysis of response time distributions. In J. T. Wixted (ed.), *Stevens' handbook of experimental psychology* (vol. *4*): *Methodology in experimental psychology* (3rd edn.), pp. 461–516. New York: Wiley.

West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. M., & Stuss, D. T. (2002). Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control. *Brain and Cognition, 49,* 402–419.

Yap, M. J., Tse, C.-S., & Balota, D. A. (2009). Individual differences in the joint effects of semantic priming and word frequency: The role of lexical integrity. *Journal of Memory and Language, 63,* 303–325.