In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Supplementary Materials

Data Preprocessing Details

First, we removed the first trial for every participant, because many participants began the task while the researcher was still in the room, or as they were leaving the room.

Next, to ensure that all items were interpreted as they were intended, we calculated the mean coherence score for each inference condition for the L1 English group. We based data preprocessing on the responses of the L1 English group because this group demonstrated greater overall English proficiency than the L2 English group. We calculated a 95% plausible values interval for each condition by adding and subtracting two standard deviations to the mean of each condition, across all items (Logical = 3.15- 5; Mental State = 1.95-5; Incoherent = 0-2.79). From there, we calculated the mean coherence score for each item in its three conditions and compared the by-item scores to the overall 95% plausible values interval. Any item whose coherence to the logical or mental state conditions was *below* its lower threshold was discarded (resulting in one discarded item). Similarly, any incoherent condition that was scored *above* its upper coherence threshold was discarded (resulting in two discarded items).

We then assessed whether any single subjects needed to be removed from the sample as a result of having a large proportion of extremely slow or extremely fast trials. First, we imposed an upper time cut-off on trials with a high global reading time (we noted that a few participants paused on a certain trial to sneeze, ask for water, or use the restroom). In visualizing the distribution of the global reading times, it seemed that 10 seconds would be an appropriate and generous upper threshold (the average number of words per item was 13). We observed that one participant would have had 35% of their trials removed with this threshold, which was double the number of trials that the next slowest participant would have had removed with the threshold. We decided to remove that particular individual. Afterwards, we removed all single trials where the global reading time exceeded 10 seconds (3% of all trials). Similarly, we imposed a lower time cut-off for fast global reading time (we noted that some individuals displayed fast global reading times and slow coherence rating times, suggesting that these individuals were prematurely pressing the spacebar and waiting to read the item only once the first judgment question appeared). Again, given an average of 13 words per item (each needing at least 250 ms for a single fixation) made it unlikely that an item would be read in less than 2500 ms. We applied a conservative lower threshold at 1000 ms. Four participants were found to average below 1000 ms on global reading time across all items, and so their global reading time and reaction times for both judgment questions were discarded. The actual judgments, nevertheless, were retained.

Lastly, we applied trial-level cleaning to each of the timed outputs. Beginning with global reading time, we removed all single trials where the global reading time was below 500 ms (1.6% of all trials). Next, we took that subset and further removed all single trials where the coherence judgment reaction time took greater than 10 seconds (2.2% of all trials). Then, we took the global reading time subset and further removed all single trials where the mentalizing judgment reaction time took greater than 10 seconds (1.3% of all trials).

In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Results of coherence and mentalizing reaction times

Across all models, language diversity was treated continuously, language group was treatment coded (L1 English = 0, L2 English = 1), and inference type was Helmert coded (Contrast 1 = Mental state vs. logical inferences, Contrast 2 = Mean of mental state + logical vs. incoherent).

Coherence Reaction Time.

We computed a linear mixed-effects regression model to predict reaction time (ms) to the coherence rating, which we log-transformed for normality. First, we detected a significant main effect of our core inference type manipulation (C1: Log/Men = $\beta = 0.067$, SE = 0.014, $t = 4.73$, $p < 0.001$; C2: Coh/Inc = $\beta = -0.109$, SE = 0.008, $t = -13.25$, $p < 0.001$). Here, coherence reaction times for mental state inferences are slower than logical inferences across all readers. Moreover, coherence reaction times for incoherent items are faster than those for all coherent items. Moreover, we detected a significant interaction between inference type and language diversity, which was qualified by a higher-order interaction with language group. Both the 2-way and the 3-way interactions were only detected at the second contrast (3-way interaction, C2: Coh/Inc = $\beta = 0.061$, SE = 0.011, $t = 5.13$, $p < 0.001$). In other words, coherence reaction times to all coherent vs. incoherent items varied as a function of language diversity and language group. We ran a series of follow-up linear regressions to determine whether the relationship between language diversity and language group was driven by differences in coherence reaction times to coherent items, incoherent items, or both. The model conducted on all coherent items (logical and mental state) did not return a significant interaction between language diversity and language group, but the model conducted on incoherent items did return a significant 2-way interaction (Incoherent: = $\beta = 0.295$, SE = 0.128, $t = 2.27$, $p = 0.027$). This confirms that the difference in coherence reaction time between all coherent and incoherent items is driven by the incoherent items. Moreover, the direction of this relationship, as indicated by the beta coefficient of the model, is such that greater language diversity in the L2 English group links to slower coherence reaction times of incoherent items than it does in the L1 English group. Overall, the fixed effects of this model accounted for 10.3% variance of the data (marginal R^2), and the fixed and random effects of this model accounted for 28.4% of variance of the data (conditional R^2).

Mentalizing Reaction Time

Similarly, we computed a linear mixed-effects regression model to predict reaction time (ms) to the mentalizing rating, which we log-transformed for normality. We again detected a significant main effect of our core inference type manipulation (C1: Log/Men = $\beta = -0.045$, SE = 0.016, $t = -2.78$, $p < 0.01$; C2: Coh/Inc = $\beta = -0.148$, SE = 0.009, $t = -15.96$, $p < 0.001$). These results indicate that mentalizing reaction times for mental state inferences were faster than logical inferences across all readers. Similarly, mentalizing reaction times for incoherent items were faster than those for all coherent items across all readers. Although we detected a significant interaction between language diversity and language group ($\beta = 0.305$, SE = 0.141, $t = 2.15$, $p = 0.036$), this did not involve our core inference type manipulation. Altogether, the fixed effects of this model accounted for 16.6% variance of the data (marginal R^2), and the fixed and random effects of this model accounted for 42.2% of variance of the data (conditional R^2).

In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Results of Global Reading Time, accounting for item length differences by random slopes

We accounted for differences in item lengths using two methods: (1) Dividing total reading time by number of characters per item, and (2) Including number of characters as a by-subject random slope. There were no differences for the manipulation check nor the individual differences model. Here, we report the results of the second approach.

Manipulation Check

```
global.rt.manip2 = lmer(log(text.RT) ~ condition+
                  (1 + scale(text.char.total) || subject) +
                  (1 | itemnum), df.rt, contrasts = list(condition = cHelmert), REML=F)
```

effect	term	estimate	std.error	t-statistic	p.value
fixed	(Intercept)	8.028	0.045	175.300	0.000
fixed	Inference C2(Coh/Inc)	0.010	0.003	2.600	0.009
fixed	Inference C1(Log/Men)	0.008	0.007	1.100	0.270

In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Individual Differences

```
global.rt2 = lmer(log(text.RT) ~ condition*scale(general.entropy)*L1.group + scale(trial.order)
+scale(percent.use.English)+
        (1 + scale(text.char.total) | subject) +
        (1 | itemnum), df.rt, contrasts = list(condition = cHelmert, L1.group = cTreatment),
REML=F)
```

effect	term	estimate	std.error	t-statistic	p.value
fixed	(Intercept)	7.948	0.074	108.090	0.000
fixed	Inference C2(Coh/Inc)	0.001	0.006	0.220	0.825
fixed	Inference C1(Log/Men)	0.001	0.011	1.040	0.297
fixed	scale(general.entropy)	-0.200	0.092	-2.190	0.033
fixed	Language (L2)	0.133	0.110	1.220	0.229
fixed	scale(trial.order)	-0.132	0.005	-24.780	0.000
fixed	scale(percent.use.English)	-0.113	0.079	-1.440	0.156
fixed	Inference C2(Coh/Inc):scale(general.entropy)	-0.005	0.006	-0.820	0.409
fixed	Inference C1(Log/Men):scale(general.entropy)	0.003	0.011	0.330	0.739
fixed	Inference C2(Coh/Inc):Language (L2)	0.008	0.009	0.940	0.346
fixed	Inference C1(Log/Men):Language (L2)	0.004	0.015	0.240	0.812
fixed	scale(general.entropy):Language (L2)	0.125	0.102	1.220	0.227
fixed	Inference C2(Coh/Inc):scale(general.entropy):Language (L2)	0.014	0.009	1.60	0.109
fixed	Inference C1(Log/Men):scale(general.entropy):Language (L2)	-0.018	0.015	-1.210	0.225

In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Full Model Outputs

Coherence Ratings

effect	term	estimate	std.error	t-statistic	p.value
fixed	(Intercept)	3.422	0.069	49.725	0.000
fixed	Inference C2(Coh/Inc)	-1.074	0.011	-99.862	0.000
fixed	Inference C1(Log/Men)	-0.237	0.019	-12.681	0.000
fixed	scale(general.entropy)	-0.050	0.078	-0.637	0.527
fixed	Language (L2)	-0.215	0.102	-2.109	0.039
fixed	scale(trial.order)	0.018	0.009	1.915	0.055
fixed	scale(percent.use.English)	-0.065	0.061	-1.071	0.288
fixed	Inference C2(Coh/Inc):scale(general.entropy)	0.001	0.012	0.100	0.920
fixed	Inference C1(Log/Men):scale(general.entropy)	-0.024	0.020	-1.213	0.225
fixed	Inference C2(Coh/Inc):Language (L2)	0.095	0.015	6.248	0.000
fixed	Inference C1(Log/Men):Language (L2)	-0.026	0.026	-1.005	0.315
fixed	scale(general.entropy):Language (L2)	0.126	0.092	1.369	0.176
fixed	Inference C2(Coh/Inc):scale(general.entropy):Language (L2)	0.012	0.015	0.772	0.440
fixed	Inference C1(Log/Men):scale(general.entropy):Language (L2)	0.049	0.027	1.815	0.070
random	Item	0.211			
random	Subject	0.296			
random	Residual	0.829			

In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Mentalizing Ratings

effect	term	estimate	std.error	t-statistic	p.value
fixed	(Intercept)	2.536	0.104	24.332	0.000
fixed	Inference C2(Coh/Inc)	-0.401	0.017	-24.282	0.000
fixed	Inference C1(Log/Men)	0.789	0.029	27.463	0.000
fixed	scale(general.entropy)	-0.029	0.120	-0.243	0.809
fixed	Language (L2)	0.044	0.157	0.278	0.782
fixed	scale(trial.order)	-0.026	0.014	-1.793	0.073
fixed	scale(percent.use.English)	-0.014	0.094	-0.145	0.885
fixed	Inference C2(Coh/Inc):scale(general.entropy)	-0.074	0.018	-4.165	0.000
fixed	Inference C1(Log/Men):scale(general.entropy)	0.087	0.031	2.827	0.005
fixed	Inference C2(Coh/Inc):Language (L2)	-0.078	0.023	-3.332	0.001
fixed	Inference C1(Log/Men):Language (L2)	-0.173	0.040	-4.276	0.000
fixed	scale(general.entropy):Language (L2)	0.106	0.142	0.746	0.459
fixed	Inference C2(Coh/Inc):scale(general.entropy):Language (L2)	0.114	0.024	4.798	0.000
fixed	Inference C1(Log/Men):scale(general.entropy):Language (L2)	0.006	0.041	0.142	0.887
random	Item	0.236			
random	Subject	0.456			
random	Residual	1.274			

In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Global Reading Time

effect	term	estimate	std.error	t-statistic	p.value
fixed	(Intercept)	3.679	0.074	49.971	0.000
fixed	Inference C2(Coh/Inc)	0.004	0.006	0.639	0.523
fixed	Inference C1(Log/Men)	0.009	0.011	0.884	0.377
fixed	scale(general.entropy)	-0.198	0.092	-2.162	0.035
fixed	Language (L2)	0.140	0.110	1.268	0.210
fixed	scale(trial.order)	-0.132	0.005	-24.851	0.000
fixed	scale(percent.use.English)	-0.112	0.079	-1.422	0.160
fixed	Inference C2(Coh/Inc):scale(general.entropy)	-0.004	0.006	-0.696	0.487
fixed	Inference C1(Log/Men):scale(general.entropy)	0.002	0.011	0.187	0.852
fixed	Inference C2(Coh/Inc):Language (L2)	0.009	0.009	1.049	0.294
fixed	Inference C1(Log/Men):Language (L2)	0.005	0.015	0.314	0.754
fixed	scale(general.entropy):Language (L2)	0.123	0.102	1.205	0.233
fixed	Inference C2(Coh/Inc):scale(general.entropy):Language (L2)	0.013	0.009	1.504	0.132
fixed	Inference C1(Log/Men):scale(general.entropy):Language (L2)	-0.018	0.015	-1.164	0.244
random	Item	0.061			
random	Subject	0.325			
random	Residual	0.450			

In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Coherence Reaction Time

effect	term	estimate	std.error	t-statistic	p.value
fixed	(Intercept)	7.548	0.067	112.220	0.000
fixed	Inference C2(Coh/Inc)	-0.109	0.008	-13.249	0.000
fixed	Inference C1(Log/Men)	0.068	0.014	4.728	0.000
fixed	scale(general.entropy)	0.041	0.084	0.488	0.627
fixed	Language (L2)	0.060	0.100	0.598	0.552
fixed	scale(trial.order)	-0.134	0.007	-18.639	0.000
fixed	scale(percent.use.English)	0.146	0.072	2.034	0.047
fixed	Inference C2(Coh/Inc):scale(general.entropy)	-0.026	0.009	-3.035	0.002
fixed	Inference C1(Log/Men):scale(general.entropy)	-0.009	0.015	-0.627	0.531
fixed	Inference C2(Coh/Inc):Language (L2)	-0.006	0.012	-0.507	0.612
fixed	Inference C1(Log/Men):Language (L2)	-0.040	0.020	-1.940	0.052
fixed	scale(general.entropy):Language (L2)	0.159	0.093	1.707	0.093
fixed	Inference C2(Coh/Inc):scale(general.entropy):Language (L2)	0.061	0.012	5.135	0.000
fixed	Inference C1(Log/Men):scale(general.entropy):Language (L2)	0.003	0.021	0.169	0.866
random	Item	0.068			
random	Subject	0.294			
random	Residual	0.599			

In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing by Tiv, M., O'Regan, E., & Titone, D.

Mentalizing Reaction Time

effect	term	estimate	std.error	t-statistic	p.value
fixed	(Intercept)	7.044	0.102	68.948	0.000
fixed	Inference type - C2: Coh/Inc	-0.148	0.009	-15.958	0.000
fixed	Inference C1(Log/Men)	-0.045	0.016	-2.782	0.005
fixed	scale(general.entropy)	-0.096	0.127	-0.752	0.455
fixed	Language (L2)	0.107	0.153	0.698	0.488
fixed	scale(trial.order)	-0.223	0.008	-27.512	0.000
fixed	scale(percent.use.English)	0.238	0.109	2.176	0.034
fixed	Inference type - C2: Coh/Inc:scale(general.entropy)	-0.003	0.010	-0.333	0.739
fixed	Inference C1(Log/Men):scale(general.entropy)	-0.029	0.017	-1.735	0.083
fixed	Inference type - C2: Coh/Inc:Language (L2)	0.012	0.013	0.876	0.381
fixed	Inference C1(Log/Men):Language (L2)	0.027	0.023	1.185	0.236
fixed	scale(general.entropy):Language (L2)	0.305	0.142	2.151	0.036
fixed	Inference type - C2: Coh/Inc:scale(general.entropy):Language (L2)	0.014	0.013	1.019	0.308
fixed	Inference C1(Log/Men):scale(general.entropy):Language (L2)	0.009	0.023	0.409	0.682
random	Item	0.047			
random	Subject	0.450			
random	Residual	0.680			