Effect of self-referent primes about language and memory on measures of working memory in individuals with and without second languages

Paul Marshall, Alex Chisholm, Stephanie De Lisio, Zahra Kamal

Department of Psychology, University of Guelph, Guelph, ON Canada. Faculty supervisor: Dr. Jeffery Spence
For correspondence, please email: spencejr@uoguelph.ca

Abstract

Canada is assumed to be a multi-cultural country. Thus, many assume that stereotypes against those with multiple languages should be minimally prevalent and minimally effective. However, studies seem to suggest that despite a multi-cultural image, Canada is predominantly an English speaking society, and that those who primarily speak English may more easily obtain jobs, with higher pay grades. This may occur because individuals whose dominant language is not English may not be as proficient at skills necessary for these jobs. Alternatively, it may be due to negative stereotypes that exist for people who speak a second language (2LS), while positive stereotypes may be present for those who primarily speak English (W2L). These stereotypes may then lead to alterations in performance and evaluation. Therefore to differentiate between these two explanations we subjected 2LS and W2L speakers to negative, neutral and positive primes about having a second language and then measured working memory (WM) with a modified reading span task. It was predicted that positive self-referent primes would enhance, whereas negative primes would disrupt WM, and that neutral primes would not differ among language groups. A significant crossover interaction between prime manipulations and conditions was demonstrated with no main effect of language group. These results suggest that while WM capacity does not differ between language groups, when these same groups are given self-referent positive and negative primes, these can act through stereotypes to reduce performance creating a self-confirming bias. This may account for why those whose first language is not English may not be chosen for jobs which require higher intellectually functioning. Thus, in order to close wage gaps between these populations specific focus may have to be paid to altering these stereotypes, as opposed to converting individuals to speak primarily English.

Keywords: Working memory, priming, second language, stereotype threat, stereotype arousal, stereotype

Introduction

Stereotypes and society

During the 21st century, individuals that immigrated to Canada accounted for 8-51% of yearly population growth (Boyd and Vickers 2000). In response to this large population increase Canada has, for the most part, adopted a multi-cultural system whereby individuals are encouraged to accept a variety of cultures and languages in their daily life. This is counter to other nations such as the United States which have, in the past, tended to hold a melting pot philosophy for immigrants, such that individuals are expected to conform to the culture and languages of the area they immigrate to. Thus, it is assumed that immigrants in Canada who possess a language other than English are not at

any disadvantage. In fact many individuals are assumed to be at a benefit in some cases because there are some professions that require or encourage second languages, such as government jobs or international business (Mady 2012).

However, despite many attesting that Canada is a multi-cultural society, represented by it having two national languages (both French and English), others contend that Canada is still an English dominant country (Mady 2012). This means that as much as a boon as dual fluency may be if one possesses it, fluency in languages other than English may not be valued as much as mastery of English. This is important because English fluency has been shown to have a direct impact on earning potential and job opportunities (Chiswick and Miller 1995; Chiswick and Miller 2003).

Together, this suggests that although Canada may appear to be a multicultural nation, biases may still exist for those with second languages, resulting in very real impacts on individuals (Judd and Park 1993).

One possibility of how this might occur is through stereotypes. These are an individual's set of personal beliefs about characteristics and attributes for a specific group (Judd and Park 1993). These beliefs are unfortunately a ubiquitous and automatic part of human behaviour because they aide in categorization, such that we may use this information to make decisions without expending significant cognitive effort (Allport 1956; Macrae, Milne, and Bodenhausen 1994). Furthermore, it has been argued that stereotypes and attitudes are a part of social tradition of society (Devine 1989). Thus, every individual is subjected to learning about the stereotypes and attitudes that are assigned to major ethnic groups.

Consequently, due to their universal nature, stereotypes are based on characteristics or traits including race, (Claude M. Steele and Aronson 1995) gender, (Rydell, McConnell, and Beilock 2009), as well as age (Chen and Bargh 1997). Notably, individuals tend to hold positive stereotypes for in-group (or members inside their same perceived group) and negative stereotypes for out-group (or members outside their own perceived group). Thus a consequence of this may be that if the dominant language is English, people who speak primarily English may possess negative stereotypes for those who speak other languages because they would not be a part of their group, whereas those with second languages may demonstrate the opposite effect (Shapiro and Neuberg 2007).

Stereotypes effects: arousal and threat

Stereotypes are judged first on whether they are relevant to the individual and then on whether they are positive or negative, from which two consequences may derive. Stereotypes that are relevant to self and perceived to be positive may facilitate performance, and this process is called stereotype arousal (Ben-Zeev, Fein, and Inzlicht 2005). However, if a stereotype is relevant to the individual but negative, this may inhibit performance, known as stereotype threat (Schmader, Johns, and Forbes 2008).

However, individuals can often belong to more than one group, which may have different consequences depending on what stereotype is primed. For example, when primed with the stereotype that females perform worse at math, Asian females performed significantly worse on a test as compared to those who were not primed. But curiously, in this same study, facilitation was shown to occur by stereotype arousal such that women, when primed with their Asian race instead of gender, tended to perform better on the math questions compared to those who were not primed at all (Rydell, McConnell, and Beilock 2009). Thus, stereotype threat depends on the nature of the stereotype that was salient to the individual.

Stereotype priming: Implicit, explicit and ability to induce

An additional illustration of differences in priming of stereotypes is that stereotypes may be primed implicitly or explicitly. For example, implicit primes were elegantly demonstrated by Bargh, Chen, and Burrows (1996) utilizing a scrambled-sentence task disguised as a test of language, which contained either an elderly stereotype (i.e. Florida, old, forgetful) or a neutral word stereotype (i.e. thirsty, clean, private). It was found that participants in the implicit elderly prime condition walked more slowly than more neutral primes when leaving the testing rooms. However, explicit primes have also been used, and can be grouped into both blatant and subtle primes, whereby the blatant primes are more obvious signals of the particular stereotypes.

These two types of explicit primes have produced mixed results; some negative with blatant primes, positive with subtle (Hess, Hinson, and Statham 2004), and others positive with subtle, and negative with blatant (Shih *et al.* 2002). Therefore, it is clear that depending on how the stereotype is activated and what the relevance is to the individual, it may have different effects. Additionally, it may be more ecologically valid to use subtle stereotype primes in experimentation as this is more likely to occur than blatant primes because of the growing societal intolerance against overtly stereotyping against individuals.

How do stereotypes effect performance: working memory?

Stereotype priming effects and the possible explanations for them vary widely in the literature. However, one prominent view is that stereotype threat or arousal can modulate memory functioning by acting on working memory (WM; Rydell, McConnell, and Beilock 2009). WM classically refers to a system which begins with incoming sensory information that is temporarily stored before becoming a part of long term memory (Miller, Galanter, and Pribram 1960). More recently, WM has been defined as the active component of processing by facilitating the interface between items stored in memory and current incoming information. It thus allows incoming sensory information to be processed or manipulated in reference to previously encountered items, which may both direct actions and thoughts (Baddeley 2003).

The problem is that WM is a limited-capacity workspace meaning that the more one taxes WM by adding more items to process or manipulate, the less efficient it is (Baddeley, 2003). This is important because WM has been demonstrated to be reduced in negative stereotypes for age, gender, and race (Levy 1996; Schmader and Johns 2003) or enhanced by priming race (Rydell *et al.*, 2009) or other positive self-relevant primes (Shih *et al.* 2002).

But even more importantly WM has been linked with everything from intelligence (Colom, Flores-Mendoza, and Rebollo 2003), reading ability (Daneman and Carpenter 1980; Carretti *et al.* 2009) and controlling higher level

cognitive processes (Lépine, Parrouillet, and Camos 2012). Thus these alterations by stereotypes are poised to disrupt many important functions and measurements necessary for individuals' success or evaluation in something such as performed when trying to obtain employment (Campion, Pursell, and Brown 1988).

But which is it; does language group predict working memory ability, or are the stereotypes about language groups effecting working memory? In response to this kind of question, Papagno and Vallar (1995) found that people who spoke 3 languages or more (polyglots) were better at phonological working memory tasks and paired associate tasks of foreign words than people who spoke 2 languages (biglots). However, no differences were found in visuo-spatial working memory, general intelligence and paired associate tasks in their native language, meaning increased performance was isolated to memory processes related to learning languages, not functioning in an already acquired language. Even though polyglots were particularly skilled at learning multiple languages, correlating with higher WM, this difference may not exist between monoglots and biglots who tend to learn languages that they are immersed in (Papagno and Vallar 1995). This indicates that although individuals with very high WM capacity may predict multiple (more than 2) language acquisition, WM may not differ between those with one or two languages.

Congruent with this idea, Alptekin, Erçetin, and Özemer (2014) tested reading span for second language students who spoke both English and Turkish in both Turkish and English versions of the task. Their results demonstrated that although WM storage capacity appears independent of which language they were tested with, WM processing may differ leading to better comprehension in their native language (Alptekin, Erçetin, and Özemer 2014). This makes intuitive sense that an individual's first language would be processed better; however, it is subtleties like this that may perpetuate negative stereotypes against those who are attempting to learn English as a second language. Although these individuals may not differ in their base WM capacity, these performance interactions combined with the threat of self-confirming stereotypes may setup a situation where individuals' capabilities are falsely appraised.

Furthermore, one issue often not explored is that individuals may differ in how likely they are to succumb to these effects. Thus, this is also an important consideration as well (McConnell and Leibold 2001)

Hypotheses and Predictions:

From the background information and previous studies, we made three main hypotheses:

H1: There is no relationship between WM ability and language group (English or multiple languages).

H2: There is a relationship between self-reference of primes and the type of prime (negative or positive) given to an individual.

H3: There will be a relationship between how readily one forms stereotypes and these effects.

Specifically, we predicted that the individuals that spoke only English would not be different on a WM task from individuals who spoke English as well as a second language (2LS). We predicted that primes that were positive for those with a second language would enhance performance of this group. Additionally, because the primes were framed in terms of one group being enhanced it was assumed that this prime would have the opposite effect on the other group (W2L) therefore an interaction between type of prime and group was expected as well. Additionally, to test the hypothesis that degree of stereotype bias may predict the size of the manipulation, a multilingual implicit association test (IAT) was utilized as a non-reactive measure of stereotype bias (Greenwald, Nosek, and Banaji 2003). From this, we predicted that those with the highest affinity for forming stereotypes, indicated by their IAT score, would have the largest effect on working memory.

Methods

Participants.

Researchers collected a convenience sample of 69 participants over the age of 18 years ($M = 27.87$ years, $SD = 12.34$), recruited by word of mouth and posting in shared social media groups. Descriptive statistics for both the W2L and 2LS groups are presented in Table 1 and Table 1a (supplementary). Information on scores between these groups can be found in Table 2

Table 1. Means and standard deviations of participants' age and reported task difficulty.

	Age		Task Difficulty	
	m	sd	m	sd
English-Dominant Speakers (W2L)	22.65	5.95	5.95	1.79
Second-Language Speakers (2LS)	33.91	14.94	5.47	1.90
Total	27.87	12.34	5.72	1.85

Table 2. Mean, standard deviation and range of scores of measures used.

	W2L				2LS			
	min	max	m	sd	min	max	m	sd
WM	1.00	9.00	5.81	2.08	2.00	9.00	5.75	1.97
IAT	-0.39	1.27	0.54	0.41	-0.49	1.04	0.44	0.35

Materials

A laptop capable of running at least windows xp, Microsoft Powerpoint 2010 as well as Inquisit version 1.33 was required. Researchers used three paragraph manipulations (see supplementary material) to prime foreign language effect (positive, negative, or neutral) on working memory similar to (Hess *et al.* 2003). A power point slide

show was designed with previously validated sentences (Daneman and Carpenter 1980). The slide show was standardized so that it automatically progressed at 15-second intervals. Researchers designed an experimental script to standardize the power point procedure. Researchers also designed a paper handout for participants to write down words recalled during each trial which could then be scored after testing. Lastly, researchers created a multilingual implicit association test (IAT) to measure proclivity for formation of stereotypes (Draine, 1997).

Procedure

1.1 Reading span

Once participants consented to be in the study after reading information about it, they were then randomly assigned to one of the three language priming conditions: foreign languages having a positive effect on working memory, foreign languages having a negative effect on working memory, or foreign languages being unrelated to working memory. After participants were primed they underwent the working memory task. An experimental script was read to participants before beginning the trials. Participants then progressed through an automated power point presentation. During the first two trials, participants practiced the procedure of reading the sentence, memorizing the sentence, and recalling the specific word from the sentence. This was to ensure that participants understood what was expected of them during the study and so the researchers could correct any errors. Additionally, participants were told that if they were having trouble they could try repeating the word over in their head, and this was done to control for effects of differing strategies due to manipulation. Participants then underwent the working memory task.

During each trial, participants were shown a sentence for 15 seconds, which they would memorize, followed by a slide asking the participant to remember a specific word from the previous sentence (Ex. remember word 4). Participants would undergo this process 3 times before they were cued with recall however, so they had to maintain these words in working memory until then. Unlike Daneman and Carpenter's (1980) methodology, the words cued to be remembered differed from trial to trial. This was done to minimize effects of strategies, and enhance difficulty of the task. The words chosen were randomly selected prior to the task and were the same for each participant.

Participants recorded the words that they were asked to remember from each sentence in the trial on a paper handout. Researchers then scored the handouts using an answer key. Trial one, two, and three were all scored out of a total of 3 words recalled correctly for a sum total of nine. The number of words recalled was then totalled for all three trials and used as the dependent variable for working memory. At the end of the task, participants were asked to rate the perceived difficulty of the task.

1.2 Implicit association task

After completion of the working memory task, participants completed the IAT which is a measure of tendency to form unconscious judgements (Draine, 1997). The computer software used to conduct the IAT was Inquisit 1.33 developed by Millisecond Software. A template provided by the software company was used to create a novel version of the IAT assessing stereotypes towards multilingualism. Words used in the Good or Bad category were the same as the ones used in the original task implemented by Draine (1997). Words for the English or Foreign category were neutral in order to avoid confusion with the Good or Bad task. Thus participants were instructed to only respond Good or Bad if the words had emotional valence (i.e. disgusting not for words like orange or chair). Half the foreign language words were derived from the Roman (English) alphabet and half the words were derived from alphabets of different origins (Cyrillic, Greek, Arabic, etc.).

The IAT procedure began with an explanation of how to sort the categories. Left categories were sorted by pressing the 'E' key on the left side of the keyboard while the right category was selected by pressing the 'I' key on the right side of the keyboard. The category headings were given at the top corners of the screen. Participants were instructed to complete the task as fast as they could without worrying about errors. If a categorization made by the participant was wrong then a red 'x' would appear on the screen until the participant made the correct categorical assignment. (Inquisit, 2003; Draine, 1997).

Table 4. Sequence of blocks in the Multilingual IAT

Block	No. Trials	Function	Items Assigned to Left Key	Items Assigned to Right Key
1	8	Practice	English Words	Non-English Words
2	8	Practice	Good Words	Bad Words
3	16	Practice	English + Good	Non-English + Bad
4	32	Test	English + Good	Non-English + Bad
5	16	Practice	Non-English Words	English Words
6	16	Practice	Non-English + Good	English + Bad
7	32	Test	Non-English + Good	English + Bad

The sequence of blocks and number of trials that the participant encountered is provided in Table 4 for additional clarification. The order of the English or good pairing phase and the English or bad pairing phase were switched randomly to control for effects of task sequence. Before the

first task was given, participants were given one practice block of sorting English/Foreign words and another practice block sorting good and bad words. Then the first test block was given and both categories were sorted simultaneously. Categories were colour-coded so the participant was not confused about the categorization required. Another practice trial was given for the English/Foreign sorting condition to practice categorizing them on switched sides. Then they were given the second dual categorization block. The measure was administered twice for reliability purposes (Draine, 1997). After the IAT procedure was completed, participants were debriefed.

The IAT test results are D_2 values which are the result of algorithm developed by Greenwald, Nosek, and Banaji (2003) that is conducted by the software. The D_2 measure had the highest associated effect size and incorporates practice blocks 3 and 6 as well as error latencies, which have both been associated with increasing correlation of IAT measures as explicit preference measures on a variety of IAT tasks (Greenwald *et al.* 2003). D_2 is calculated by the following formula where $\bar{x}$ and sd denote means and standard deviations of response latencies of each block.

$$D_2 = \left[\frac{\bar{x} \text{ block 3} - \bar{x} \text{ block 6}}{sd \text{ blocks 3 and 6}} + \frac{\bar{x} \text{ block 4} - \bar{x} \text{ block 7}}{sd \text{ blocks 4 and 7}} \right] / 2$$

In our design specifically, the higher the score the more likely one would be to form a negative stereotype for those with a second language, and positive values the opposite.

Statistical Analyses

All experiments behavioural data between groups were analyzed by Analysis of Variance (ANOVA). Significant interactions or main effects were further analyzed by multiple comparisons using the Student Newman Keuls method ($\alpha = 0.05$). Statistical analyses were performed using SigmaStat (v. 3.5 for Windows, SPSS Inc) and SPSS version 21.

Results

A 3 X 2 ANOVA comparing condition (negative, neutral, positive) and language (2LS, WSL) for total number of correct words recalled on the WM task revealed a significant interaction between manipulation (neutral, negative, positive) and condition (2LS and WSL), $F(2, 63) = 3.89$, $p = 0.03$, partial $\eta^2 = .11$ (**Figure 1**). There was no main effect of condition $F(1, 63) = 0.01$, $p = 0.94$, partial $\eta^2 = .00$ and there was no main effect of manipulation $F(2, 63) = 0.35$, $p = 0.74$, partial $\eta^2 = .01$.

Multiple comparisons that tested the effect of language group for each prime manipulation revealed that while there was no significant difference between the language groups for the neutral condition $q = 0.59$, $p = 0.68$, there was a significant difference for the positive condition $q = 3.00$ $p = 0.04$, and it was approaching significance for the negative condition $q = 2.51$, $p = 0.08$. Additionally, multiple

Figure 1. Mean (SEM) number of correct responses for participants who either had a second language (2LS) or did not have a second language (W2L) and were exposed to either negative, neutral or positive primes about how having a second language would effect memory.

comparisons of the effect for condition within each language group revealed that these prime effects, as the interaction suggested are dependent on the group. Such that, for the 2LS group the postive prime produced scores that were higher, but not statistically different, than control $q = 1.22$, $p = 0.40$, and negative primes $q = 2.86$, $p = 0.12$. And for the W2L group the postive second language prime produced scores that were lower, but not statistically different, than control $q = 2.51$, $p = 0.08$ and negative primes $q = 2.67$, $p = 0.15$. Thus, postive primes about second languages produced higher scores for the 2LS group, but lower for the W2L group, but negative primes produced scores that were lower for the 2LS group and higher for the W2L group. However, it is important to stress here that statistical power was low due to only 9-12 subjects per conditon.

A 3 X 2 ANOVA comparing condition (negative, neutral, positive) X language (2LS, WSL) for IAT score revealed no significant interaction between manipulation (neutral, negative, positive) and condition (2LS and WSL), $F(2, 63) = 0.26$, $p = 0.78$, partial $\eta^2 = .01$. There was also no main effect of condition $F(1, 63) = 1.03$, $p = 0.31$, partial $\eta^2 = .02$ and there was no main effect of manipulation $F(2, 63) = 0.60$, $p = 0.55$, partial $\eta^2 = .02$. When evaluating IAT score, it became clear that IAT did not have a direct relationship with WM score, but may possibly mediate other variables. Specifically, in order to conclude that mediation has occurred one must show that there is a significant relationship between $A \rightarrow B$ and $B \rightarrow C$ such that when you look at the relationship of $A \rightarrow C$ controlling for B there is no significant relationship (Mallinckrodt *et al.* 2006). In our study it was found that IAT score significantly predicted perceived difficulty $F(1, 67) = 6.21$, $p = 0.02$. This accounted for approximately 9% of the variance, $R^2 = 0.09$, in difficulty such that the higher someone's IAT score the higher their

perceived difficulty. Additionally, difficulty rating was a significant predictor of WM scores while holding IAT constant $t(66) = 3.65$, $p = 0.001$ such that the lower difficulty ratings had higher WM total scores. However, IAT was not a significant predictor for total score while holding difficulty constant $t(66) = 0.82$, $p = 0.42$. Therefore, IAT score was concluded to have a relationship with WM score, as predicted, but this relationship was mediated by perceived difficulty.

Additionally, age also is a significant predictor of IAT score $F(1, 67) = 4.41$, $p = 0.04$ such that the older one is the lower the IAT score. However, age was not a significant predictor of WM score alone $F(1, 67) = 0.01$, $p = 0.95$. Thus overall it appears as if there was a significant effect of manipulation on working memory moderated by whether one is a 2LS or W2L. Additionally IAT score, which is predicted by age, may have a relationship with WM but it is mediated by difficulty.

Discussion

Self-referent prime effects on working memory

Canada is assumed to be a multi-cultural country. Thus, many assume that stereotypes against those with multiple languages should be minimally prevalent and minimally effective. However, studies seem to suggest that despite a multi-cultural image Canada is a English dominant society and those that primarily speak English may more easily obtain jobs, with higher pay grades (Chiswick and Miller 1995; Chiswick and Miller 2003). Two possibilities exist for why this occurs. One is that it is possible that individuals may not obtain these jobs because they don't have the English skills necessary to obtain these jobs, however it is also possible that stereotypes against these individuals may generate self-confirming biases towards this conclusion (Allen *et al.* 2009). Therefore to differentiate between these two hypothesis we exposed individuals to negative, positive and neutral primes about how having a second language effected their memory and then measured their performance on a working memory task.

In support of the interpretation that stereotypes drive this effect, not differences in ability, it was found that there was a significant interaction between type of prime and language group (2LS, W2L) and no significant main effect of language. This means that participants in the 2LS group performed better on the WM task when told that foreign languages were positively related to memory, however they performed poorly when told that foreign language was negatively related to memory. Mirroring this, individuals in the W2L condition performed better on the WM task when told that having a second language was negatively related to memory and performed poorly on the WM task when told that memory was positively related to having a second language. These results provide support that the effect of a prime is determined by how personally relevant it may be to an individual (Shih *et al.* 2002).

Additionally, the lack of a difference between 2LS and W2L groups on the neutral primes, or main effect between the two suggests these groups don't differ in baseline abilities. Although we did not perform tests prior to manipulations that would conclusively rule out this interpretation, the results strongly suggest against any underlying differences in WM storage capacity from the outset (Alptekin, Erçetin, and Özemir 2014).

Additionally, it is interesting to note that although age was included as a variable in our analysis it was not related to WM scores as some previous studies have found when manipulating age primes (Hess *et al.* 2003). This indicates that our manipulation was selective to the specific language stereotype activated instead of a causing generalized changes in performance from the positively or negatively worded statements, or a spurious effect due to age (Levy 1996; Hess *et al.* 2003; Hess, Hinson, and Statham 2004).

Mechanisms of threat and arousal

Our results are also in line with traditional views on stereotype threat and arousal such that primes which should have activated negative stereotypes about the participants groups tended to inhibit performance, whereas those primes that activated positive stereotypes enhanced performance (Ben-Zeev, Fein, and Inzlicht 2005; Schmader, Johns, and Forbes 2008). It has been hypothesized that the mechanisms for how stereotype arousal and threat work includes modifying: anxiety, physiological arousal, motivation, strategy type, or the incidence of disruptive thoughts (Beilock, Rydell, and McConnell 2007; Schmader, Johns, and Forbes 2008). But given our data set, it seemed most logical that anxiety may be the most relevant explanation as it has been suggested by (Steele and Aronson (1995).

We approached this question indirectly by building in several checks into our design. One way that this was done was to ensure that all participants were trained and given equal opportunity to learn the best strategy to perform the task. They were also told that their performance was anonymous and encouraged to do their best to potentially equalize motivation to perform the task. Therefore, combined these two methods were meant to eliminate the explanations of using alternate strategies under threat and arousal as well as differing motivation, with the aim of leaving anxiety and activation of unrelated thoughts (Beilock, Rydell, and McConnell 2007; Schmader, Johns, and Forbes 2008). However, it seems that both of these can be collapsed into one encompassing domain of anxiety and physiological arousal effects on performance triggered by both stereotype threat and arousal which has been suggested by (Osborne 2007).

The analysis of our perceived difficulty scores indicate that perceived difficulty was significantly related to WM scores which further supports an interpretation of physiological arousal and anxiety as mediating the effects from stereotypes (Ben-Zeev, Fein, and Inzlicht 2005). Therefore, although this relationship is correlational in

nature, and we did not measure baseline difficulty perceptions, it seems plausible that this is the driving mechanisms of our observed effects. Interestingly, as well, perceived difficulty scores had a significant relationship with measured IAT scores.

Interactions with IAT

IAT scores were originally taken to measure potential to form negative stereotypes for those with second languages (Greenwald, Nosek, and Banaji 2003). However upon analysis of the results, the relationship of IAT with WM appeared to depend on a mediated relationship with difficulty, but not age. Specifically, it was found that age had a significant relationship with IAT scores such that more positive (or stereotyped scores) tended to occur with younger individuals. As well perceived, difficulty had a significant relationship with IAT. Although in a multiple regression model with age, difficulty and IAT as predictors of WM, only difficulty had a relationship with IAT when holding the other two constant. Therefore, this suggests that perceived difficulty may mediate the relationship of IAT on WM such that those who are in a high state of arousal from negative stereotypes may tend to make more post-hoc external explanations for failure (i.e. it was too hard), compared to those who do well (Koch, Müller, and Sieverding 2008). To further extend these findings it may also be important to measure participants attitudes towards the task as if they interpret the task as a challenge over an impossible task. This may also reduce or amplify the facility or inhibitory effects on WM (Schmader and Johns 2003).

Strengths, limitations and future directions

In our design, we used a subtle instead of a blatant stereotype prime which has been found to have differing effects (Hess, Hinson, and Statham 2004). This provided support to address the discrepancy between subtle manipulations and their effectiveness suggesting that they may be effective to induce stereotype threat and subsequent effects (Shih *et al.* 2002; Hess, Hinson, and Statham 2004). This is important because priming plays a major role in the formation of stereotypes (Bargh and Williams 2006) and in the society that we live in today, overtly stereotyping is not acceptable practice. One other important point about subtle stereotypes is that individuals may not be as consciously aware of them and may therefore act in stereotype confirming ways without conscious recognition and therefore are much less able cease such behaviour (Chen and Bargh 1997). A future direction may be to test the threshold level of conscious awareness of stereotypes and ways to increase this in order to aide in people's ability combat them.

It is important to note that this study was not without limitations as researchers utilized a convenience sample, and were very underpowered for group effects (13%). Therefore, if anything, this may be a underestimation of these effects. As well this experiment failed to measure baseline IAT scores and aforementioned difficulty perceptions therefore

these may also be added to future studies to improve interpretation of results.

Conclusions

Taking into account the limitations of the results, it was still shown that WM capacity does not differ between language groups. However, when these same groups are given self-referent positive and negative primes these can induce stereotype threat and arousal which have clear effects on WM. Therefore altogether, these results suggest that immigrants with multiple languages may be subject to stereotypes and prevailing ideas about second languages. It should be important to note the consequences of these results, as it must specifically be eliminated as a judgement factor through both policy changes and sufficient education on its potential effects, especially if we Canadians wish to call ourselves a multicultural nation.

References.

- Allen TJ, Sherman JW, Conrey FR, Stroessner SJ. 2009. Stereotype strength and attentional bias: Preference for confirming versus disconfirming information depends on processing capacity. *J. Exp. Soc. Psychol.* 45:1081–1087.
- Allport, Gordon W. (1956). *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.
- Alptekin C, Erçetin G, Özemir O. 2014. Effects of Variations in Reading Span Task Design on the Relationship Between Working Memory Capacity and Second Language Reading. *Mod. Lang. J.* 98:536–552.
- Baddeley A. 2003. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4:829–839.
- Bargh J a, Williams EL. 2006. The Automaticity of Social Life. *Curr. Dir. Psychol. Sci.* 15:1–4.
- Bargh JA, Chen M, Burrows L. 1996. Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *J. Pers. Soc. Psychol.* 71:230–244.
- Beilock SL, Rydell RJ, McConnell AR. 2007. Stereotype threat and working memory: mechanisms, alleviation, and spillover. *J. Exp. Psychol. Gen.* 136:256–76.
- Ben-Zeev T, Fein S, Inzlicht M. 2005. Arousal and stereotype threat. *J. Exp. Soc. Psychol.* 41:174–181.
- Boyd M, Vickers M. 2000. 100 years of immigration in Canada. *Can. Soc. Trends*:2–13.
- Campion M a, Pursell ED, Brown BK. 1988. Structured interviewing: Raising the psychometric properties of the employment interview. *Pers. Psychol.* 41:25–42.

- Carretti B, Borella E, Cornoldi C, De Beni R. 2009. Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learn. Individ. Differ.* 19:246–251.
- Chen M, Bargh J a. 1997. Nonconscious Behavioral Confirmation Processes: The Self-Fulfilling Consequences of Automatic Stereotype Activation. *J. Exp. Soc. Psychol.* 33:541–560.
- Chiswick BR, Miller PW. 1995. The Endogeneity between Language and Earnings: International Analyses. *J. Labor Econ.* 13:246–288.
- Chiswick BR, Miller PW. 2003. The complementarity of language and other human capital: Immigrant earnings in Canada. *Econ. Educ. Rev.* 22:469–480.
- Colom R, Flores-Mendoza C, Rebollo I. 2003. Working memory and intelligence. *Pers. Individ. Dif.* 34:33–39.
- Daneman M, Carpenter PA. 1980. Individual differences in working memory and reading. *J. Verbal Learning Verbal Behav.* 19:450–466.
- Draine, S. C. (1997). Analytic limitations of unconscious language processing. Doctoral Dissertation, University of Washington.
- Devine PG. 1989. Stereotypes and Prejudice: Their Automatic and Controlled Components. 56:5–18.
- Greenwald AG, Nosek B a, Banaji MR. 2003. Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85:197–216.
- Hess TM, Auman C, Colcombe SJ, Rahhal T a. 2003. The Impact of Stereotype Threat on Age Differences in Memory Performance. *Journals Gerontol. Ser. B Psychol. Sci. Soc. Sci.* 58:3–11.
- Hess TM, Hinson JT, Statham J a. 2004. Explicit and implicit stereotype activation effects on memory: do age and awareness moderate the impact of priming? *Psychol. Aging* 19:495–505.
- Inquisit 1.33 [Computer software]. (2003). Seattle, WA: Millisecond Software.
- Judd CM, Park B. 1993. Definition and assessment of accuracy in social stereotypes. *Psychol. Rev.* 100:109–128.
- Koch SC, Müller SM, Sieverding M. 2008. Women and computers. Effects of stereotype threat on attribution of failure. *Comput. Educ.* 51:1795–1803.
- Lépine R Ile, Parrouillet P, Camos V. 2012. What makes working memory spans so predictive of high-level cognition? *Psychon. Bull. Rev.* 12:165–170.
- Levy B. 1996. Improving memory in old age through implicit self-stereotyping. *J. Pers. Soc. Psychol.* 71:1092–1107.
- Macrae CN, Milne AB, Bodenhausen G V. 1994. Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *J. Pers. Soc. Psychol.* 66:37–47.
- Mady C. 2012. Official language bilingualism to the exclusion of multilingualism: immigrant student perspectives on French as a second official language in “English-dominant” Canada. *Lang. Intercult. Commun.* 12:74–89.
- McConnell AR, Leibold JM. 2001. Relations among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes. *J. Exp. Soc. Psychol.* 37:435–442.
- Miller, G. A., Galanter, E., and Pribram, K. H., (1960). Plans and the structure of behavior. New York: Holt, Rinehart and Winston.
- Osborne JW. 2007. Linking Stereotype Threat and Anxiety. *Educ. Psychol.* 27:135–154.
- Papagno C, Vallar G. 1995. Verbal Short-term Memory and Vocabulary Learning in Polyglots. *Q. J. Exp. Psychol. Sect. A* 48:98–107.
- Rydell RJ, McConnell AR, Beilock SL. 2009. Multiple social identities and stereotype threat: imbalance, accessibility, and working memory. *J. Pers. Soc. Psychol.* 96:949–66.
- Schmader T, Johns M. 2003. Converging evidence that stereotype threat reduces working memory capacity. *J. Pers. Soc. Psychol.* 85:440–452.
- Schmader T, Johns M, Forbes C. 2008. An integrated process model of stereotype threat effects on performance. *Psychol. Rev.* 115:336–356.
- Shapiro JR, Neuberg SL. 2007. From stereotype threat to stereotype threats: implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Pers. Soc. Psychol. Rev.* 11:107–130.
- Shih M, Ambady N, Richeson J a., Fujita K, Gray HM. 2002. Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation. *J. Pers. Soc. Psychol.* 83:638–647.

Steele CM, Aronson J. 1995. Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* 69:797–811.

Steele CM, Aronson J. 1995. Stereotype Threat and the Intellectual Test Performance of African Americans. 69:797–811.

Supplementary Information

Table 1a. Demographic Information: Numbers and Percentages

	W2L		2LS		Total	
	n	%	n	%	n	%
Gender						
Male	17	45.9	13	40.6	30	43.0
Female	20	54.1	19	59.4	39	56.0
Ethnicity						
Caucasian	30	81.1	19	51.4	49	71.0
African Canadian	2	5.4	1	2.7	3	4.3
Hispanic/Latino	3	8.1	1	2.7	4	5.8
South Asian	1	2.7	9	24.3		
East Asian	1	2.7	1	2.7	2	2.9
Oceanic	0	0	1	2.7	1	1.4
Nationality						
Canadian	36	97.3	19	51.4	49	71.0
Italian	0	0	8	21.6	8	11.6
Polish	0	0	3	8.1	3	4.3
Great Britain	1	2.7	0	0	1	1.4
Germany	0	0	1	2.7	1	1.4
Iranian	0	0	1	2.7	1	1.4

Figure 2. The graph displays the time course of the experiment. During informed consent many descriptive statistics were acquired. Although not assigned participants either had a second language or did not. Therefore the paragraph manipulations were self-referent for those with a second language and positive for paragraphs which said having a second language was beneficial, and similarly negative for those without one. Reciprocally they were self-referent and negative for the negative prime about second languages for those with a second language, but positive for those without. Neutral was not in either direction. Following this a working memory task was performed. Immediately afterward a multilingual implicit association test measuring proclivity for negative stereotypes against those with second languages was performed. Following both tasks participants were debriefed.

Instructions for 3-Sentence Condition:

"In the upcoming slides, you will undergo the following steps:

1. You will see a sentence. You will have 15 seconds to read and memorize this sentence.
2. You will then see a slide indicating the number of word, which you must remember from the sentence that you just saw. For example, if the number is "4", then you must remember the 4th word in the sentence that you just saw.
3. You will undergo these steps 3 times, before you will be asked to recall them.
4. You will then be asked to type out each of the 3 words that you were asked to remember from each sentence.
5. You will undergo this process 3 times.

Practice Trial

Using the instructions just given to you, please complete a practice trial before beginning the study.

- [if practice trial is completed successfully] – Great! Let's get started.
- [if practice trial is completed unsuccessfully] – That's okay. Next time, perhaps try saying the sentence to yourself in your head, so that when you are asked to remember a specific word, it will be easier for you to remember! Let's try one more time.
- [if second practice trial is completed successfully] – Great! Let's get started.
- [if second practice trial is completed unsuccessfully] – (Coach participant based on their specific neW2L).
- Experimental script for conditions

Experiment Answer Sheet

Gender: _____ Age: _____

Practice 1

Practice 2

Test Set 1

Test Set 2

Test Set 3

How difficult did you find this task? (Scale: 1-10): _____

[This part is for the experimenter only]

Trial 1 Score _____

Trial 2 Score _____

Trial 3 Score _____

Trial handouts

Second languages appear to impair people's ability to solve problems

There has been some speculation over the advantage of multiple language ability on problem solving. This is due to linguistic relativity by Sapir and Whorf (1921, 1939), which states that different languages influence cognitive processing differently, causing the formation of different worldviews. Having different perspectives of a problem may help with problem solving. Francis (2008) found that students that were fluent in two or more languages often showed poor performance on exams compared to their monolingual counterparts. This can be explained by cognitive load, where having multiple perspectives may tax working memory and effectively impair the ability to manipulate and process components of a problem (Sweller, 1988; Reisberg, 2009).

Second languages appear to facilitate people's ability to solve problems

There has been some speculation over the advantage of multiple language ability on problem solving. This is due to linguistic relativity by Sapir and Whorf (1921, 1939), which states that different languages influence cognitive processing differently, causing the formation of different worldviews. Having different perspectives of a problem may help with problem solving. Francis (2008) found that students that were fluent in two or more languages often showed significantly better performance on exams compared to their monolingual counterparts. This supports the theory that linguistic relativism allows multi-linguals to look at a problem from multiple perspectives (Bialystok and Majumder, 1998).

Second languages appear to be unrelated to people's ability to solve problems

There has been some speculation over the advantage of multiple language ability on problem solving. This is due to linguistic relativity by Sapir and Whorf (1921, 1939), which states that different languages influence cognitive processing differently, causing the formation of different worldviews. Having different perspectives of a problem may help with problem solving. Francis (2008) found that students that were fluent in two or more languages often showed performance that was no different to their monolingual counterparts. This may be due to one culture dominating over others, which forces other worldviews to be lost over the one of the dominate culture, meaning there is no advantage to problem solving over peers (Newell and Simon, 1972).