Language and transient emotional states affect implicit cultural bias
Ellis, Ceri; Hadden, Lowri; Jones, Manon Wyn

Quarterly Journal of Experimental Psychology

DOI: 10.1177/1747021818792412

Published: 01/05/2019

Peer reviewed version

Cyswllt i'r cyhoeddiad / Link to publication

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Hawliau Cyffredinol / General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Language and transient emotional states affect implicit cultural bias

*Running head: Bilinguals in the mood for culture*

Ceri Ellis¹, Lowri Hadden², Manon Wyn Jones³

¹Division of Neuroscience & Experimental Psychology, University of Manchester, UK
²School of Psychology, Cardiff University, UK. HaddenL@cardiff.ac.uk
³School of Psychology, Bangor University, UK. Manon.jones@bangor.ac.uk

Address for correspondence:
Dr Ceri Ellis,
Division of Neuroscience & Experimental Psychology,
Faculty of Biology, Medicine & Health,
Room 3.316 Jean McFarlane Building,
University of Manchester,
Oxford Rd,
Manchester.
M13 9PL
Tel: 0161 306 7964
Email: ceri.ellis@manchester.ac.uk

**Word Count:** 3,125 (excluding references)

**Acknowledgments**
This work was supported by the Coleg Cymraeg Cenedlaethol [C. Ellis and M. W. Jones; www.colegcymraeg.ac.uk]. We thank Catrin Nickson, Naomi Morris, Jennifer Barlow and Emma Wyn Owen for their assistance with data collection. We also thank Rob Ward and Guillaume Thierry for their helpful comments on the manuscript.

**Disclosure of interest**
The authors report no conflicts of interest.
Abstract

Bilinguals react to cultural information in a language-dependent fashion; but it is unknown whether this is influenced by the individual’s emotional state. Here, we show that induced mood states increase cultural bias – measured using the Implicit Association Test (IAT) – but this effect occurs asymmetrically across languages. In the native language, bilinguals show a strong cultural bias, which is not influenced by mood. But in the non-native language, a relatively low cultural bias significantly increases as a function of a positive or negative mood. Our findings suggest that the native language promotes an inherent cultural bias, which is impervious to fluctuations in the bilingual’s mood state. In the second language, however, bilinguals are culturally impartial, unless they are in a heightened mood state.

Keywords: language, culture, emotion, bias, Implicit Association Test
Introduction

Language affects how we perceive, process and judge information that pertains to our native and non-native culture (Briley, Morris, & Simonson, 2005). In bilinguals, the native language has been shown to enhance cultural biases (Danziger & Ward, 2010; Ellis et al., 2015; Ogunnaike, Dunham, & Banaji, 2010), which likely reflects more vivid recollection of pertinent cultural memories and norms (Ellis, Thierry, Vaughan-Evans, & Jones, 2017; Marian & Kaushanskaya, 2004; Marian & Neisser, 2000; Schrauf, Pavlenko, & Dewaele, 2003). Here, we question the assumption that cultural biases are fixed attributes of the bilingual’s languages, and examine whether such prejudicial attitudes can instead fluctuate according to a bilingual’s mood.

Several studies now show that mood states\(^1\) arising from external sources unrelated to culture or group membership influence how we automatically judge and stereotype others. Positive mood states such as happiness induce an intuitive, heuristic style of cognitive processing, which promotes greater cultural bias and reliance on stereotypes (Bodenhausen, Mussweiler, Gabriel & Moreno, 2001; DeSteno, Dasgupta, Bartlett & Cajdric, 2004; Cottrell & Neuberg, 2005; Tiedens & Linton, 2001). Similar effects are found for negative mood states, but only when the state is associated with intergroup conflict (e.g., Dasgupta, DeSteno, Williams, & Hunsinger, 2009). Often coupled with increased autonomic arousal, these fast, automatic processing styles are deemed adaptive, allowing quick responses to environmental stimuli (Clark & Fiske, 2014; van Kleef & Fischer, 2016). In bilinguals, recent work suggests that the native language elicits this heuristic approach for making judgements and decisions, whilst reduced emotional resonance in the foreign language attenuates intuitive thinking and moral judgements (Keysar, Hayakawa, & An, 2012; Costa, Foucart, Arnon, Aparici,\(^1\)

---

\(^1\) For clarity, we refer to ‘mood’ as an incidental, transient emotional response to an environmental stimulus (see Cox, 2002; Damasio, 1994; Keltner & Gross, 1999; Stirling & Kerr, 2006).
However, reduced emotional resonance in these studies is necessarily tied to the linguistic expressions used in the experimental task, and whilst compelling, these findings tell us little of the effect of mood when disentangled from lexical factors.

Our aim in this study is to ascertain whether cultural biases can be modulated by an elevated mood state *per se*: That is, one that is elicited independently of the experimental task, and not bound to early experiences and recollections of words in one language over another. To this end, we used the Implicit Association Test (IAT; e.g., Greenwald, Poehlman, Uhlmann, & Banaji, 2009) to measure implicit cultural biases in Welsh-English bilinguals who identified as culturally Welsh. An adapted version of the IAT (Danziger & Ward, 2010; Ogunnaike, Dunham & Banaji, 2010) poses a culturally-relevant categorisation task, in which faster and more accurate responses to related concepts signal an automatic bias. Prior to completing the IAT, we manipulated the bilinguals’ mood state: Audio-video stimuli with no linguistic, cultural or socially biasing components were pre-selected on the basis of ‘positive’, ‘negative’ or ‘neutral’ mood responses (Positive and Negative Affect Schedule, PANAS; Watson, Clark, & Tellegen, 1988). Importantly, ‘negative’ stimuli tended to elicit agitated mood states associated with fear and anxiety (Diener, 2009), which are also associated with intergroup conflict (Fiske, 2002; Halperi et al., 2012). Overall, we predicted stronger implicit cultural bias in the native compared with the non-native language, as per Ogunnaike et al. (2010) and Danziger and Ward (2010). Given that the native language is also more susceptible to emotional resonance (e.g., Keysar et al., 2012; Costa, Foucart, Arnon, Aparici & Apesteguia, 2014; Dewaele, 2004), we also predicted maximal cultural bias following an elevated positive or negative mood state in L1.
Method

Participants

Our sample comprised twenty-nine participants, all of whom identified as native L1 Welsh, with at least one Welsh-speaking parent. English was acquired at an early age ($M_{years} = 3.83$, $SD = 2.21$). A further two participants were excluded for having outlier scores according to the exclusion procedures (Greenwald, Nosek, & Banaji, 2003), and one participant presented with incomplete data.

Reading, writing, speaking and comprehension scores were high in both Welsh ($M_{Grand} = 9.03$, $SD = 1.44$), and English ($M_{Grand} = 8.54$, $SD = 1.46$), as measured by a composite self-reported language proficiency index on a scale of 1 (not literate) to 10 (very literate). Participants also reported more daily use of Welsh ($M = 67.69\%$, $SD = 17.67\%$) than English ($M = 35.41\%$, $SD = 16.67\%$).

Participants took part in the experiment if they identified as culturally Welsh on the Multi-group Ethnic Identity Measure (MEIM; Phinney, 1992; Roberts et al., 1999). Participants then indicated, on two subscales ranging from 1 (disagreement with cultural statement) to 4 (strong agreement with cultural statement): 1) degree of engagement with the Welsh culture (‘Identity Search’; $M = 2.98$, $SD = 0.54$, $\alpha = .69$), and 2) sense of cultural pride (‘Affiliation and Belonging’; $M = 3.58$, $SD = 0.31$, $\alpha = .74$). Scores indicated above average cultural affinity on both subscales. The School of Psychology ethics committee at Bangor University granted ethical approval and all participants gave informed consent.

Stimuli and Procedure

---

2 A power analysis performed in G*Power indicated a conservative sample size of $n=19$ (repeated measures ANOVA with 6 levels; power = .80; alpha = .05; effect size = .25 (Faul, Erdfelder, Lang & Buchner, 2007; Faul, Erdfelder, Buchner & Lang, 2009), consistent with previous IAT studies that typically report an effect of language at $d = .81$, e.g., Ogunnaike et al., 2010).
The Implicit Association Test (IAT)

Two identical versions of the cultural attitude IAT were used (Danziger & Ward, 2010), one in Welsh, and the other in English; each implemented with Inquisit 4.0 Millisecond software.

Participants categorised Welsh (Branwen, Cerys, Ieuan, Dafydd, Rhys) and English names (Alice, Mary, Henry, John, David), as well as “good” trait words (Welsh: da, clyfar, glan, hapus, cryf; English: good, smart, clean, happy, strong), and “bad” trait words (Welsh: drwg, twp, budr, gwan, blin; English: bad, dumb, dirty, weak, angry), via a binary choice keyboard press (see Table 1). Participants first completed two training blocks in which they correctly categorised names and traits separately. Then, they completed two combined names/traits blocks (see Fig. 1). Congruent combined blocks contained trials in which names and traits were compatible, considering the anticipated cultural bias of the participants (i.e., Welsh-good; English-bad). Following a reverse names training block, two incongruent combined blocks were presented in which names and traits were incompatible (i.e., Welsh-bad; English-good.). Names and traits used in the Welsh and the English IATs were matched for word frequency, valence and arousal (cf. Hadden, 2014).

Table 1. Block sequence in the Welsh/English language IAT

<table>
<thead>
<tr>
<th>Block</th>
<th>N trials</th>
<th>Task</th>
<th>Response key</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>Name categorisation</td>
<td>Welsh</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>Trait categorisation</td>
<td>good</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>Congruent combined block</td>
<td>Welsh, good</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>Congruent combined block</td>
<td>Welsh, good</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>Reverse name categorisation</td>
<td>English</td>
</tr>
<tr>
<td>6</td>
<td>25</td>
<td>Incongruent combined block</td>
<td>English, good</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>Incongruent combined block</td>
<td>English, good</td>
</tr>
</tbody>
</table>

Note. Congruency of critical combined blocks (i.e., blocks 3-4 and 6-7) and language of the IAT were counterbalanced between participants. Data obtained from bold-highlighted blocks served as ‘test’ critical blocks in the analyses (see ‘IAT data analysis’ section below).
Mood manipulation

Short film clips were used to elicit a specific mood state (Rottenberg, Ray, & Gross, 2007; Hewig et al., 2005; Coan & Allen, 2007; Egidi & Nusbaum, 2012). Films comprised short clips from movies or stock videos, based on a format used by Rottenberg, Ray and Gross (2007). Two ‘positive’ films depicted pleasant scenes (cf. Table 2); two ‘negative’ films depicted unpleasant scenes, and two ‘neutral’ films depicted scenarios typically devoid of any particularly strong emotion (Gross & Levenson, 1995; Stanton, Reeck, Huettel, & LaBar, 2014; Wang, LaBar, & McCarthy, 2006). Each film was presented with music reflecting the intended valence (positive/negative/neutral), and all types of linguistic cues (characters talking, written signs etc.) were deliberately avoided.
Table 2. Summary of videos used in the experiment

<table>
<thead>
<tr>
<th>Version</th>
<th>Duration (min)</th>
<th>Music</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive 1</td>
<td>03:26</td>
<td>“One Day Like This” by Elbow.</td>
<td>A new-born baby smiling; People smiling; People dancing in different contexts; Family enjoying a barbeque; Friends watching a sunset.</td>
</tr>
<tr>
<td>Positive 2</td>
<td>03:29</td>
<td>“Stars” by Basspartout.</td>
<td>Exercising; Nature; Sunshine; A young couple in love; People reuniting at an airport.</td>
</tr>
<tr>
<td>Neutral 1</td>
<td>03:33</td>
<td>“Opening” by Andrea Guerra.</td>
<td>Office work; Reading; Microwave countdown; Passengers waiting at a train station platform; Cleaning.</td>
</tr>
<tr>
<td>Neutral 2</td>
<td>03:34</td>
<td>“Dead Already” by Thomas Newman.</td>
<td>A tube platform; Traffic; A shopping trolley being pushed around a store; Man watching the clock to leave work.</td>
</tr>
<tr>
<td>Negative 1</td>
<td>03:26</td>
<td>“Cold” by Jorge Mendez.</td>
<td>Children fleeing from war; Homelessness; Domestic violence; Robber committing crime; Failed attempt at escape being watched by woman.</td>
</tr>
<tr>
<td>Negative 2</td>
<td>03:29</td>
<td>“Funeral For a Tree” by John Powell.</td>
<td>A destroyed city; Cancer patient losing hair; A retired soldier visiting graves of war heroes.</td>
</tr>
</tbody>
</table>

Experiment procedure

The IAT experiment was conducted in two separate 45-minute testing sessions. Both sessions were identical (one in Welsh; one in English). Session order was counterbalanced between participants, and for each participant, Welsh/English sessions were conducted a week apart.

Each session began with the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988; presented online on Bristol Online Surveys) to establish baseline affect scores. The PANAS consists of ten positive (e.g. ‘enthusiastic’) and 10 negative (e.g. ‘jittery’) adjectives, which participants rate – on a Likert scale from 1-5 (1= very slightly to 5= extremely) – as to the adjective’s congruency with their current emotional state. None of our participants displayed baseline abnormal affect scores worthy of data exclusion, cf. Crawford & Henry, 2004. The procedure then followed three sequences, illustrated in Figure 2.
Figure 2. Schematic of experimental procedure. Note. In each sequence, a film was followed by presentation of the IAT, followed by presentation of the PANAS.

A ‘high mood’ film was always presented in the first and last sequence, interspersed with the neutral sequence. At the end of each sequence, a short break allowed completion of the language history questionnaire and the MEIM (Phinney, 1992). Short filler tasks (grammar and spelling worksheets, consistent with the test language) also reduced the likelihood that the affect elicited in one sequence would spill over into another sequence, and de-emphasised the overall role of mood and evaluation of mood.

IAT data analysis

Data from congruent and incongruent blocks were analysed using the improved scoring algorithm (Greenwald, Nosek & Banaji, 2003). The ‘IAT effect’ comprised a proxy $D$ score, herein referred to as the ‘IAT-$D$’, calculated via the following steps: (1) Participant exclusion if more than 10% of trials had response latencies < 300 ms; (2) Calculation of the “inclusive” standard deviation for response latencies in Blocks 3 and 6 (‘practice’ critical blocks) and in Blocks 4 and 7 (‘test’ critical blocks); (3) Calculation of mean response latencies for each congruent/incongruent block (Blocks 3, 4, 6 and 7); (4) Calculation of mean differences between the ‘practice’ and ‘test’ blocks ($\text{Mean}_{\text{Block 6}} - \text{Mean}_{\text{Block 3}}$ and $\text{Mean}_{\text{Block 7}} - \text{Mean}_{\text{Block 4}}$); (5) Mean difference scores were then divided by the “inclusive” standard deviation; (6)
Calculation of the IAT-D score from the equal-weight average of these scores. Response latencies for each participant per block were trimmed to within 2 SD following step 1 (Danziger & Ward, 2010). Since congruent responses are subtracted from incongruent responses, a larger IAT-D score here indicates a more favorable implicit attitude towards the cultural in-group (Welsh), and a less favourable attitude towards the cultural out-group (English). The IAT-D score also indicates the strength of the effect size: thus, values of .15, .35, and .60 correspond to small, medium, and large effects, respectively (Rudman, 2011).

Results

Following the data trimming procedure outlined above, 93% of the data was included in the analysis. A within-subjects repeated measures ANOVA was conducted with Language (Welsh, English) and Mood (positive, neutral, negative) as independent variables.

A main effect of Language showed that, when the IAT was administered in Welsh, participants showed a greater overall cultural bias than when it was presented in English ($F_{(1, 28)} = 7.04, p = .013, \eta^2_p = .201$; Fig. 3). Mood did not affect the overall strength of cultural bias ($F_{(2, 56)} = 0.83, p > .250, \eta^2_p = .029$), but there was a significant Language*Mood interaction ($F_{(2, 56)} = 5.55, p = .006, \eta^2_p = .165$): A post-hoc analysis, split by language, showed that for the English language IAT, mood had a significant effect on cultural bias ($F_{(2, 56)} = 4.34, p = .018, \eta^2_p = .134$), such that both the positive and negative mood conditions elicited a stronger bias relative to the neutral condition ($t_{(28)} = 2.40, p = .023$; $t_{(28)} = 2.45, p = .021$), but were nevertheless not significantly different from each other ($t_{(28)} = 0.28, p > .250$). No such differences in mood emerged in the Welsh language IAT, however ($F_{(2, 56)} = 1.65, p = .200, \eta^2_p = .056$).
In order to ascertain actual mood changes during the experiment sessions, global affect scores (positive minus negative sub-scales; Fig. 4) derived from PANAS were submitted to an ANOVA examining differences as a function of Language and Mood. Global affect scores were not affected by the language of the test session ($F_{(1, 28)} = 0.07, p > .250, \eta_p^2 = .002$), but they were affected by mood condition ($F_{(2, 56)} = 24.91, p < .001, \eta_p^2 = .471$): As expected, the positive mood condition ($M = 17.47, SE = 1.49$) elicited a larger overall positive affect score compared to neutral ($M = 12.16, SE = 1.71; p < .001$) and negative ($M = 7.19, SE = 2.11; p < .001$) mood conditions. Neutral and negative scores also significantly differed ($p = .002$). There was no significant interaction between language and mood condition ($F_{(2, 56)} = 0.52, p > .250, \eta_p^2 = .018$). This analysis confirms that participants’ mood was altered by the manipulation in the expected way, and that alterations in mood were equivalent in both languages.
Figure 4. Global affect scores by Language and Mood. Positive values indicate a greater overall positive affect score. Error bars represent SE.

Discussion

Using a measure of implicit attitudes, we show that whilst bilinguals’ cultural biases are inherently stronger in the native language, transient mood states can also enhance bias in the second language. In Welsh (L1), bilinguals showed a consistently strong cultural bias, which was already present in the neutral mood condition, and was not further modulated by a transient emotional state. In English (L2), bilinguals showed a weaker bias at baseline, which was made stronger as a result of either a positive or negative mood state. Our findings therefore reveal bilingual cultural bias in the native language but not in the second language, when these individuals are in a neutral mood. Cultural bias emerges in both languages as a result of elevated mood. Crucially, this effect can be attributed to affect per se, since the mood manipulation did not comprise any linguistic or cultural associations.

Recent studies have shown that the native language promotes cultural stereotyping in favour of the in-group (Danziger & Ward, 2010; Ogunnaike et al., 2010). Our results both support these findings and provide a crucial new layer of information: that cultural bias in the
bilingual’s second language is malleable, depending on the affective state of the individual. One mechanistic account for the asymmetric language-cultural link is in terms of the strength and saliency of culturally-relevant episodic memories when consolidated via the native language (Danziger & Ward, 2010; Marian & Kaushanskaya, 2004; Marian & Neisser, 2000; Schrauf, Pavlenko, & Dewaele, 2003).

On the basis of our findings, we posit that the intrinsic native language-culture link gears the bilingual towards a more intuitive, heuristic processing style, even in the absence of a high affect state. However, the bilingual’s second language is inherently less intuitive (cf. Costa et al., 2014; Ellis et al., 2017; Hadjichristidis, Geipel, & Surian, 2018; Keysar et al., 2012; Gao, Zika, Rogers, & Thierry, 2015; also see Pavlenko, 2012 for a review), and becomes heuristic only as a result of increased affect (e.g., Dewaele, 2004; Hadjichristidis, Geipel, & Savadori, 2015). Importantly, our findings indicate that elevated mood – disassociated from the emotional and episodic connotations of words in either language – can override the foreign language effect found in previous work (Costa et al., 2014; Keysar et al., 2012), and promotes a similarly heuristic processing style in both of the bilinguals’ languages.

Both positive and negative mood manipulations exerted a similar effect in this study, consistent with previous findings showing an increase in automatic cultural bias as a function of positive emotions such as happiness or specific negative emotions (Dasgupta et al., 2009; Lambert, Khan, Lickel, & Fricke, 1997; DeSteno et al., 2004; Park & Banaji, 2000). In the current study, ‘negative’ mood was associated with fear, anxiety and a general feeling of agitation (as measured by the PANAS; Diener, 2009), states that are commonly associated with group conflict (Fiske, 2002; Halperin et al., 2012). We note that other, typically low-arousal negative mood states, such as sadness, have been found to promote an analytical processing style, leading to reduced bias (Bodenhausen et al., 2001; Bless & Fiedler, 2006).
Our findings are consistent with previous work in which social stereotyping is associated with high arousal states, which likely promote fast, automatic ‘approach or avoid’ reactions to environmental stimuli (cf. Bradley, Codispoti, Cuthbert, & Lang, 2001; Panksepp, 2003; Panksepp & Biven, 2012), but with the constraint that these states influence intergroup judgements only when they are also related to intergroup interactions (e.g., DeSteno et al., 2004; Dasgupta, DeSteno, Bartlett & Cajdric, 2009).

The current findings show, for the first time, that bilinguals’ cultural biases are not only inherently stronger in the native language, but that in the second language, the level of bias is flexible and can be increased with temporarily heightened affect. Given that automatic cultural prejudices are a key factor in international relations, both historically and in projecting future policies, these findings present with the intriguing possibility that cultural bias may be moderated in the second language by diffusing affect.
References


