Supplementary Online Materials
I bet I'll remember "biochemistry" – Meta-memory as a function of lexical features and language background

Grace C. Lin[1], Rachel N. Smith[1], Chelsea Parlett-Pelleriti[2], Masha R. Jones[1], and Susanne M. Jaeggi[1]

1 – University of California, Irvine, School of Education
2 – Chapman University Schmid College of Science and Technology
graceCL@uci.edu

Table S1. Sample Size- Language Group and Version

	EO	EDB	ENDB	Total
Randomized	22	24	20	66
Non-Randomized	23	26	20	69
Total	45	50	40	135

Note. EO = English only. EDB = English-dominant bilingual. ENDB = English non-dominant bilingual. Among the EDB participants, 18 (36%) speak Spanish as their other language, 14 (28%) speak Chinese (4 Mandarin, 5 Cantonese, and 5 unspecified Chinese language), and five (10%) speak Tagalog. The remaining languages include Korean (3 students), Farsi (2), Arabic (2), Vietnamese (2), Gujarati (1), Telugu (1), Tigrinya (1), and Hindi (1). Among the ENDB participants, 34 (85%) speak Chinese (29 unspecified Chinese, 4 Mandarin, 1 Cantonese) as their dominant language, and the remaining dominant languages of ENDB participants are Arabic (2 students), Farsi (2), Korean (1), and Vietnamese (1).

Formation of the Language Groups (EO, EDB, and ENDB)

Participants filled out a Language Background questionnaire modified from Torres (2012). The survey includes questions such as the country in which participants were born, whether they speak a non-English language, the age at which participants started learning English, and the language most used in different periods of their life (0-5 years old, 6-12, 13-18, and over 18 years old). Participants who indicated that they speak another language were first grouped into the "bilingual" category. Then, we used their responses to other questions, e.g., the age at which they first started learning English and the language they most used at different periods of their lives, to determine if the participant should be classified as EDB or ENDB. For example, all ENDB participants were born in a non-English speaking country, did not attend primary school in the United States, and indicated their most used language as something other than English during the ages of 0-5.

Table S2. Word Lists and Valence

List 1	List 2	List 3	List 4	List 5
cigarette	trouble	cruelty	cheat	stress
meanness	crack	deception	selfishness	intolerance
annoyance	epidemic	nervousness	bully	anger
hate	violence	death	fear	lie
welder	gallon	antelope	gusto	temperature
goose	time	bowl	step	ounce
diver	people	brush	location	prince
class	biochemistry	solid	glass	table
epic	holiday	appreciation	value	magic
delight	love	idea	laugh	amusement
creativity	meal	dance	child	comfort
party	swimming	refund	peace	sandwich

Note. Negative Neutral Positive. In the non-randomized version, the words are presented in the order listed above. In the randomized-version, the order presentation is randomized within the list.

Word List Creation

We used multiple databases in the construction of the word lists. First, we merged the databases containing different lexical features of English lemmas. The databases are as follows:

- Polysemy and frequency databases: 62,954 words and their polysemy values were extracted from WordNet (Fellbaum, 1998). The words in this database are drawn from the 100,000 most common words in English according to the Corpus of Contemporary American English (Davies, 2008). In addition to the frequency from the COCA corpus, we also obtained the frequency index from the *Corpus of American Soap Operas* for these words.
- Valence database: 13,915 English lemmas from Warriner, Kuperman, & Brysbaert (2013)
- Imageability: 2,645 words from Bird, Franklin, & Howard (2001)

After the merging process, we removed all words that contained any missing values of these word characteristics from the potential word pool. Because of potential parts of speech influences, we also limited the potential word pool to only *nouns*.

Because valence is the lexical feature of interest, we examined the valence distribution of the words in the potential pool, and kept the words that have the highest and lowest valence ratings (the top and bottom 10th percentile) as well as the words in the middle 10 percentile in

order to create three groups of distinctly valenced words.

We then examined the words in terms of their lexical characteristics (frequency, polysemy, imageability) as a function of the valence groups to ensure that the negative, neutral, and positive words do not differ in their frequency (both COCA and SOAP), polysemy, and imageability. In some cases, words were manually selected or removed in order to ensure the balance of these other lexical features. The final word list of 60 words (20 in each valence group) was balanced in terms of frequency, polysemy, and imageability (cf. Table S3).

Unfortunately, despite originally limiting word length to 4-7 letters, this criterion reduced the potential word pool such that the three valence groups were not equally balanced on other lexical features. Thus, we loosened the word length criterion to 3-12 letters, while ensuring that the words within the three valence groups (20 negative, 20 neutral, and 20 positive words) did not differ significantly on word length.

Table S3. Descriptive Statistics of Lexical Features by Valence Groups

|	n	Mean	SD	Min	Max
Log Frequency (COCA)					
Negative	20	1.94	1.77	-1.71	5.32
Neutral	20	2.10	2.44	-2.66	7.44
Positive	20	3.17	1.72	-1.39	5.63
Log Frequency (SOAP)					
Negative	20	1.16	2.58	-4.61	5.60
Neutral	20	0.78	2.84	-3.91	6.64
Positive	20	2.44	2.34	-2.41	6.37
Polysemy					
Negative	20	3.45	2.80	1	13
Neutral	20	4.30	3.71	1	15
Positive	20	3.40	1.82	1	7
Imageability					
Negative	20	4.42	0.76	3.25	6.45
Neutral	20	4.96	0.98	2.92	6.16
Positive	20	4.86	1.15	2.89	6.44

Note. There are no significant differences as a function of valence group in any of the lexical features (all *p*s > .12).

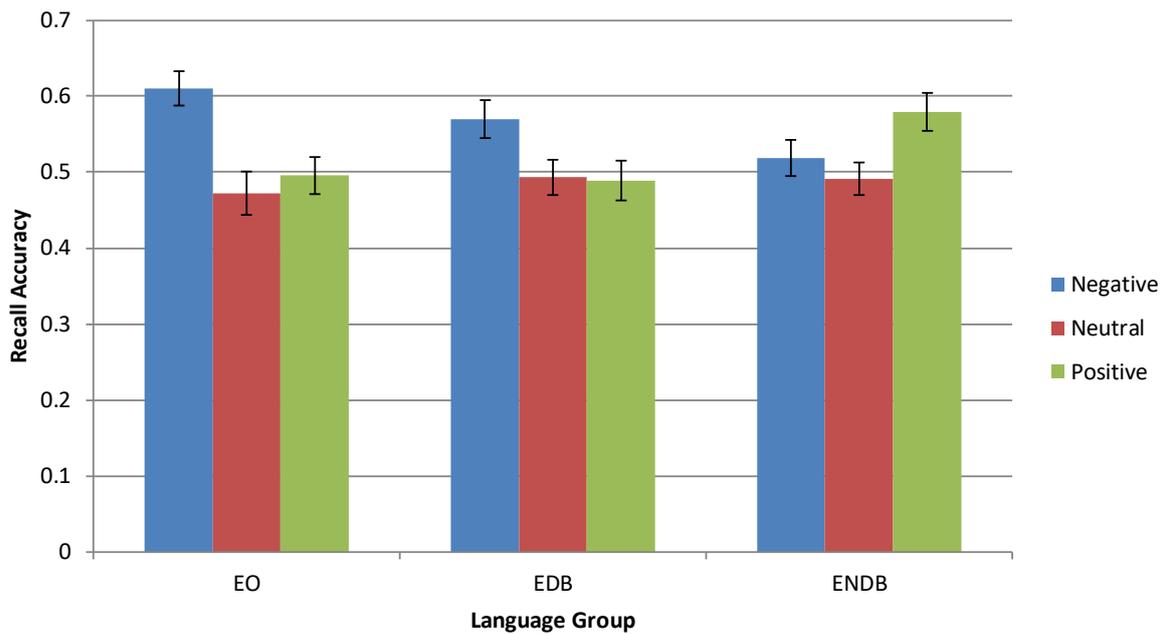**Figure S1.** Stimuli and Procedure for the MM Task.
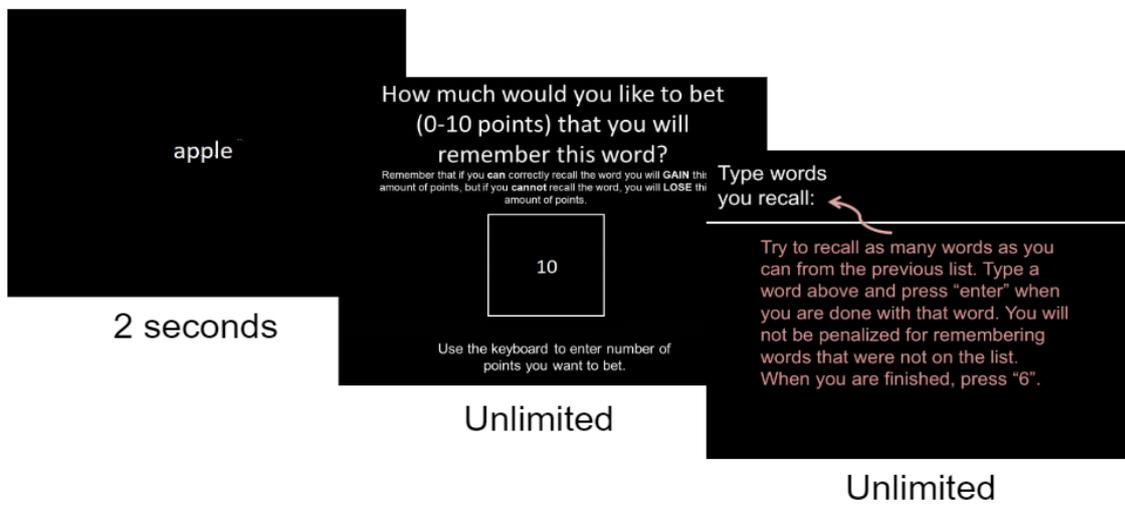
The MM task was administered via the PsychoPy program on PC computers in a laboratory setting.

Figure S2. The average value of bets as a function of language group and valence.
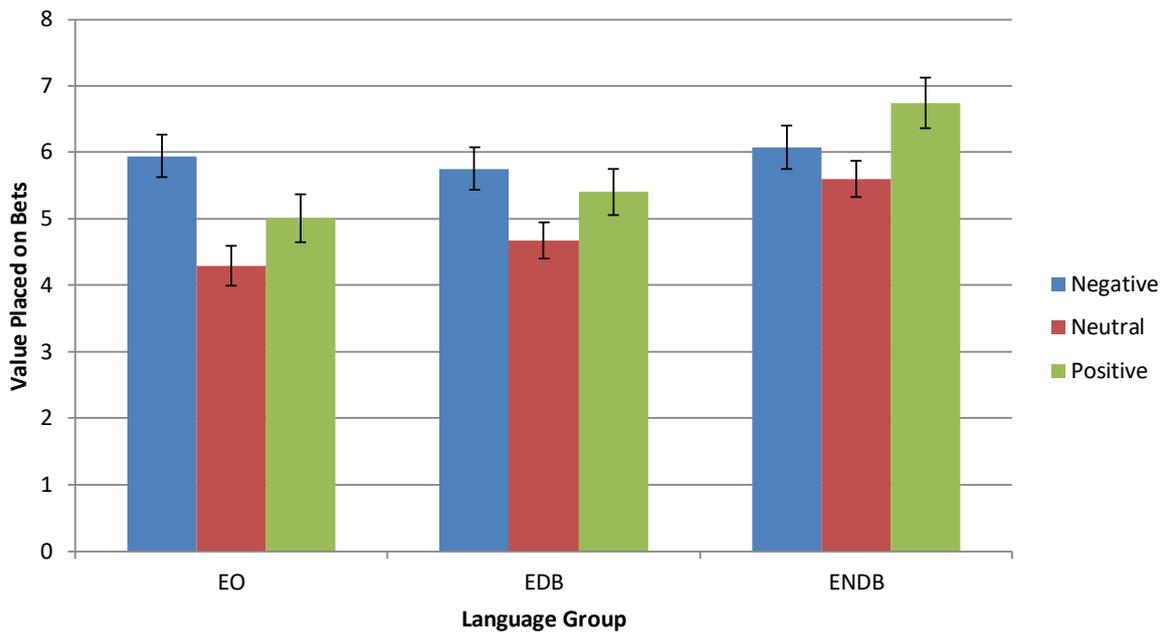
Figure S3. The average recall accuracy as a function of language group and valence.

References

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments & Computers, 33*(1), 73-79. doi: 10.3758/BF03195349

Davies, Mark. (2008-) *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. Available online at https://corpus.byu.edu/coca/.

Davies, Mark. (2011-) *Corpus of American Soap Operas: 100 million words*. Available online at https://corpus.byu.edu/soap/.

Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Torres, J. (2012). Questionnaires in English for heritage speakers of Spanish. *NHRC Language Background Questionnaires*. Retrieved from http://nhlrc.ucla.edu/nhlrc/data/questionnaires

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191-1207. doi: 10.3758/s13428-012-0314-x