

Common and Difference of Mindfulness Interpretation in China and the United States

Eric Qu

Duke Kunshan University

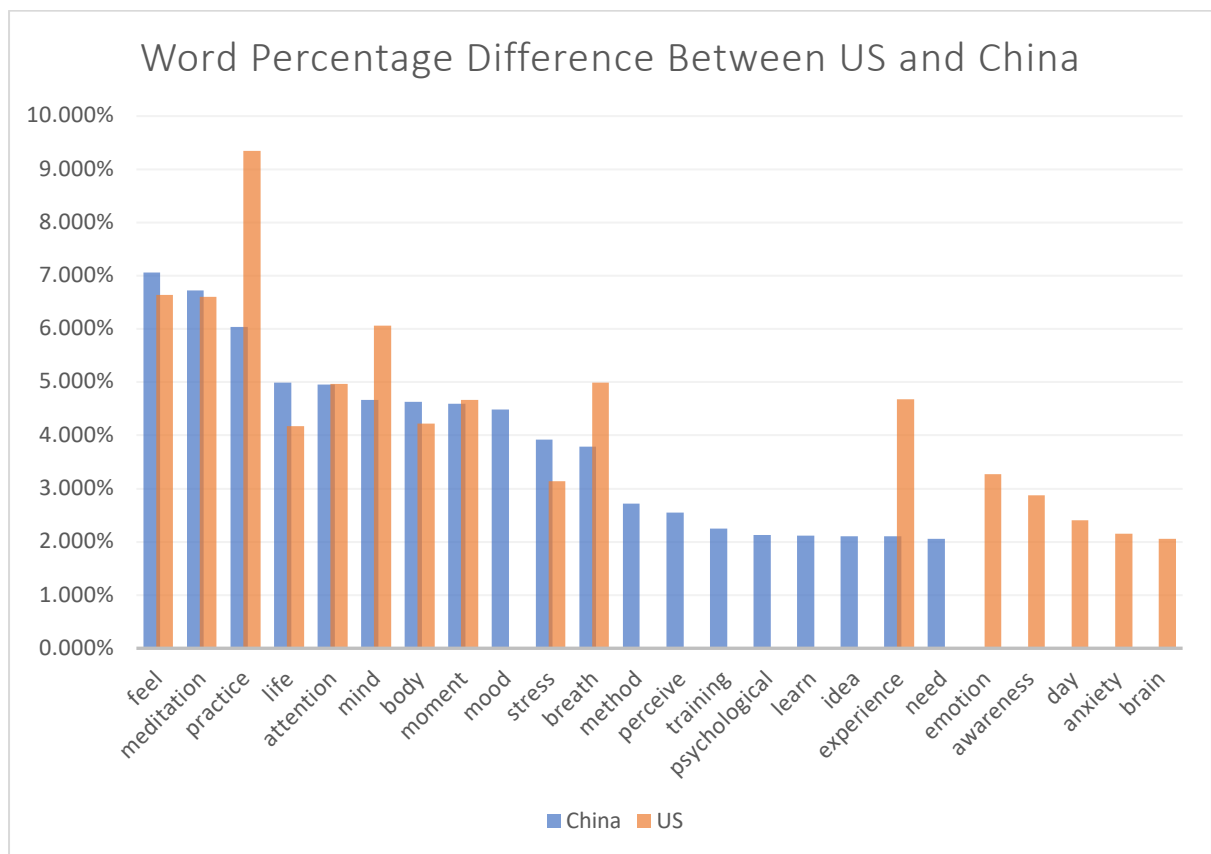
Mindfulness, Stress, and Health Report

This report contains analyses of the word frequency of websites related to mindfulness from China and the United States. The results are Chinese are more interested in the definition of mindfulness, but the practice of which is more prevalent in the US, which indicates the concept and effect of mindfulness is more well-spread in the US than China.

The Origin of Data

The data is collected from the website related to mindfulness. I get the list of websites of US and China from the top search results of the search engine Google and Baidu, respectively. Then translated the words from the Chinese websites into English and lemmatized the words from English websites to make them match. The usable words collected are about a million (95,769) in total. To make it easier to analyze, I chose the words that appeared more than 200 times and listed below. The word “mindfulness” is removed from the graph because it makes the graph looks strange and do not have effects on the results. I also manually changed and merged some synonyms to get more apparent results.

	China	US		China	US
feel	7.060%	6.637%	training	2.245%	
meditation	6.725%	6.606%	psychological	2.130%	
practice	6.037%	9.342%	learn	2.111%	
life	4.996%	4.176%	idea	2.102%	
attention	4.958%	4.962%	experience	2.102%	4.676%
mind	4.662%	6.065%	need	2.054%	
body	4.633%	4.217%	emotion		3.267%
moment	4.595%	4.666%	awareness		2.879%
mood	4.490%		day		2.410%
stress	3.917%	3.135%	anxiety		2.154%
breath	3.793%	4.993%	brain		2.052%
method	2.723%				
perceive	2.551%		mindfulness	26.118%	24.740%



Summary of Data

The words that have clear difference are “practice”, “life”, “mind”, “stress”, “breath”, “experience”. The words that only appear more than 200 times in China are “mood”, “method”, “perceive”, “training”, “psychological”, “learn”, “idea”, “need”. And the words only in the US are “emotion”, “awareness”, “day”, “anxiety”, “brain”.

Analysis of Data

Firstly, the words in the US are more about practice. The most obvious is the word “practice”, but there are also other words. For example, the “breath” is 1.2% more than China’s, indicates that there are more instructions about how to practice mindfulness. Another obvious example is the “day”, it’s not even in China’s list. In the

original data, China's "day" count is 118 versus the US's 236. The "day"'s appearance meaning there are specific guidelines on the websites. And the search engine would list the most popular websites in front, which means that people in the US are more care about how to practice mindfulness.

Then, the Chinese are more about the definition of mindfulness. Such as "life", "mood", "need", etc. The most noticeable word is "psychological" (223 versus 67), which is a clear definition word. Another example is "perceive" (267 versus 3), which is a highly technical word. All these words are more likely to appear in the definition page rather than instruction page.

But what causes the difference? I think the reason is that the concept of mindfulness is familiar in the US but not so well-spread in China. So Chinese people are more interested in what is mindfulness. And Americans already know the concept and effects of mindfulness, which makes them want to know how to practice mindfulness.

There are some more interesting points in the data. Firstly, the word "brain" appears a lot in the US but no so much in China (201 versus 139). But the "brain" should appear more in the definition. Maybe the Americans are more interested in the mechanism of mindfulness, or they have better educations than Chinese, or just simply because the Google search engine is more academic than Baidu.

Next, there are some words appear almost the same in both results, and they have both high percentages. Such as "feel", "attention", "moment". It indicates that these words are important in both the definition and practice of mindfulness, which may be essential concepts in mindfulness.

What Can Be Improved

Firstly, I used python to simulate the requests. Although with the fake header, there are changes that major websites are blocking me from accessing their contents.

Then, the translation can cause some problems. I tried to use the same translation library to translate both English and Chinese to cross exam the results, but it is inevitable that there are some mistakes in the translation, letting some key information missing.

Moreover, the structure of the two languages is different. The words in English are separated, but they are linked in sentences in Chinese. I have to use some stopwords to separate the Chinese sentences, which could cause mistakes in the statistics. And the words in Chinese may be translated to a phrase in English, creating more bias.

Finally, the form of words may have effects on the statistics. Because the words collected are original from the websites, the form of which could differ in the contents. I used a lemmatization library to stem the words to its original form, but it can't transform the verb or adjective to the noun. Although they may have the same meaning, they were counted differently.