

# INSTRUCTIONS

There are two types of questions in this survey:

1. First we ask you to read a short text and come up with two words inspired by it.
2. Then you need to evaluate how similar two words are within the context of the same particular text.

The whole survey needs to be finished in one go, you cannot save your progress and you cannot go back to correct answers either. Please make sure you allocate plenty of time for it.

## 1. TWO WORDS INSPIRED BY A SHORT TEXT:

The first task in each page will be to read a short text and come up with two words inspired by it. No need to think too long about it, the most important thing is that you read the whole text. The words you chose can describe the general topic of the text, be related to the context or simply come to your mind while reading the text.

## 2. SCORING SIMILARITY:

Highly similar words usually represent the same type or category of thing. We say two words are **synonymous** when they are extremely similar:

- cup / mug
- glasses / spectacles
- envy / jealousy

However similarity is not black and white, there are degrees of it. For example word pairs that are not synonymous may still be quite similar:

- love / affection
- frog / toad

We can find degrees of similarity even with words that are not synonymous at all. A cat and a dog are more similar than a cat and a shark. They are all animals, but cats and dogs share a lot more characteristics (mammals, pets...). However a cat and a shark are more similar than a cat and a ball, they are both living animals after all.

In contrast, although the following word pairs are **related**, they are **not very similar**. The words represent entirely different types of thing:

- car / motorway
- car / crash

In this survey, you are asked to compare two words and to rate how similar they are **within the context of a specific group of sentences**. The rating goes from 0 to 6, 0 being words that don't have any similarity at all and 6 being words that mean basically the same.

You will have to estimate the degrees in between. A useful guide is to consider **how close the words are (or are not) to being synonymous**. Remember, things that are related are not necessarily similar.

## 3. THE EFFECTS OF CONTEXT:

In the second task we are asking you to rate the similarity of words **within a specific text**.

A particular context can have a strong effect in the meaning of words. As a consequence context can change how we score the similarity between two words.

Please read the next two examples and make sure you understand them, if you have questions please contact us before you start the survey.

### Example 1 - Room and Cell:

Some words have different senses. Here we see one example **cell** could mean a biological term, like in "human cell", or a particular kind of room, like in "prison cell". By selecting one of these senses, a particular context can have a big impact in the meaning of one of these words.

Sentence 1: Her prison **cell** was almost an improvement over her **room** at the last hostel.

In Sentence 1 the words **room** and **cell** both refer to rooms in a building, they should be scored as very similar words.

Sentence 2: His job as a biologist didn't leave much **room** for a personal life. He knew much more about human **cells** than about human feelings.

However in Sentence 2 the words are being used in different senses ('room' as an abstract concept and 'cell' as a biological term) the words mean very different things:

The scoring of similarity should be a lot lower in the context of Sentence 2 than it was in the context of Sentence 1.

### Example 2 - People and Population:

Even when we are looking at word in which the sense doesn't change, a particular context can make us think of concepts in slightly different ways.

Sentence 1: The **population** of India is actually bigger than most **people** expect.

In Sentence 1 population and people both refer to human beings.

Sentence 2: The **population** of bison became a lot smaller when **people** settled in the valley.

Now population refers to a collective of animals instead.

The similarity of these words should be higher in the context of Sentence 1 than in the context of Sentence2 .

### Example 3 - Butter and Oil:

Another example where the sense of the words doesn't need to change for the similarity score to change.

Sentence 1: Americans love to pour **butter** over their popcorn. Spaniards use olive **oil** for the same thing.

In Sentence 1 below, the butter is referred as being "poured" which makes us realise we are talking about "liquid butter". In addition we are talking about butter and oil being used in very similar ways. The score of similarity should be very high

Sentence 2: The high content of animal saturated fat makes **butter** solid at room temperature. Olive **oil** only contains vegetable unsaturated fat and needs to be a lot colder to become solid.

In contrast, in Sentence 2, we are focusing on differences: animal vs vegetable, saturated fat vs unsaturated fat, solid vs liquid.

Again, the two words should be scored as more similar in Sentence 1 than in Sentence 2.

The effects of context (and the rating of similarity itself) can be very difficult to quantify, these questions are very subjective and there are not perfect answers. Please try not to overthink your responses and **use your intuition and gut feeling** instead.

Thanks a lot for participating in our study!

15%

\*Read the following text and write down two words inspired by it:

Automatically and dynamically allocated objects are only initialized if an initial value is explicitly specified; otherwise they initially have indeterminate values (typically, whatever bit pattern happens to be present in the storage, which might not even represent a valid value for that type). If the program attempts to access an uninitialized value, the results are undefined. Many modern compilers try to detect and warn about this problem, but both false positives and false negatives occur.

First word:

Second word:

[Next](#)

16%

\*Read the sentences again and then score the similarity between the words **occur** and **happen** when compared **within this specific text**:

Automatically and dynamically allocated objects are only initialized if an initial value is explicitly specified; otherwise they initially have indeterminate values (typically, whatever bit pattern **happens** to be present in the storage, which might not even represent a valid value for that type). If the program attempts to access an uninitialized value, the results are undefined. Many modern compilers try to detect and warn about this problem, but both false positives and false negatives **occur**.

0 = Not similar at all

6 = Extremely similar

0     1     2     3     4     5     6

[Next](#)