

TIP

To get up to speed on JSON if you're unfamiliar, W3Schools [has a good introduction](#).

For image generation models, format is very important, because the opportunities for modifying an image are near endless. They range from obvious formats like stock photo, illustration, and oil painting, to more unusual formats like dashcam footage, ice sculpture, or in minecraft (see [Figure 1-7](#)).

Input:

```
business meeting of four people watching on MacBook on top of  
table, in minecraft
```

The output is shown in [Figure 1-7](#).



Figure 1-7. Business meeting in minecraft

When setting a format it is often necessary to remove other aspects of the prompt that might clash with the specified format. For example, if you supply a base image of a stock photo, the result is some combination of stock photo and the format you wanted. To some degree, image generation models can generalize to new scenarios and combinations they haven't seen before in their training set, but in our experience, the more layers of unrelated elements, the more likely you are to get an unsuitable image.

There is often some overlap between the first and second principles, Giving Direction and Specifying Format. The latter is about defining what type of output you want, for example JSON format, or the format of a stock photo. The former is about the style of response you want, independent from the format, for example product names in the style of Steve Jobs, or an image of a business meeting in the style of Van Gogh. Where

there are clashes between style and format, they're often best resolved by dropping whichever is less important to your final result.

3. Provide Examples

The original prompt didn't give the AI any examples of what you think *good* names look like. Therefore the response is approximate to an average of the internet, and you can do better than that. Researchers would call a prompt with no examples *zero-shot*, and it's always a pleasant surprise when AI can even do a task zero shot: it's a sign of a powerful model. If you're providing zero examples, you're asking for a lot without giving much in return. Even providing one example (*one-shot*) helps considerably, and it's the norm among researchers to test how models perform with multiple examples (*few-shot*). One such piece of research is the famous GPT-3 paper [Language Models are Few-Shot Learners](#), the results of which are illustrated in [Figure 1-8](#), showing adding one example along with a prompt can improve accuracy in some tasks from 10% to near 50%!

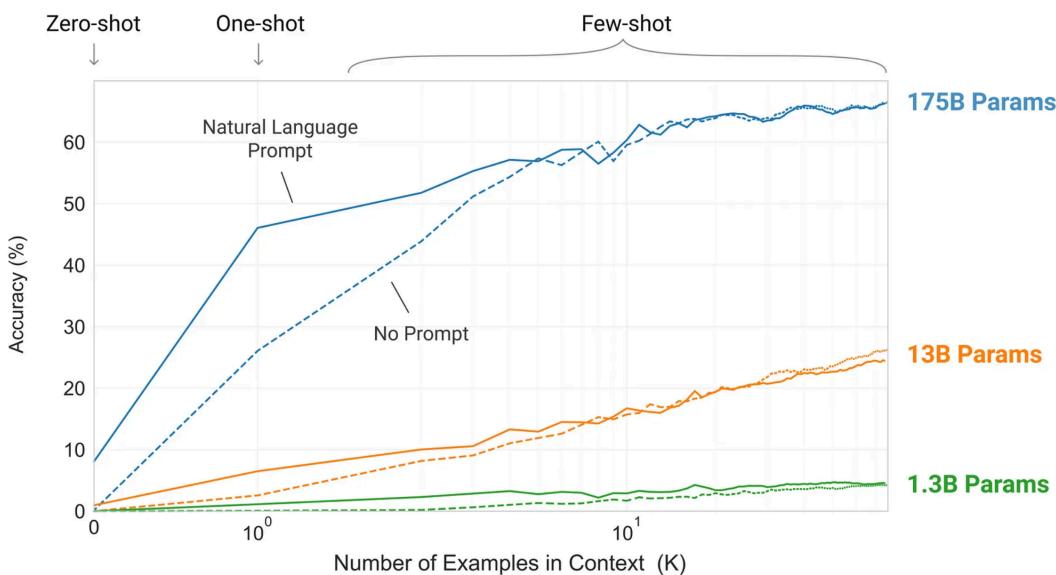


Figure 1-8. Number of examples in context

When briefing a colleague or training a junior employee on a new task, it's only natural that you'd include examples of times that task had previously been done well. Working with AI is the same, and the strength of a prompt often comes down to the examples used. Providing examples can

sometimes be easier than trying to explain exactly what it is about those examples you like, so this technique is most effective when you are not a domain expert in the subject area of the task you are attempting to complete. The amount of text you can fit in a prompt is limited (at the time of writing around 6,000 characters on Midjourney, and approximately 32,000 characters for the free version of ChatGPT), so a lot of the work of prompt engineering involves selecting and inserting diverse and instructive examples.

There's a tradeoff between reliability and creativity: go past 3-5 examples and your results will become more reliable, while sacrificing creativity. The more examples you provide, and the lesser the diversity between them, the more constrained the response will be to match your examples. If you change all of the examples to animal names in the previous prompt, you'll have a strong effect on the response, which will reliably return only names including animals.

Input:

```
Brainstorm a list of product names for a shoe that fits any foot size.
```

```
Return the results as a comma separated list, in this format:
```

```
Product description: A shoe that fits any foot size
```

```
Product names: [list of 3 product names]
```

```
## Examples:
```

```
Product description: A home milkshake maker.
```

```
Product names: Fast Panda, Healthy Bear, Compact Koala
```

```
Product description: A watch that can tell accurate time in space.
```

```
Product names: AstroLamb, Space Bear, Eagle Orbit
```

```
Product description: A refrigerator that dispenses beer
```

```
Product names: BearFridge, Cool Cat, PenguinBox
```

Output:

Product description: A shoe that fits any foot size

Product names: FlexiFox, ChameleonStep, PandaPaws

Of course this runs the risk of missing out on returning a much better name that doesn't fit the limited space left for the AI to play in. Lack of diversity and variation in examples is also a problem in handling edge cases, or uncommon scenarios. Including 1-3 examples is easy and almost always has a positive effect, but above that number it becomes essential to experiment with the number of examples you include, as well as the similarity between them. There is some evidence ([Hsieh et al, 2023](#)) that direction works better than providing examples, and it typically isn't straightforward to collect good examples, so it's usually prudent to attempt the principle of Giving Direction first.

In the image generation space, providing examples usually comes in the form of providing a base image in the prompt, called *img2img* in the open-source [Stable Diffusion](#) community. Depending on the image generation model being used, these images can be used as a starting point for the model to generate from, which greatly affects the results. You can keep everything about the prompt the same, but swap out the provided base image for a radically different effect, as in [Figure 1-9](#).

Input:

```
stock photo of business meeting of four people watching on  
white MacBook on top of glass-top table, Panasonic, DC-GH5
```

The output is shown in [Figure 1-9](#).



Figure 1-9. Stock photo of business meeting of four people

In this case, by substituting for the image shown in [Figure 1-10](#), also from Unsplash, you can see how the model was pulled in a different direction, and incorporates whiteboards and sticky notes now.