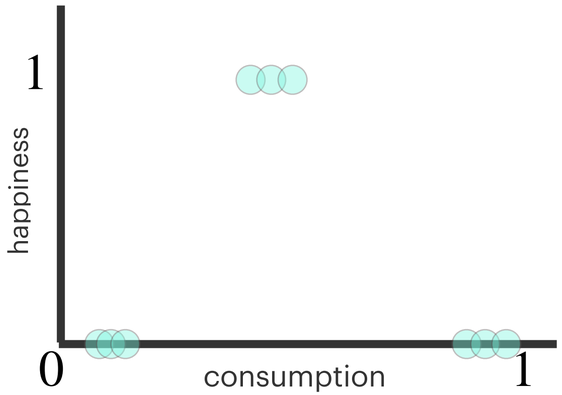
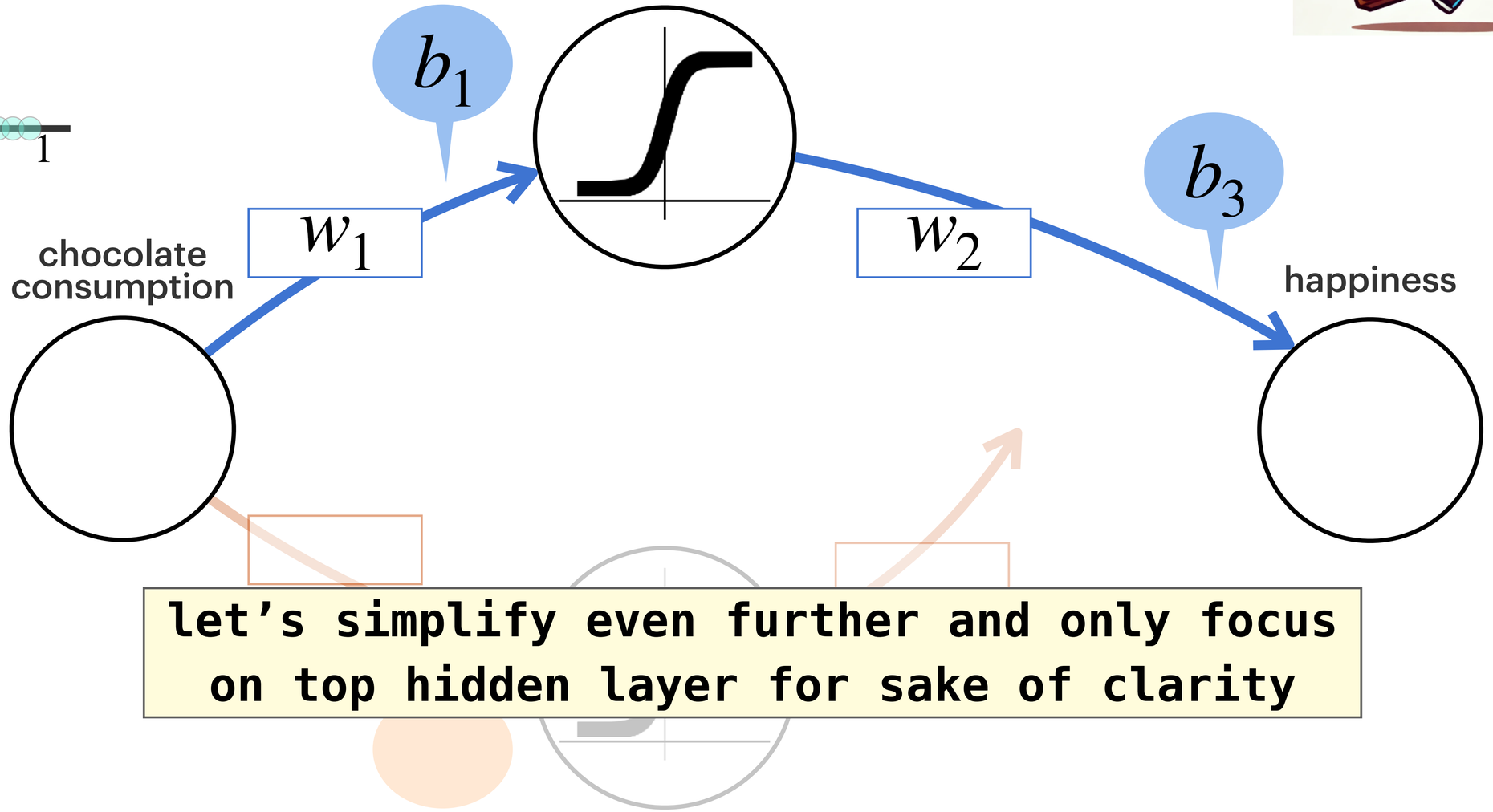
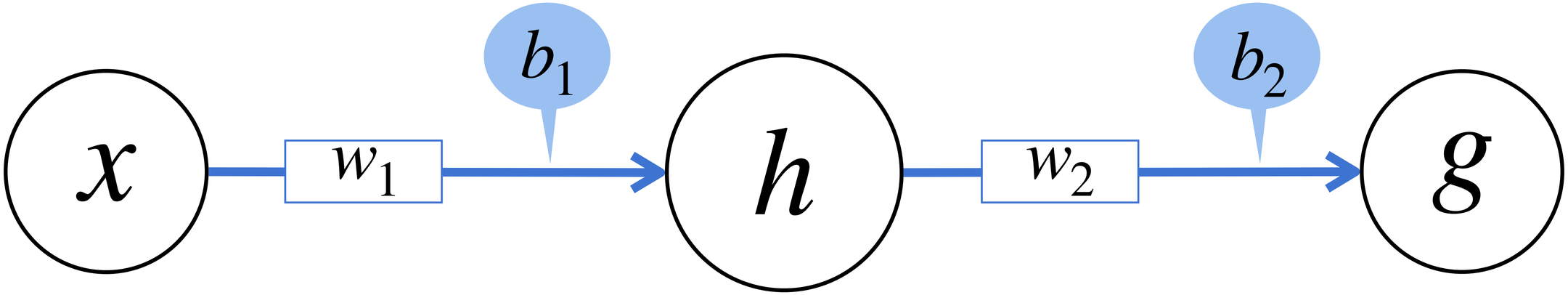
Back to Our Neural Network





let's simplify even further and only focus on top hidden layer for sake of clarity

Backpropagation



Chain rule:

If we want to know how changing x affects $f(g(x))$ we first need to think about how changing x affects $g(x)$ and then how changing $g(x)$ affects$f(g(x))$:

$$\frac{\partial f(g(x))}{\partial x} = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial x}$$

the derivative of the outer function f evaluated at $g(x)$ multiplied by the derivative of the inner function $g(x)$