Temperature Calculation: Step-by-Step

Given: Logits = [2.0, 1.0, 0.5, 0.2]
Tokens = ["cat", "dog", "bird", "fish"]

Step 2: Scale by T=1.0 (UNCHANGED)
Scaled = [2.0, 1.0, 0.5, 0.2]
Scaled = [4.0, 2.0, 1.0, 0.5, 0.5, 0.21, 0.2, 0.5]
= Softmax → [0.63]
Softmax → [0.19, 0.12, 0.10]
Softmax 53% on "cat" (BALANCED)
→ **73% on "cat" (VERY FOCUSED)**

Step 3: Scale by T=2.0 (FLATTER)
Scale [5, 0.1]
Softmax [8, 0.17, 0.15]
38% ... CH FLATTER)

$$p_i = \frac{\exp(logit_i/T)}{\sum ...}$$

Lower T → More confident (peaky)
Higher T → More random (flat)