But wait... (mark II)

Both GloVe and word2vec use **context window weighting**:

- word2vec: **samples a window size** $\ell \in [1, L]$ for each token before extracting counts
- GloVe: weighs contexts by their **distance from the target word** using the harmonic function $\frac{1}{d}$

What is the impact of using the different context window strategies between methods?

Levy, Goldberg, and Dagan **did not** investigate this, instead using word2vec's method across all of their evaluations.