Two (Representation) by Two (Encoding Type) Model Performances

Table 1:

Best and average performance on direct items (experiment 1)

Model	Best Performance		Average Performance	
	co-occurrence	similarity	co-occurrence	similarity
Space	1	.703	.459(.504)	-.104 (.328)
Graph	1	.703	.477 (.554)	-.103 (.328)

Note. Scores are organized by data structure (rows) and encoding type (columns). Values in parentheses are standard deviations.

For each of the four types of models, we averaged performance across models that vary only in minor parameters (e.g. window direction, normalization type, etc.). However, doing so may obscure the possibility that one model with one specific set of minor parameters performed best overall, so we also present the best performing model of each of the four major types. These results are shown in Table 1.

To compare the average performance between model types, we first used Fisher transformation, a method to scale Pearson or Spearman correlation scores so that their distribution approximate normal distribution (Zar, 2014), to convert each model's spearman correlations in every single run to a 'z-score'. Then we conducted a mixed effect model analysis using the z-scores as the dependent variables. We used the model's Encoding Type (co-occurrence or similarity) and Representational Structure (space or graph) as predictor variables and had each run of the model as the random factor. In this analysis, we found a significant main effect of Encoding Type (with co-occurrence outperforming similarity, $t = 87.58$ $p<.001$), but we did not find significant main effect of the Representational Structure ($t = 1.79$, $p = .08$), or of the interaction ($t=1.66$, $p = .10$). When considering the best performing models, we found that the

best co-occurrence graphical and co-occurrence space model (z = 4.952) the similarity models (similarity space:z = 1.110, similarity graph z = 1.109).

In terms of the average performance of each four model types, the co-occurrence space models was the best, followed by co-occurrence graph, and then similarity space and similarity graph (co-occurrence graph: mean z = 1.759, std = 2.083, range [-.631, 4.952]; co-occurrence space: mean z = 1.810, std = 2.207, range [-.265, 4.952]; similarity graph: mean z = -.136, std = .425, range [-1.462, 1.109]; similarity space: mean z = -.137, std = .425, range [-1.462, 1.110]). The overall outperformance of co-occurrence models, and especially co-occurrence space models is expected, as these models directly encoded the co-occurrence, which are what the 'direct' tasks asked for. Our core interest lies in the 'indirect' condition: whether the models utilize these direct co-occurrence to form indirect relations and make inference. We present the result of experiment 2 in the same format. The mean and best performance of the four major types in experiment 2. are shown in Table 2.

Table 2:

Best and average performance on indirect items (experiment 2)

Model	Best Performance		Average Performance	
	co-occurrence	similarity	co-occurrence	similarity
Space	.758	.723	.564(.168)	.437 (.390)
Graph	.901	.750	.450 (.353)	-.007 (.273)

Note. Scores are organized by data structure (rows) and encoding type (columns). Values in parentheses are standard deviations.

Similarly we compared the average performance between model types after fisher-z transformation. We found a significant main effect of Encoding Type (with co-occurrence outperforming similarity, $t = 58.92$ $p<.001$), a significant main effect of the Representational

Structure ($t = 9.83$, $p < .001$), and of the interaction ($t = 30.91$, $p < .001$). When considering the best performing models, we found that the best co-occurrence graphical model ($z = 1.484$) outperformed all three other models (co-occurrence space: $z = .991$, similarity space: $z = .914$, similarity graph $z = .975$). For the average performance of each four model types, the co-occurrence space models was the best, followed by co-occurrence graph, and then similarity space and similarity graph (co-occurrence space: mean $z = 0.675$, std = 0.256, range [.349, .991]; co-occurrence graph: mean $z = .553$, std = .439, range [-0.458, 1.484]; similarity space: mean $z = 0.531$, std = 0.476, range [-0.890, 0.914]; similarity graph: mean $z = -.007$, std = 0.302, range [-0.988, 0.975]). These results show that in inferring the indirect relations, co-occurrence models maintain their advantage over the similarity models. Moreover, while co-occurrence space models performed better on average, the co-occurrence graph models dominated the top.