Transformer models employing the self-attention mechanism have reached state of the art performance in a range of computer vision and NLP tasks. Typical explanations of the self-attention function claim that it works by computing a semantically meaningful “similarity score” or “compatibility function” between input tokens and using that value as a weight in further computations. In this paper, several modifications are made to the standard self-attention function to remove any token-to-token interactions while maintaining the overall structure of computation. When tested, a number of these modified self-attention functions achieve similar and even superior performance to standard self-attention on both image classification and natural language translation tasks.