Transformer test loss on random tasks with different wildcard alphabet sizes