Attention Heads Effect (Noising)
Letter-String Analogy Task ('+1' vs 'Swap' Rule)