

Mean frequency of consistency on DSMi, DSMu, and MiN across the six different question orders

Temperature: 0, Condition: chain of thought

