

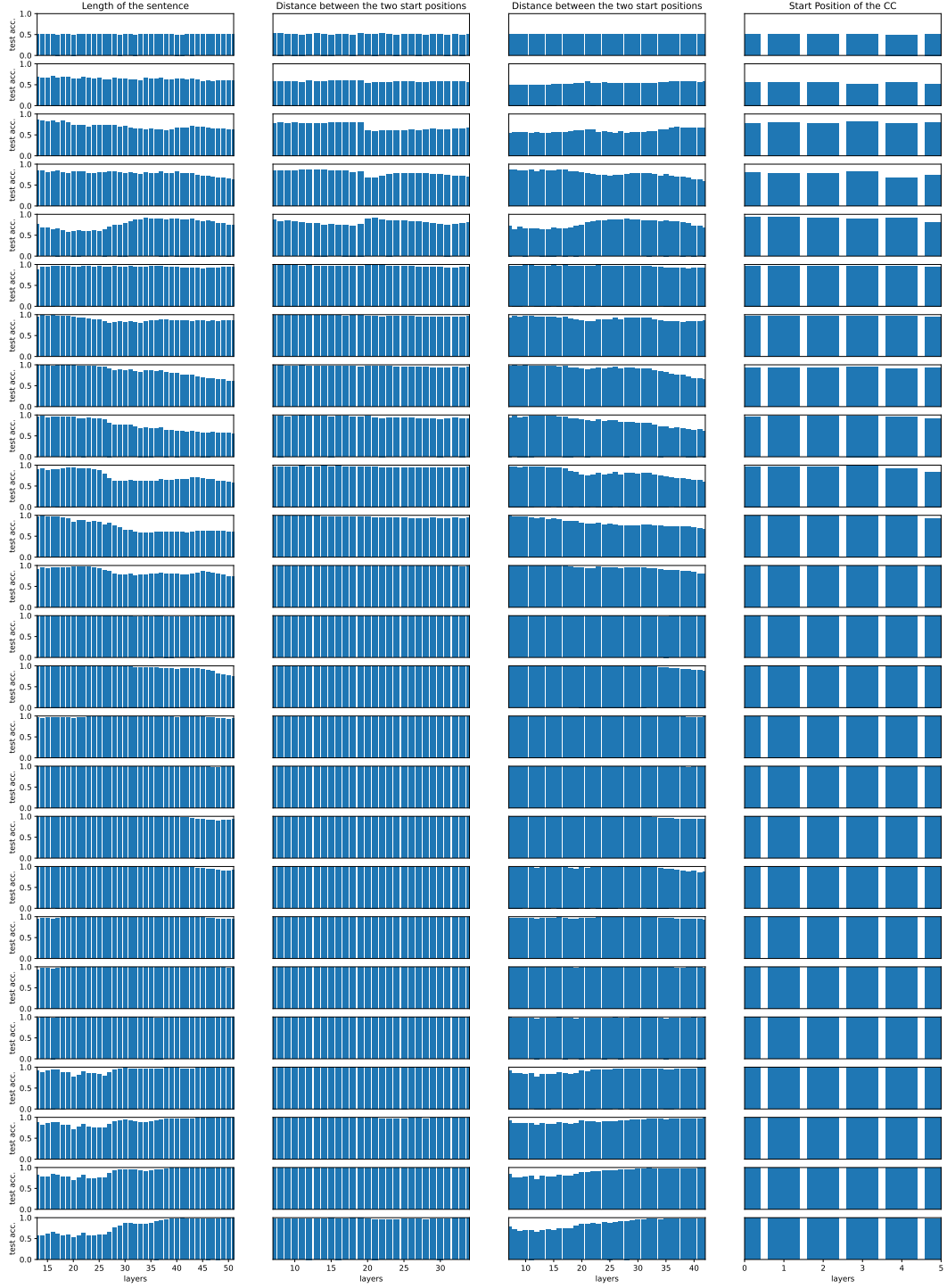


Figure A1: Full results for BERT_{LARGE} on artificial data. Columns indicate the variable that the training and test set controls for.

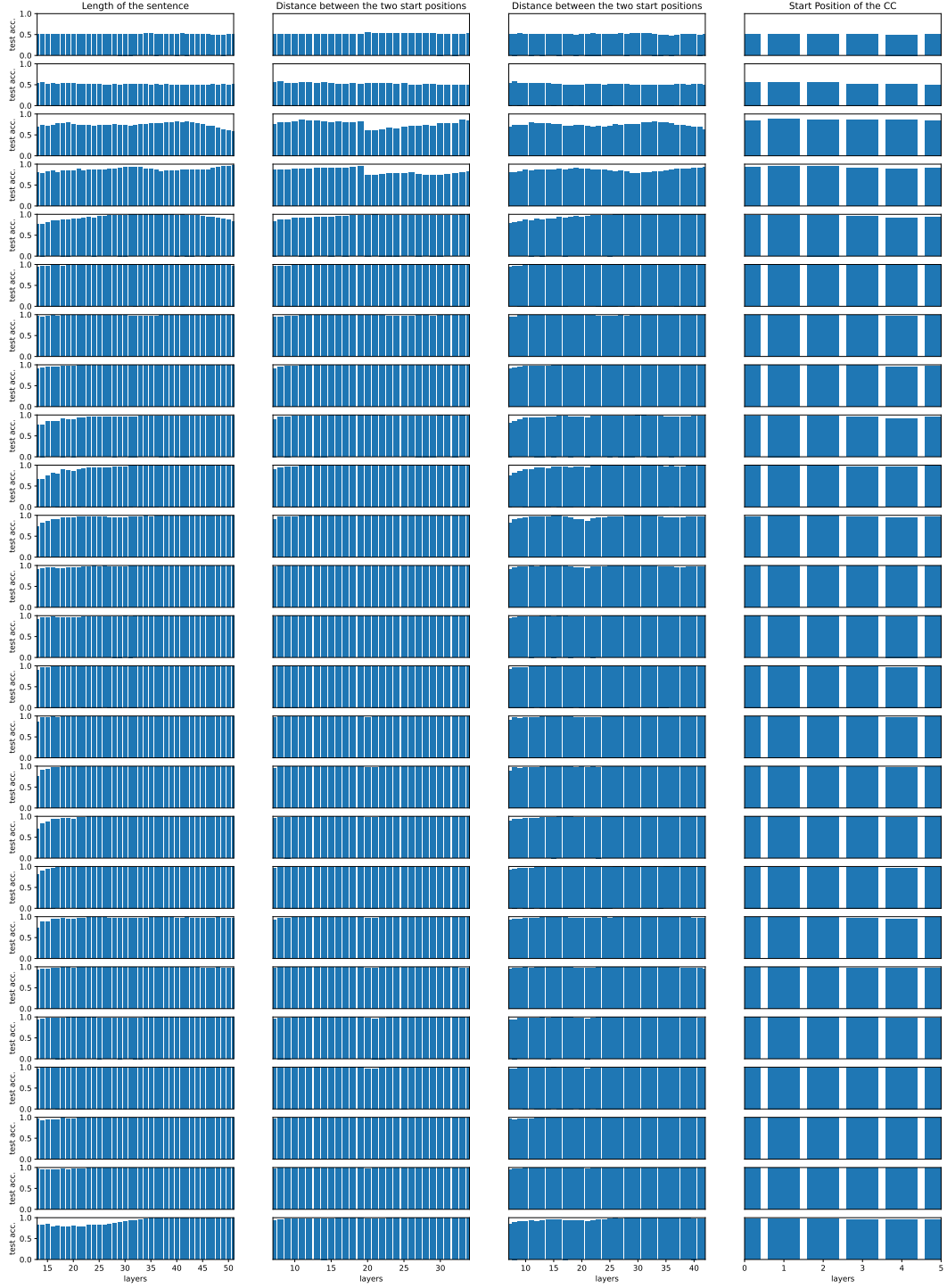


Figure A2: Full results for RoBERTa_{LARGE} on artificial data. Columns indicate the variable that the training and test set controls for.

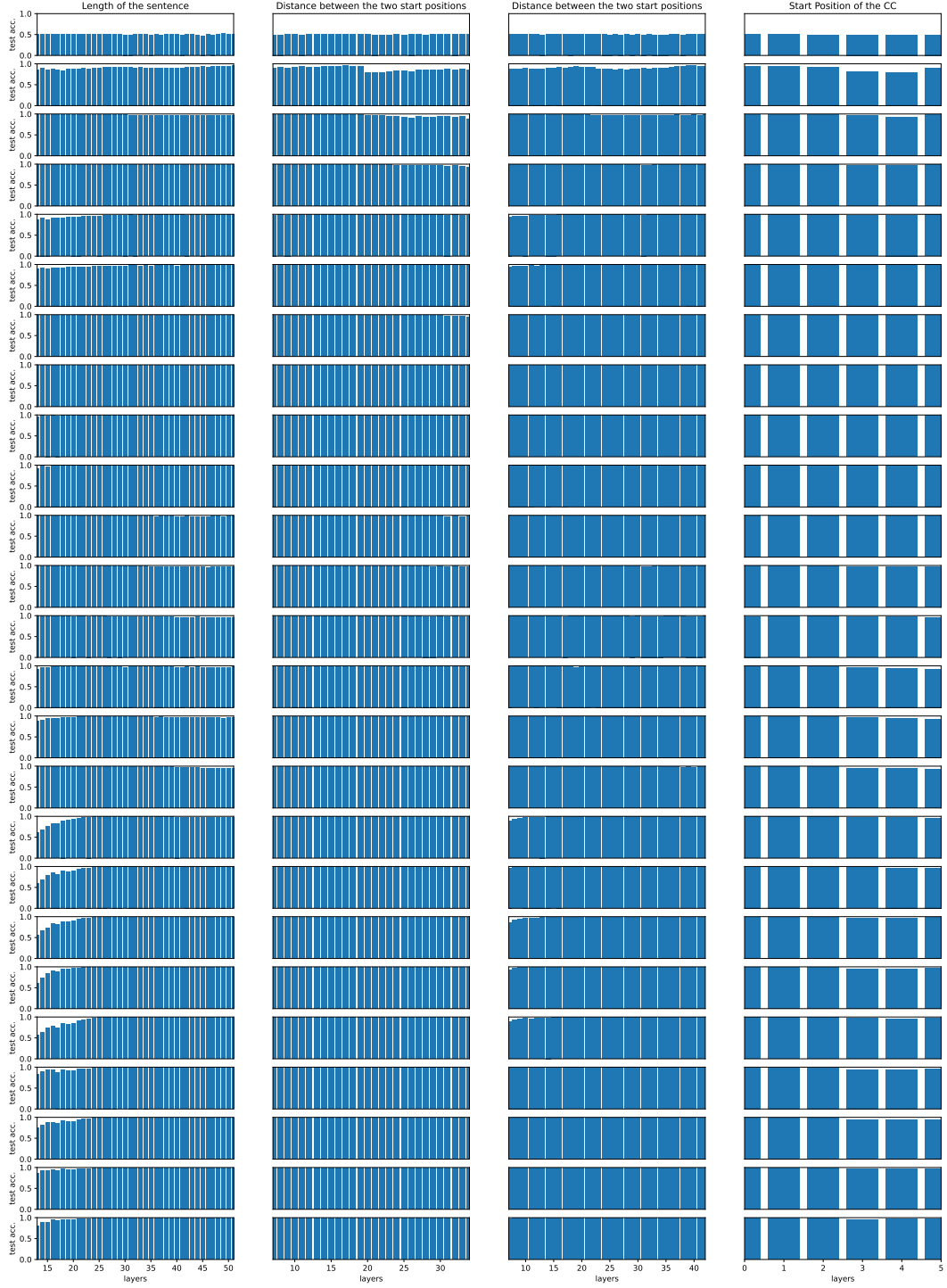


Figure A3: Full results for DeBERTa_{LARGE} on artificial data. Columns indicate the variable that the training and test set controls for.

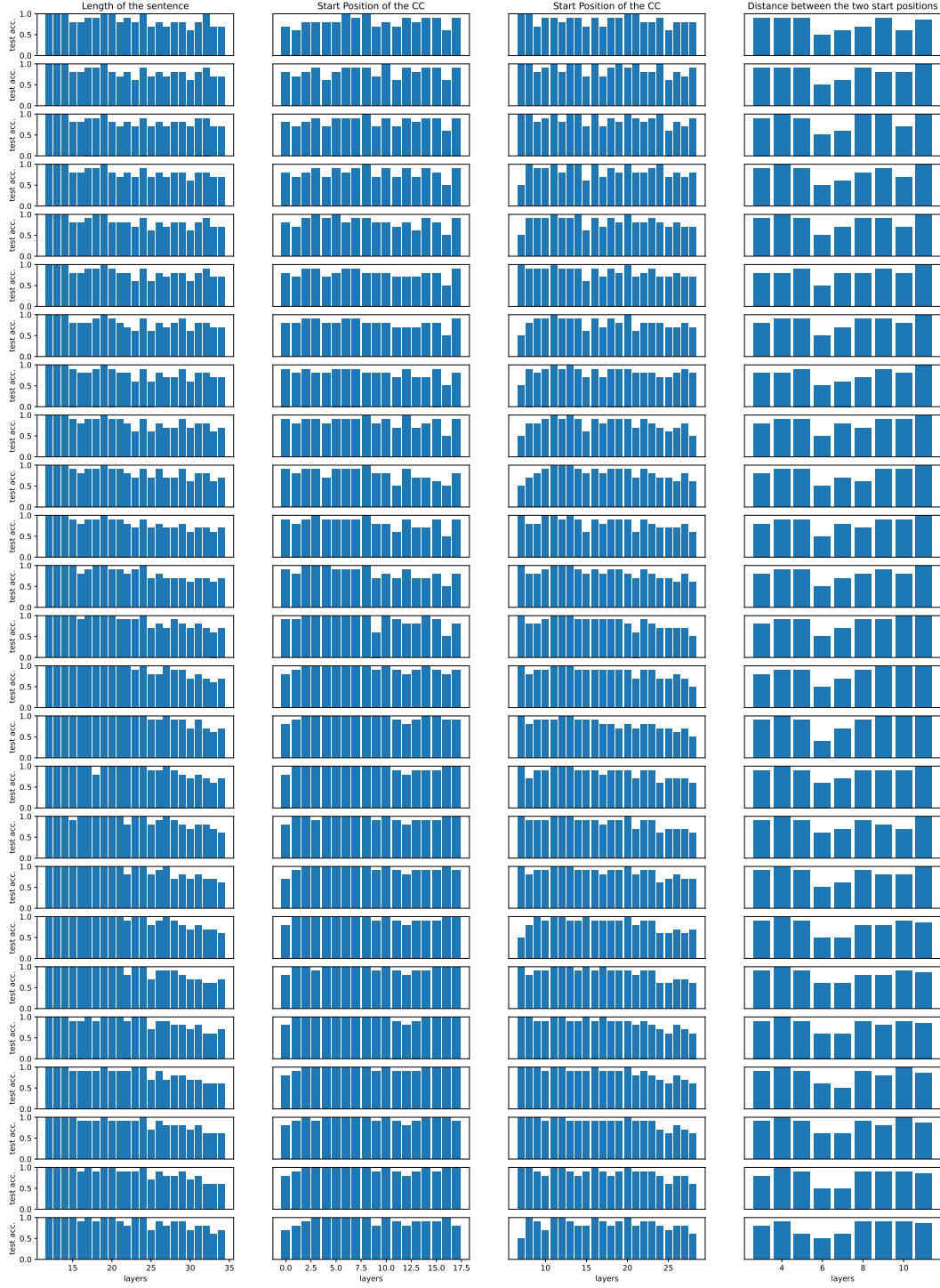


Figure A4: Full results for BERT_{LARGE} on corpus data. Columns indicate the variable that the training and test set controls for.

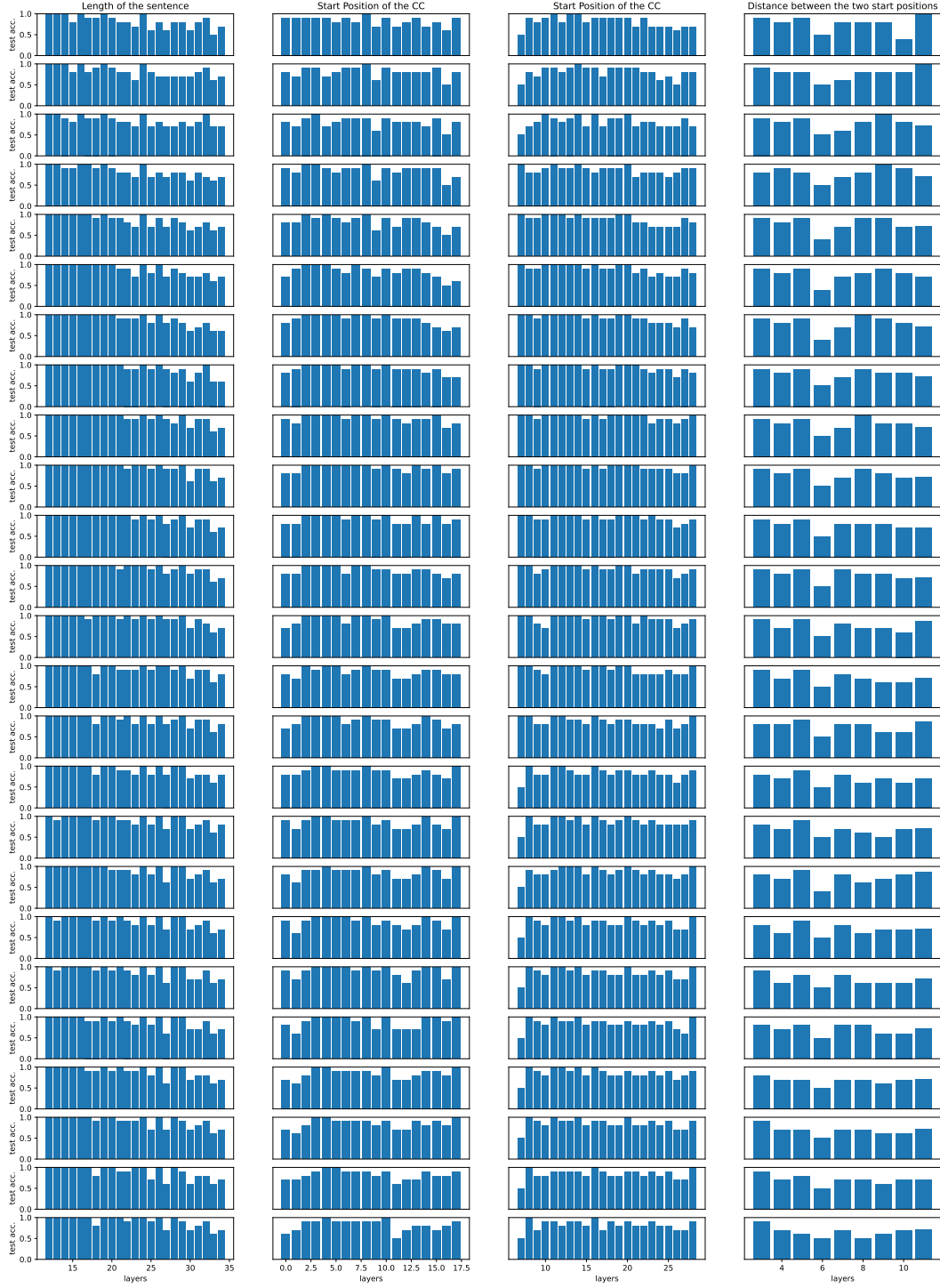


Figure A5: Full results for RoBERTa_{LARGE} on corpus data. Columns indicate the variable that the training and test set controls for.

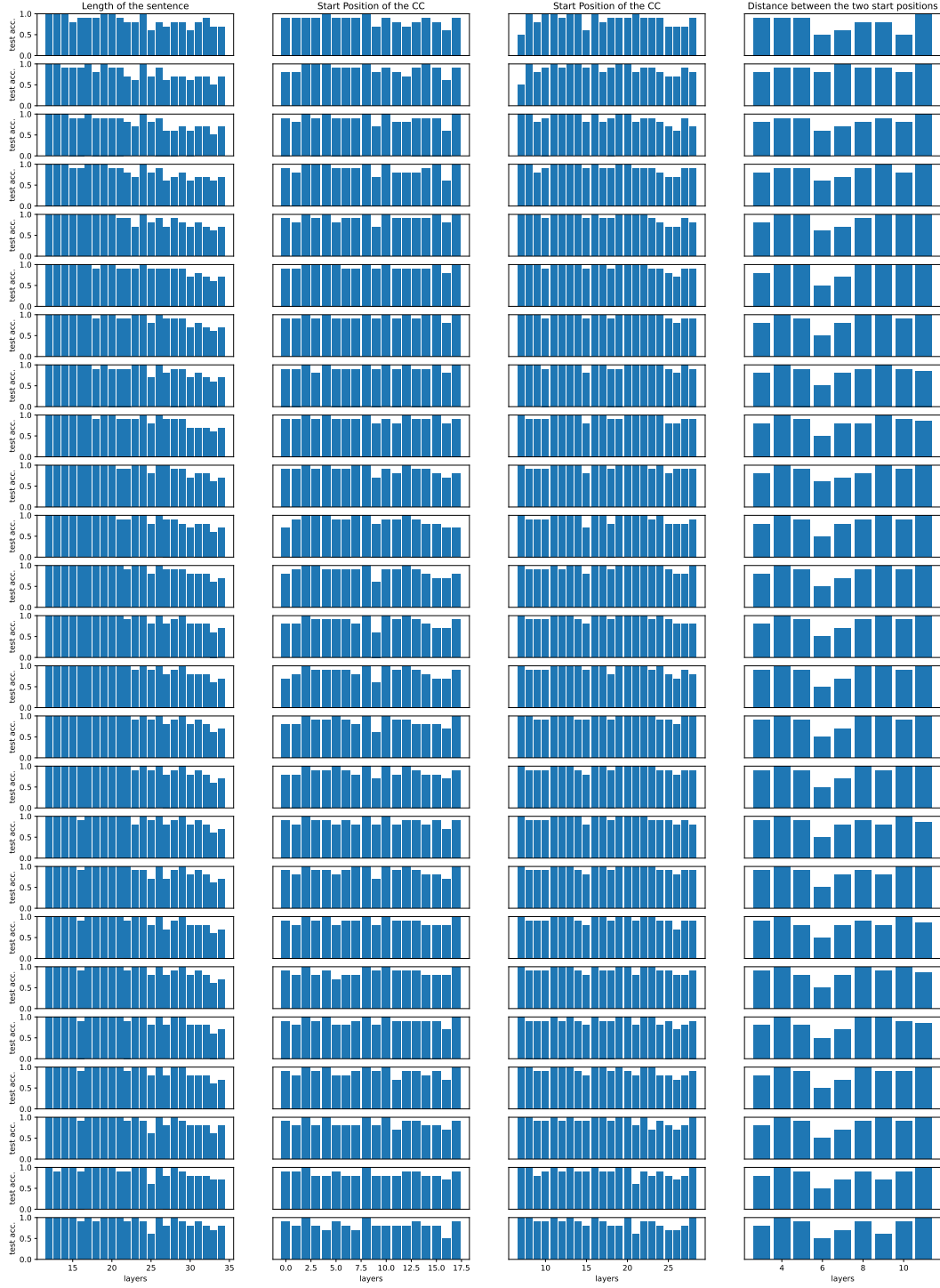


Figure A6: Full results for DeBERTa_{LARGE} on corpus data. Columns indicate the variable that the training and test set controls for.