

% Heldout Test Cases Passed

Validation Score

100%  
75%  
50%  
25%  
0%

Manual

autogen-0%

autogen-5%

autogen-10%

autogen-20%

