1 Experiment Set Difference: Conv (25-50 words) v. Conv (3-50 words)

1.1 Average Treatment Effect (Per Experiment in Experiment Set)

cpu_hours	0.0813406 +/- 0.0130409 (stderr) (p=4.36932e-06, Welch's t-test)
gpu_hours	0.0669757 +/- 0.00239887 (stderr) (p=5.57528e-07, Welch's t-test)
estimated_carbon_impact_kg	0.0048772 +/- 0.000697035 (stderr) (p=1.93072e-06, Welch's t-test)
total_power	0.0160142 +/- 0.00183235 (stderr) (p=3.77302e-07, Welch's t-test)
kw_hr_gpu	0.00826282 +/- 0.000841254 (stderr) (p=2.38469e-07, Welch's t-test)
kw_hr_cpu	0.00187278 +/- 0.000340574 (stderr) (p=1.30387e-05, Welch's t-test)
exp_len_hours	0.0803547 +/- 0.0129589 (stderr) (p=4.57027e-06, Welch's t-test)
average_realtime_carbon_intensity	-16.2576 +/- 3.51527 (stderr) (p=0.00109106, Welch's t-test)

1.2 Cumulative Difference (Per Experiment Set)

cpu_hours	0.40670277777777786
gpu_hours	0.3348787200059824
estimated_carbon_impact_kg	0.02438599036448791
total_power	0.08007124283852213
kw_hr_gpu	0.04131410557575092
kw_hr_cpu	0.009363896220782049
exp_len_hours	0.40177337878280217
average_realtime_carbon_intensity	-81.28823319644539