· Track Width / Track Curvature- Unsure if tile size is increasing (but number of tiles is the same) or number of tiles are increasing (but tile size is the same). Therefore, using track turn rate.
· Linear Velocity / Angular Velocity of car is not constant- only obtained at the end of an evaluation run and dependent on last frame in track. Therefore, using number of tiles visited as a metric- to measure how long car stays on track along with grass time vs total time ratio
· CarRacing Reward Structure- Rewarded for number of tiles covered; and time in which they are covered, penalty for wasting time; going into grass
· CarRacing Obstacles Reward Structure- Rewarded for number of tiles covered; time in which they are covered, penalty for wasting time; going into grass; colliding with obstacles

I. Manual Curricula:

1. CarRacing Original- (Default Track Turn Rate=0.31)

For curriculum learning:

- Training phase: turn rate was increased using varied linear pacing

Track Turn Rate in Curriculum Training Environment	Timesteps/1000
0.31	0-230 (230)
0.41	230-460 (230)
0.51	460-756 (296)
0.61	756-1000 (244)

- Evaluation phase: Best model was chosen based on generalization performance of trained curriculum agent over evaluation environments with random turn rates sampled from [0.31,0.41,0.51,0.61,0.71] after every 50000 training timesteps

2. CarRacing Obstacles- (Default Obs_prob=0.05)

For curriculum learning:

- Training phase: obstacle probability was increased using varied linear pacing

Obstacle Probability in Curriculum Training Environment	Timesteps/1000
0.05	0-120 (120)
0.07	120-240 (120)
0.09	240-660 (420)
0.11	660-1000 (340)

- Evaluation phase: Best model was chosen based on generalization performance of trained curriculum agent over evaluation environments with random obstacle probabilities sampled from [0.05,0.07,0.09,0.11,0.13] after every 50000 training timesteps

3. CarRacing Obstacles- (Default Obs_prob=0.05, Track Turn Rate=0.31)

For curriculum learning:

- Training phase: turn rate and obstacle probability were increased using varied linear pacing

Track Turn Rate in Curriculum Training Environment	Obstacle Probability in Curriculum Training Environment	Timesteps/1000
0.31	0.05	0-198 (198)
0.41	0.07	198-396 (198)
0.51	0.09	396-775 (385)
0.61	0.11	775-1000 (225)

- Evaluation phase: Best model was chosen based on generalization performance of trained curriculum agent over evaluation environments with random turn rates sampled from [0.31,0.41,0.51,0.61,0.71] and random obstacle probabilities sampled from [0.05,0.07,0.09,0.11,0.13] after every 50000 timesteps

II. BO Curricula:

Number of trials: n_warmup=5, N=18-22
Acquisition Function used: Upper Confidence Bound
Used Gaussian Process Surrogate: Matern 5/2 Kernel (due to Quadratic dependencies)

1. CarRacing Original- (Default Track Turn Rate=0.31)

Bayesian Optimization was run for 18 trials
Exploration factor (kappa=2)

For curriculum learning:

- Training phase: turn rate was increased using varied linear pacing (range0: [150,250], range1: [360,460], range2: [660,760])

Track Turn Rate in Curriculum Training Environment	Timesteps/1000
0.31	0-239 (239)
0.41	239-442 (203)
0.51	442-711 (269)
0.61	711-1000 (289)

- Evaluation phase: Best model was chosen based on generalization performance of trained curriculum agent over evaluation environments with random turn rates sampled from [0.31,0.41,0.51,0.61,0.71] after every 50000 training timesteps

2. CarRacing Obstacles- (Default Obs_prob=0.05)

Bayesian Optimization was run for 22 trials
Exploration factor (kappa=3.1)

For curriculum learning:

- Training phase: obstacle probability was increased using varied linear pacing (range0: [100,150], range1: [220,260], range2: [630,700])

Obstacle Probability in Curriculum Training Environment	Timesteps/1000
0.05	0-101 (101)
0.07	101-238 (137)
0.09	238-670 (432)
0.11	670-1000 (330)

- Evaluation phase: Best model was chosen based on generalization performance of trained curriculum agent over evaluation environments with random obstacle probabilities sampled from [0.05,0.07,0.09,0.11,0.13] after every 50000 training timesteps

3. CarRacing Obstacles- (Default Obs_prob=0.05, Track Turn Rate=0.31)

Bayesian Optimization was run for 18 trials
Exploration factor (kappa=1.9)

For curriculum learning:

- Training phase: turn rate and obstacle probability were increased using varied linear pacing (range0: [150,250], range1: [330,450], range2: [730,830])

Track Turn Rate in Curriculum Training Environment	Obstacle Probability in Curriculum Training Environment	Timesteps/1000
0.31	0.05	0-161 (161)
0.41	0.07	161-418 (257)
0.51	0.09	418-737 (319)
0.61	0.11	737-1000 (263)

- Evaluation phase: Best model was chosen based on generalization performance of trained curriculum agent over evaluation environments with random turn rates sampled from [0.31,0.41,0.51,0.61,0.71] and random obstacle probabilities sampled from [0.05,0.07,0.09,0.11,0.13] after every 50000 timesteps

Results: (All averaged over 500 testing/evaluation environments due to more randomness)-

1. PPO Performance with different training regimes (CarRacing Original).
   - Default Environment has a Track Turn Rate of 0.31
   - Hard Environment has random Track Turn Rates in [0.31,0.41,0.51,0.61,0.71]

Training Environment	Testing Environment	Values	Number of tiles visited	Time on grass ratio
Default	Default	724 ± 283	238	0.168
Default	Hard	667 ± 305	221	0.227
Manual Curriculum	Hard	682 ± 121	227	0.157
BO Curriculum	Hard	793 ± 118	259	0.042

2. PPO Performance with different training regimes (CarRacing Obstacles).
   - Default Environment has an Obstacle Probability of 0.05- Average of about 5 obstacles
   - Hard Environment has random Obstacle Probabilities in [0.05,0.07,0.09,0.11,0.13]

Training Environment	Testing Environment	Values	Collision/ Obstacle Ratio	Number of tiles visited	Time on grass ratio	Collisions
Default	Default	422 ± 185	0.166	168	0.322	1.002
Default	Hard	386 ± 179	0.183	163	0.336	1.418
Manual Curriculum	Hard	664 ± 134	0.120	236	0.031	0.932
BO Curriculum	Hard	696 ± 200	0.105	245	0.1246	0.84

3. PPO Performance with different training regimes (CarRacing Obstacles)
   - Default Environment has a Track Turn Rate of 0.31 and Obstacle Probability of 0.05
   - Hard Environment has random Track Turn Rates in [0.31,0.41,0.51,0.61,0.71] and random Obstacle Probabilities in [0.05,0.07,0.09,0.11,0.13]

Training Environment	Testing Environment	Values	Collision/ Obstacle Ratio	Number of tiles visited	Time on grass ratio	Collisions
Default	Default	422 ± 185	0.166	168	0.322	1.002
Default	Hard	368 ± 181	0.185	159	0.343	1.514
Manual Curriculum	Hard	667 ± 167	0.055	230	0.082	0.454
BO Curriculum	Hard	696 ± 113	0.133	248	0.032	1.088

Performance Comparison across Levels

Defaults:
Turnrates:
729 ± 282 (0.31), 686 ± 311 (0.41), 662 ± 314 (0.51), 669 ± 310 (0.61), 659 ± 309 (0.71)
Obstacles:
400 ± 180 (0.05), 393 ± 165 (0.07), 393 ± 169 (0.09), 369 ± 185 (0.11), 364 ± 158 (0.13)
Both:
400 ± 180 (0.31,0.05), 369 ± 190 (0.41,0.07), 341 ± 171 (0.51,0.09), 316 ± 195 (0.61,0.11), 337 ± 159 (0.71,0.13)

Manual Curricula:
Turnrates:
687 ± 131 (0.31), 682 ± 131 (0.41), 681 ± 118 (0.51), 687 ± 132 (0.61), 690 ± 125 (0.71)
Obstacles:
685 ± 112 (0.05), 660 ± 131 (0.07), 667 ± 117 (0.09), 658 ± 130 (0.11), 652 ± 135 (0.13)
Both:
688 ± 164 (0.31,0.05), 676 ± 168 (0.41,0.07), 654 ± 176 (0.51,0.09), 643 ± 172 (0.61,0.11), 628 ± 181 (0.71,0.13)

BO Curricula:
Turnrates:
792 ± 138 (0.31), 798 ± 124 (0.41), 801 ± 110 (0.51), 784 ± 134 (0.61), 785 ± 123 (0.71)
Obstacles:
704 ± 226 (0.05), 711 ± 197 (0.07), 693 ± 199 (0.09), 690 ± 201 (0.11), 687 ± 193 (0.13)
Both:
748 ± 87 (0.31,0.05), 696 ± 146 (0.41,0.07), 685 ± 128 (0.51,0.09), 662 ± 126 (0.61,0.11), 659 ± 138 (0.71,0.13)