Competition for novelty in an information-sampling game

Leonid Tiokhin, Maxime Derex

Open Science 2019

Trier, Germany

American
SOCIOLOGICAL
Review

Official Journal of the American Sociological Society

PRIORITIES IN SCIENTIFIC DISCOVERY: A CHAPTER IN THE SOCIOLOGY OF SCIENCE *

ROBERT K. MERTON
Columbia University

"In biology, it is the long-standing practice to append the name of the first describer to the name of a species, a custom which greatly agitated Darwin since, as he saw it, this put "a premium on hasty and careless work" as the "species mongers" among naturalists try to achieve an easy immortality by "miserably describing a species in two or three lines."

(Merton, 1957, p. 644)

“Scientific research can be a cutthroat business, with undue pressure to publish quickly, first, and frequently. The resulting race to publish ahead of competitors is intense and to the detriment of the scientific endeavor…we are formalizing a policy whereby manuscripts that confirm or extend a recently published study (“scooped” manuscripts, also referred to as complementary) are eligible for consideration at PLOS Biology.”

(The PLOS Biology Staff Editors, 2018)

eLIFE
elifesciences.org

EDITORIAL

SCIENTIFIC PUBLISHING

Beyond scoops to best practices

Authors submitting a manuscript to eLife are encouraged to upload it to a recognized preprint server at the same time in order to make their results available as quickly and as widely as possible.

EVE MARDER

PLOS | BIOLOGY

EDITORIAL

The importance of being second

The *PLOS Biology* Staff Editors*

Public Library of Science, San Francisco, California, United States of America

* plosbiology@plos.org

Registered Report:

Competition for novelty in an information-sampling game.

Royal Society Open Science (Stage 2 Review).

UNIVERSITY OF EXETER

Experimental Design

Confirmatory Hypotheses

Criteria for Evaluating Evidence for/against Hypotheses

Inclusion / Exclusion Criteria

Theoretical Model

Statistical Models

Pilot Study + Quality Checks

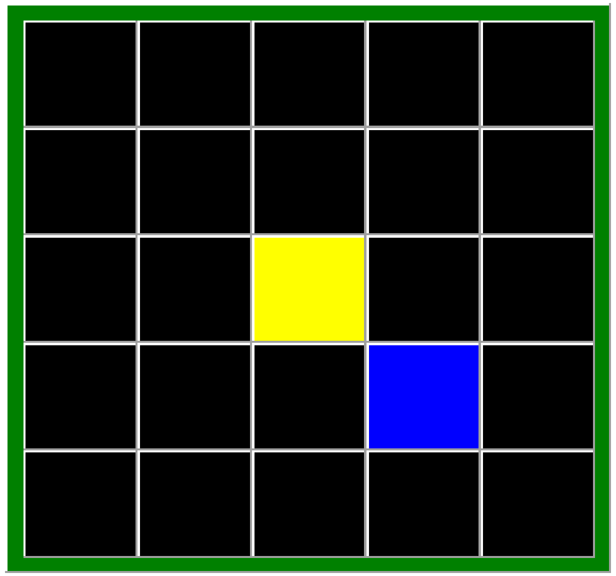
Data-Collection Stopping Rules

Power Analysis

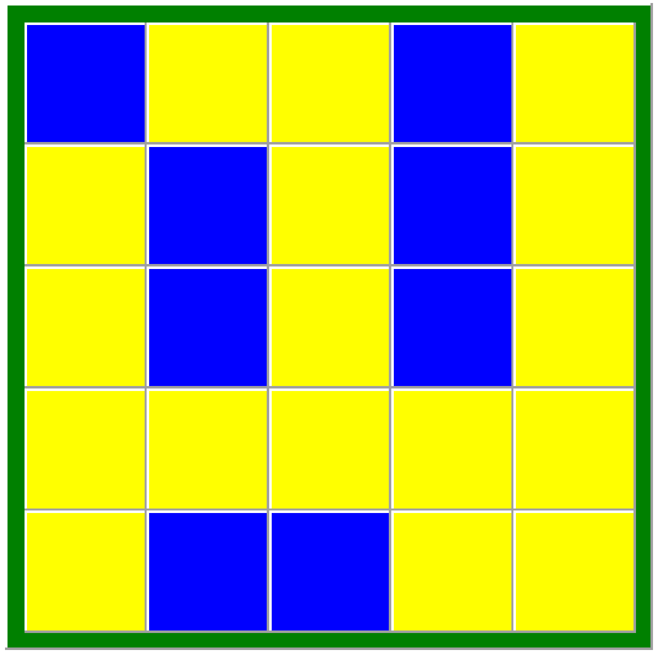
PREREGISTERED

Puzzle

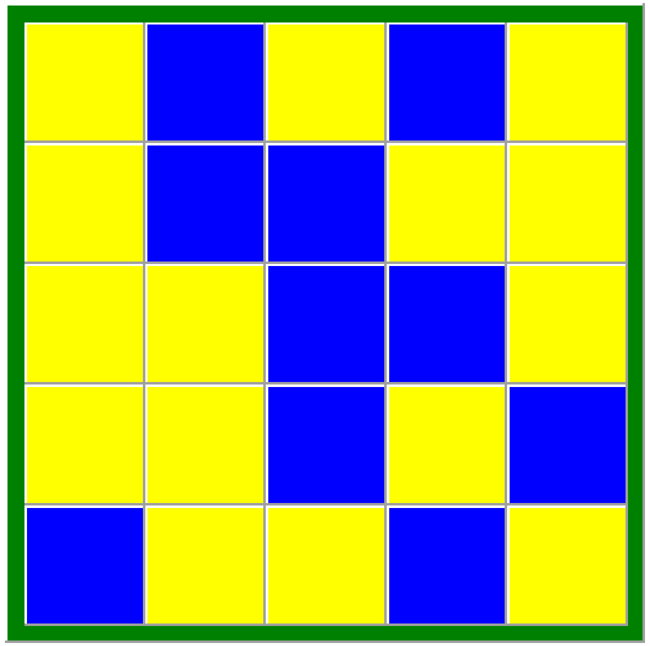
Score 0



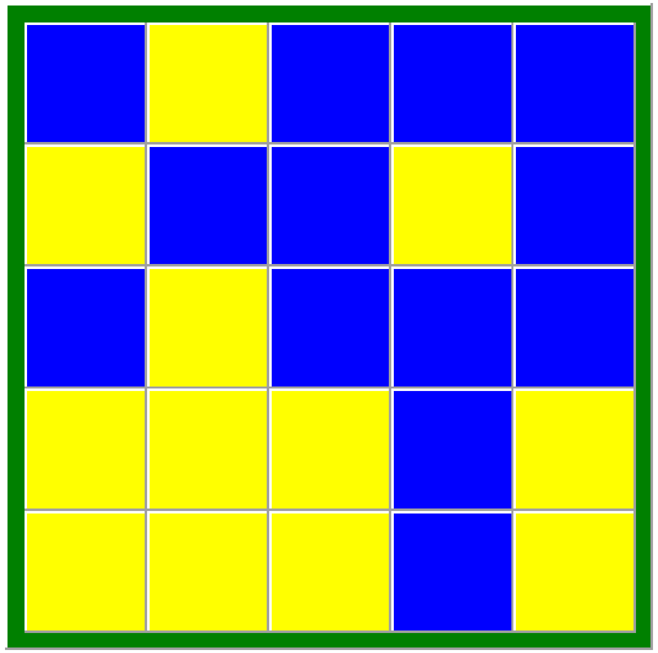
Which color is the most common?



Large Effect



Medium Effect



Small Effect

260 students from Arizona State University (130 M / 130 F).

85 – 99% Statistical Power to detect effects.

Participants fully informed of payoff structure and told that their earnings depend on their performance.

Tile-sequences deterministic.

	No Comp	Comp
No Effort		
Effort		

	No Comp	Comp
No Effort		
Effort		

Experiment lasts 20 minutes.

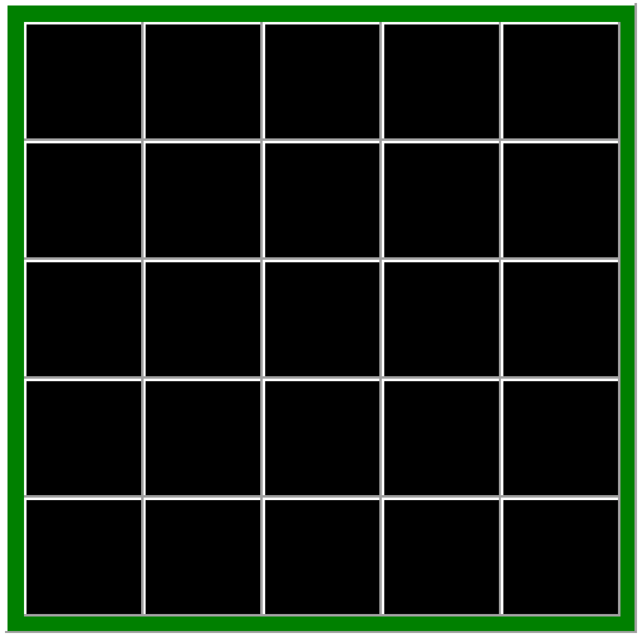
Players can reveal 1 tile every 1 second.

+1 for correct guesses.

-1 for incorrect guesses.

Puzzle

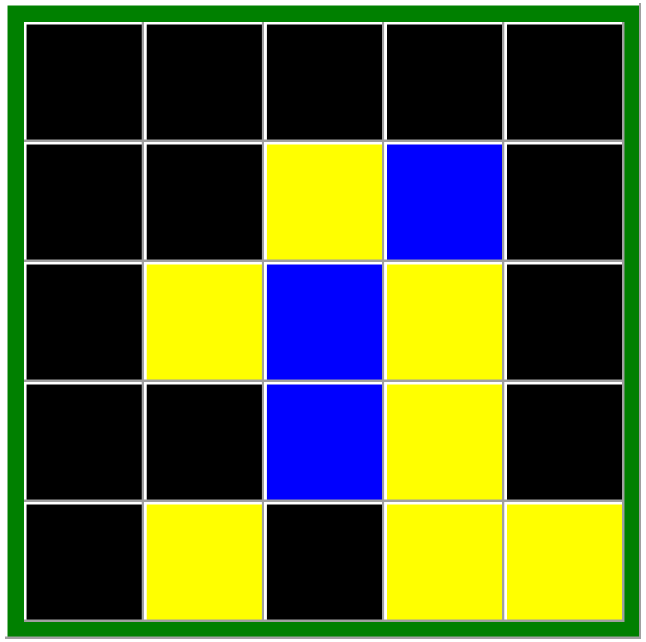
Score 1



Which color is the most common?

Puzzle

Score 1



Which color is the most common?

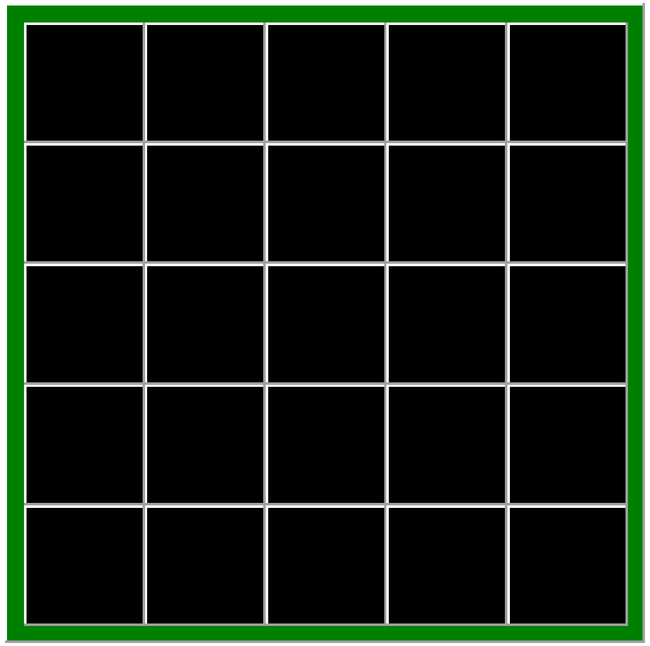
Correct! You gained 1 point.

Please wait for the next problem

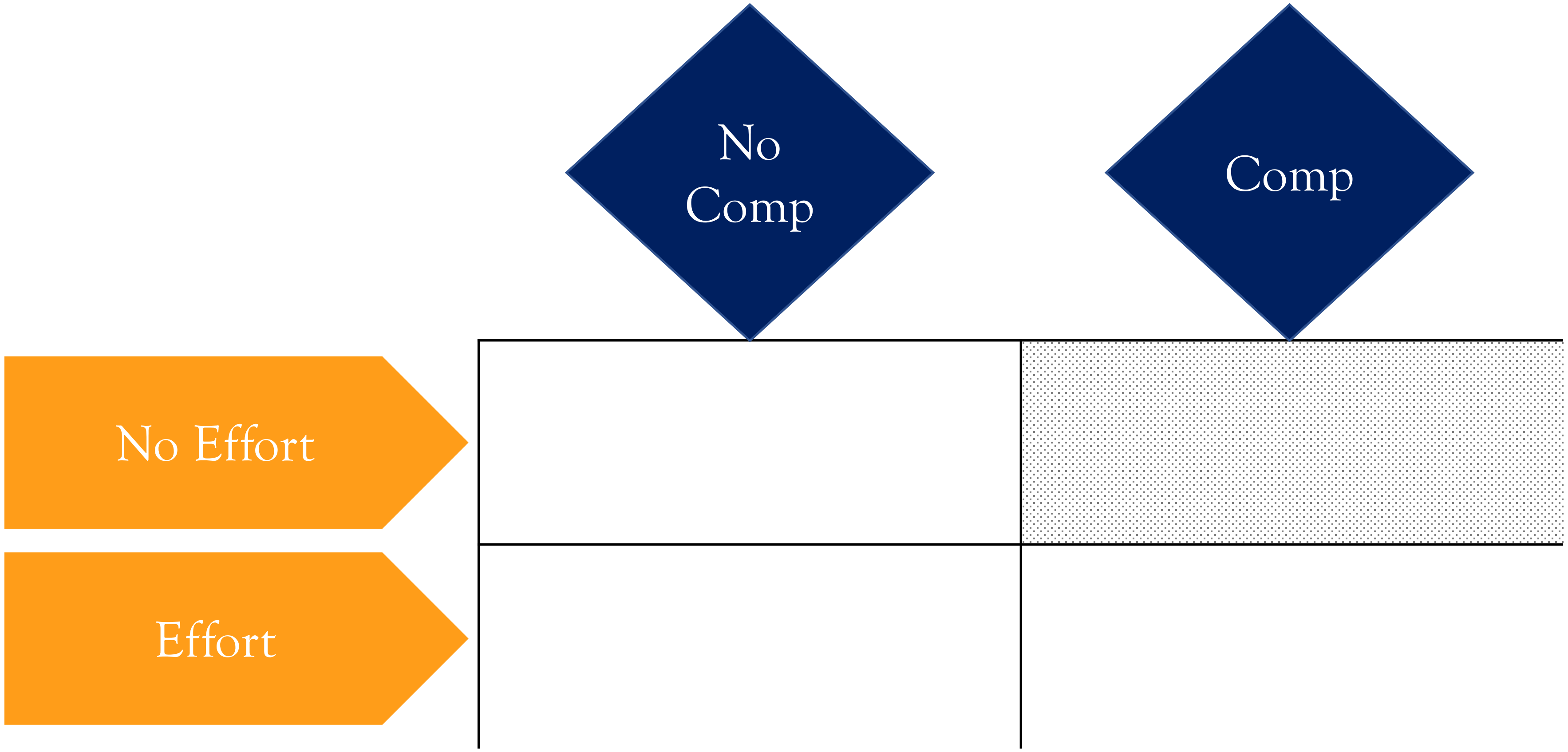
.

Puzzle

Score 2



Which color is the most common?


No Comp
Comp
No Effort
Effort

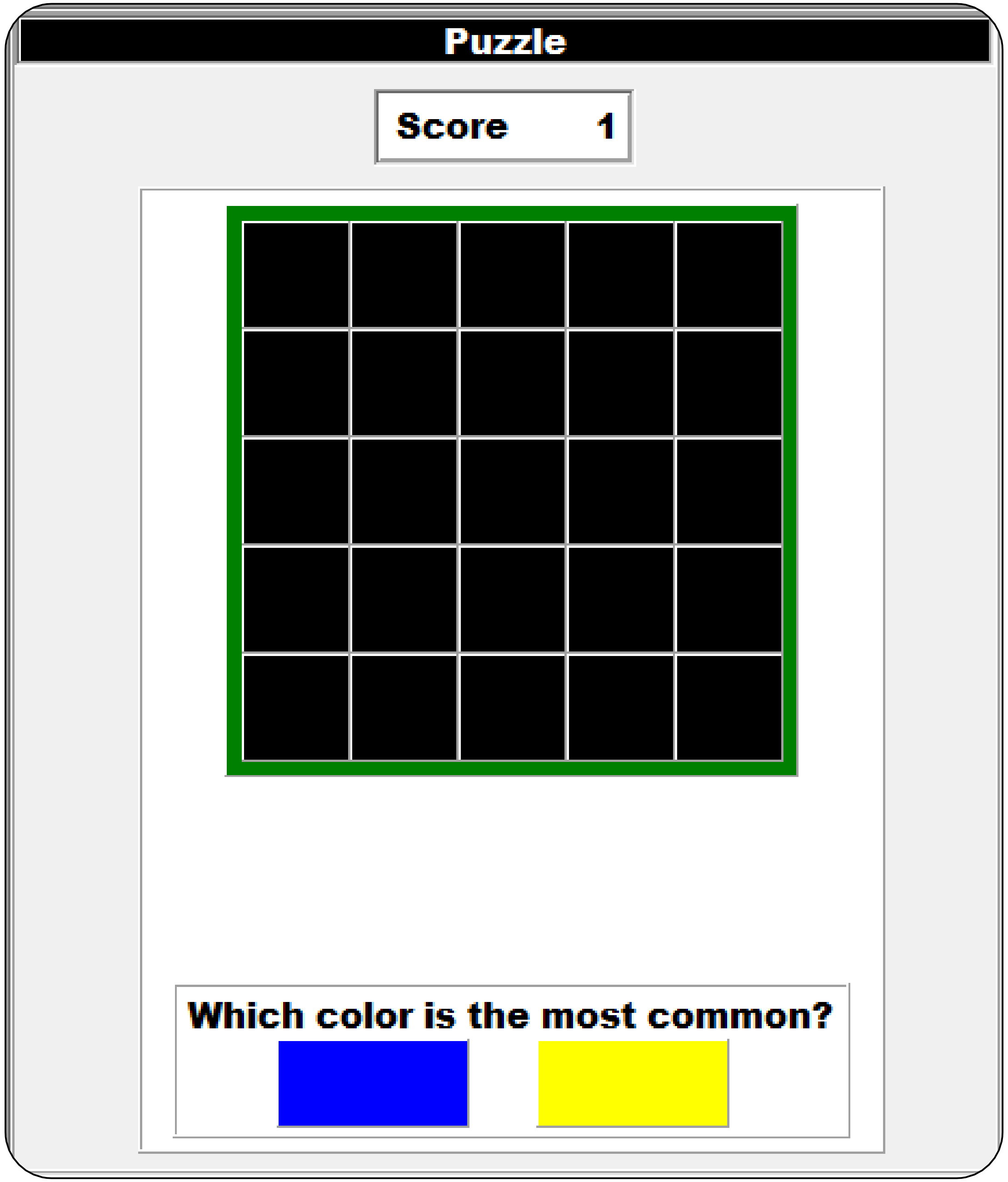
Experiment lasts as long as it takes players to complete the same grids as their opponent.

Players can reveal 1 tile every 1 second.

+1 for correct guesses.

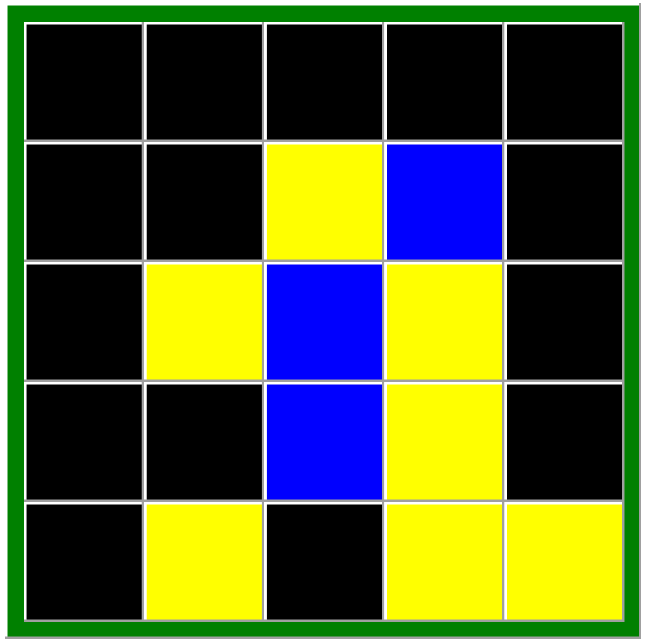
-1 for incorrect guesses.

0 points when scooped.


Puzzle
Score 1
Which color is the most common?

Puzzle

Score 1

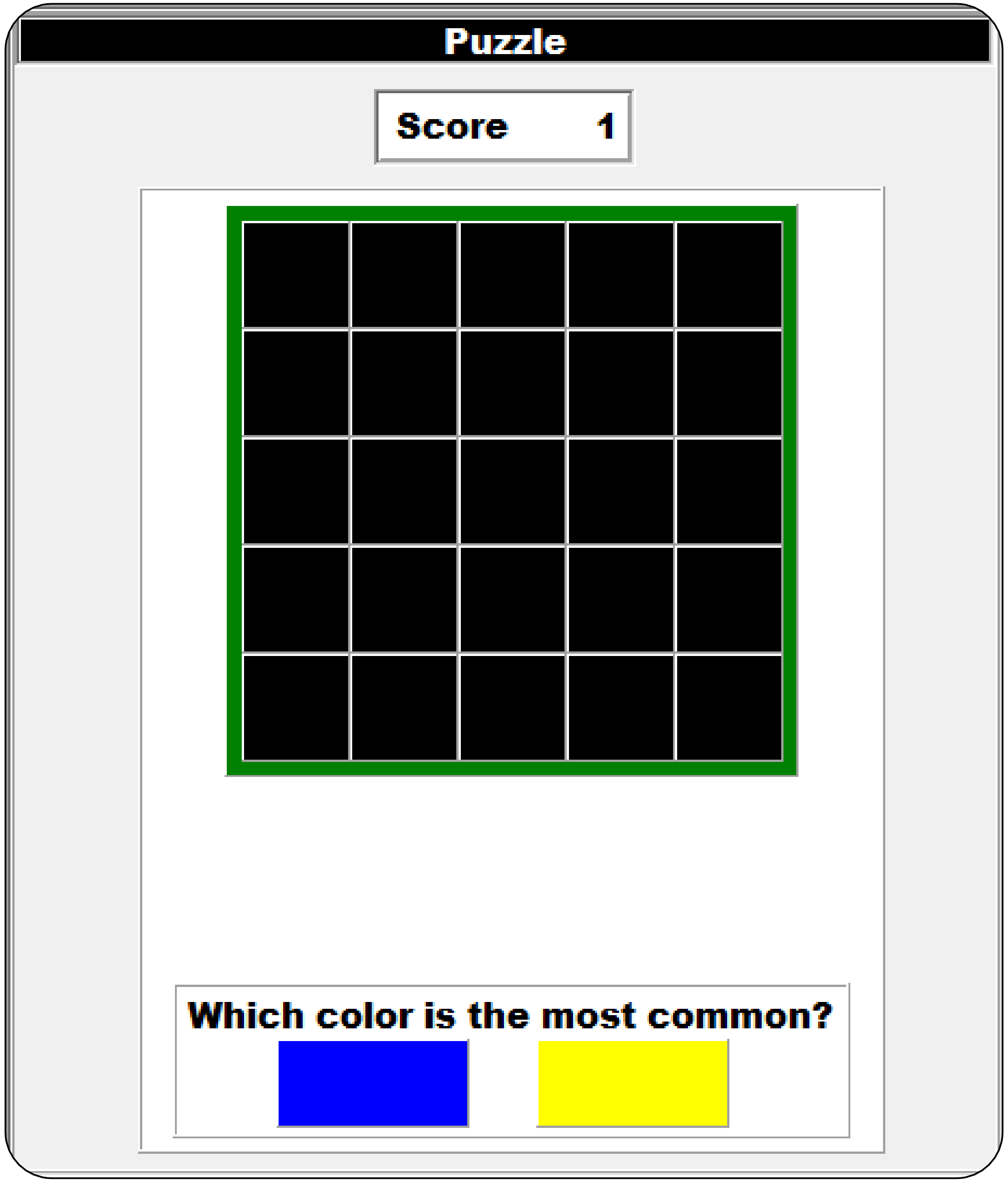


Which color is the most common?

Correct. However your competitor already solved the problem, so you did not gain any points.

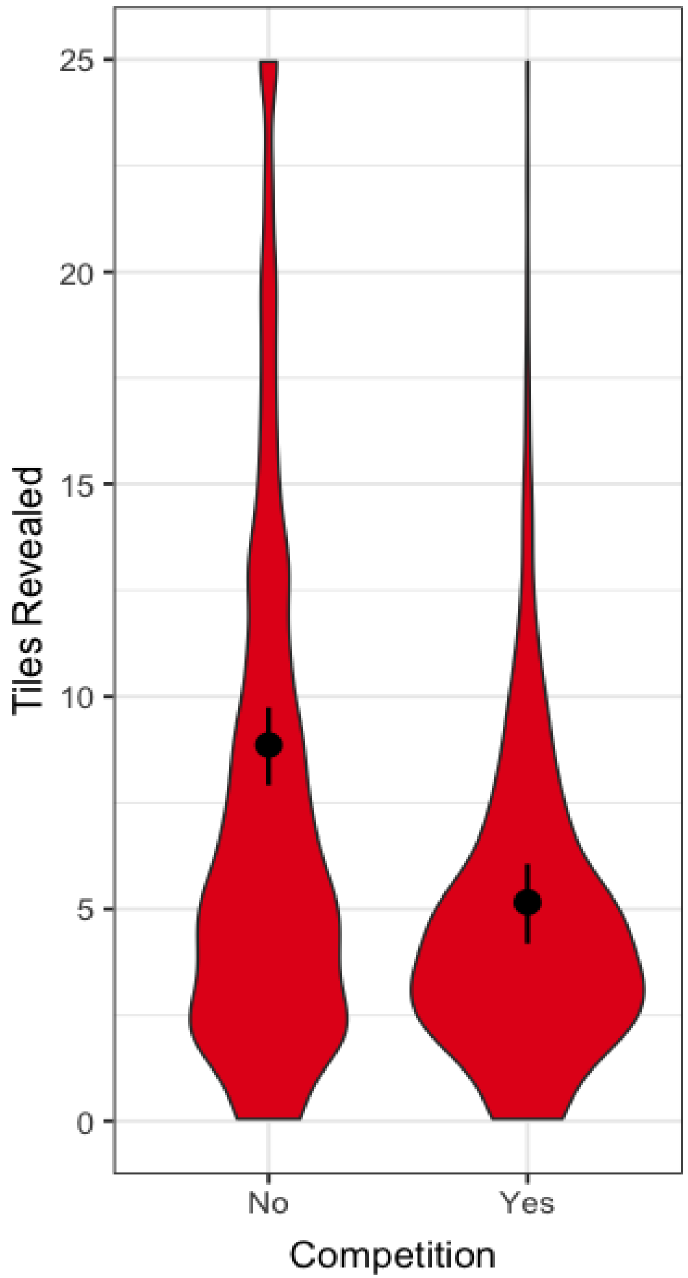
Please wait for the next problem

.


Puzzle
Score 1
Which color is the most common?

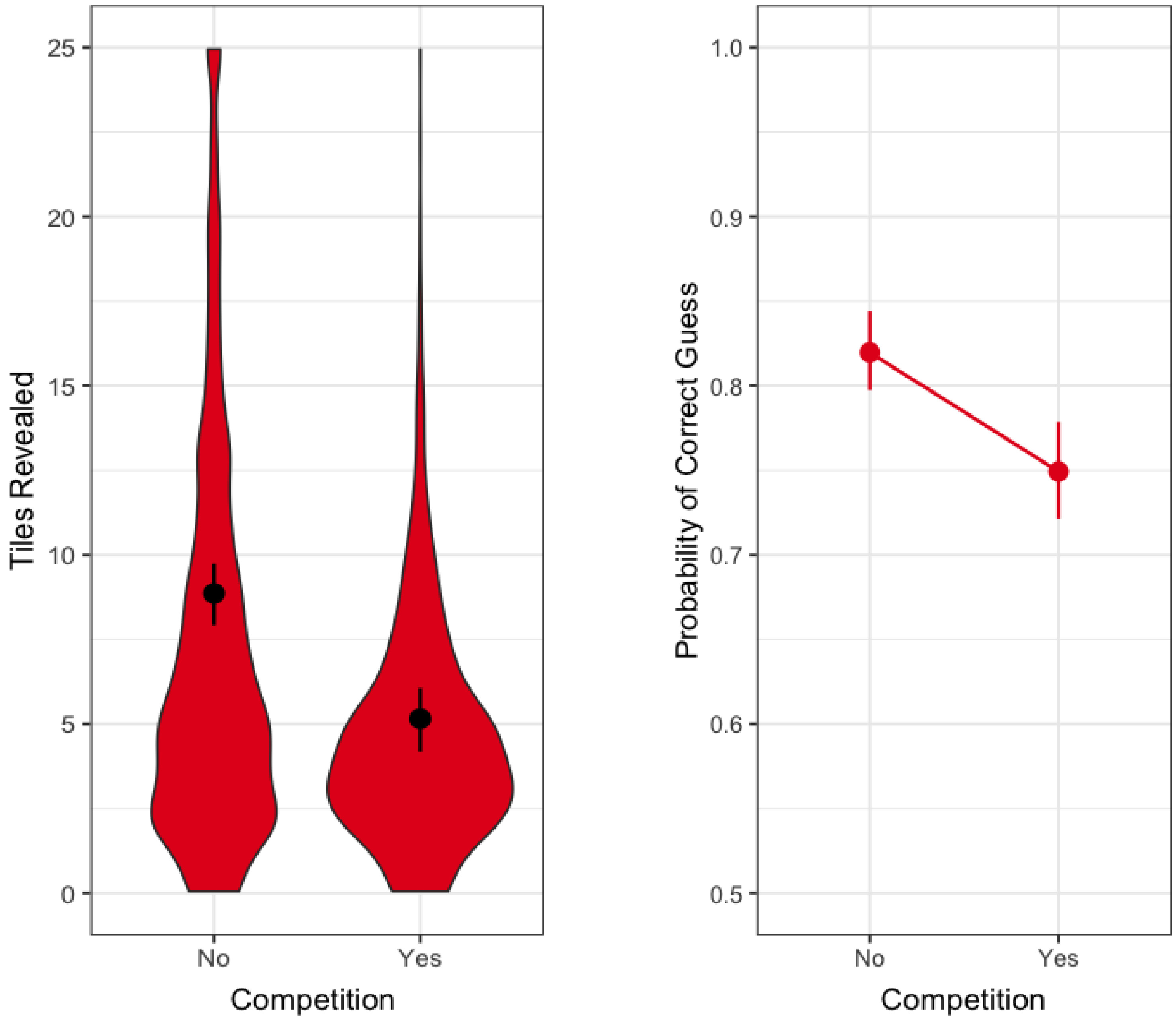
Hypothesis 1:

Competition will cause players to guess with less information and be less accurate



β = -3.70

95% HPDI: (-5.03, -2.39)



β = -3.70

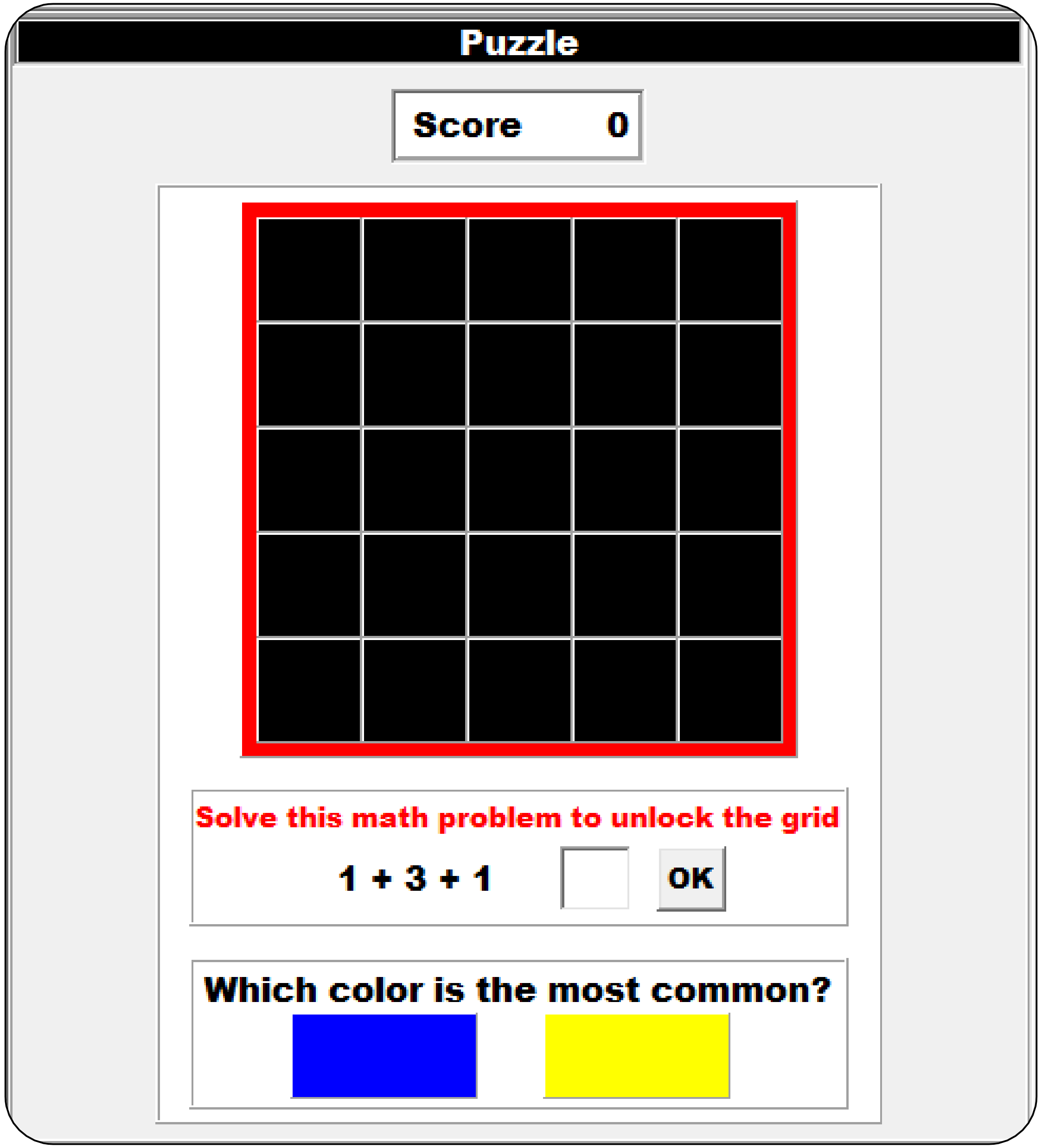
95% HPDI: (-5.03, -2.39)

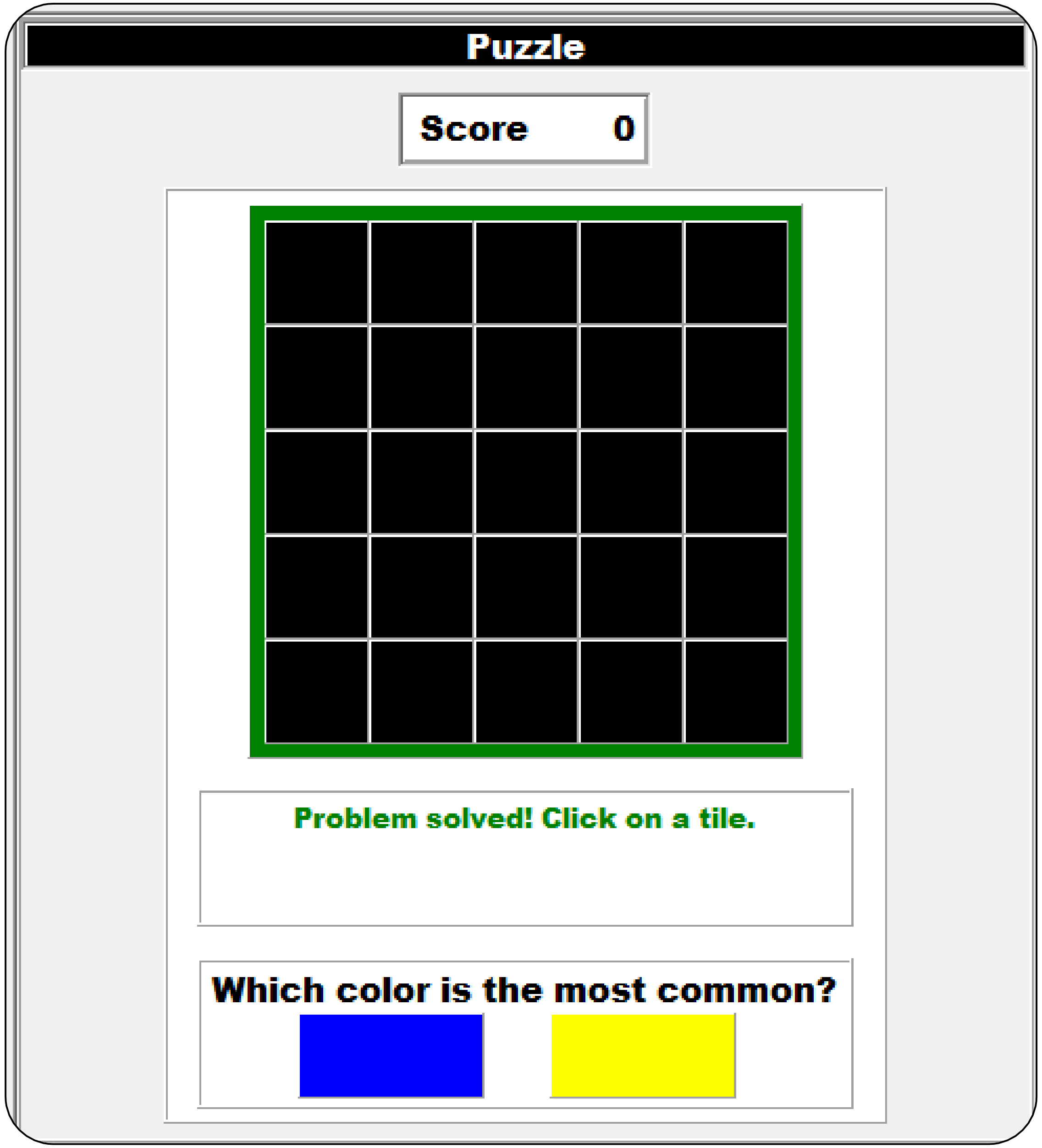
β = -0.07

95% HPDI: (-0.11, -0.03)

What if players can potentially improve research efficiency by increasing effort?

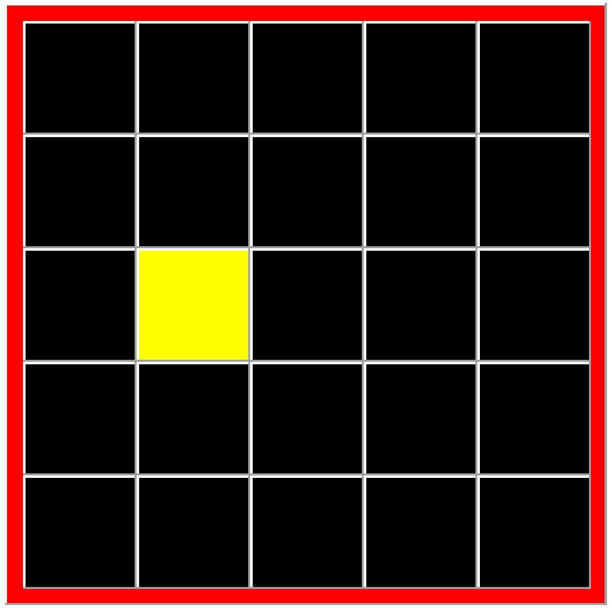
	No Comp	Comp
No Effort		
Effort		


Puzzle
Score 0
Solve this math problem to unlock the grid
1 + 3 + 1
OK
Which color is the most common?


Puzzle
Score 0
Problem solved! Click on a tile.
Which color is the most common?

Puzzle

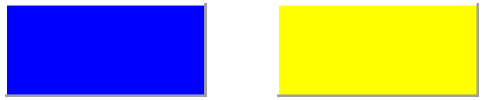
Score 0



Solve this math problem to unlock the grid

4 + 1 + 7 OK

Which color is the most common?



Hypothesis 2:

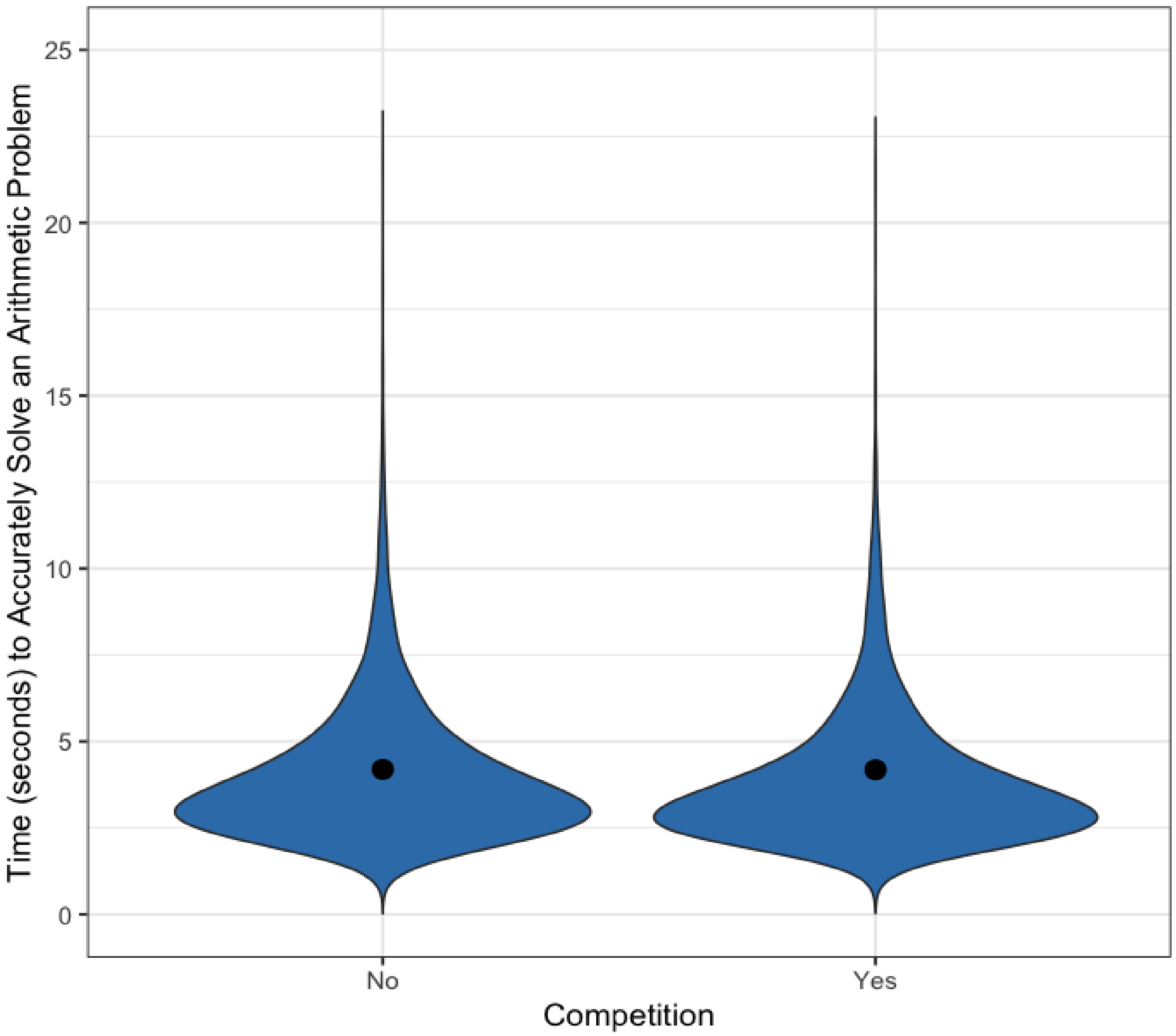
When players can potentially increase research efficiency by adjusting effort (i.e. arithmetic-problem solving speed), competition will cause players to increase effort.

Hypothesis 2:

When players can potentially increase research efficiency by adjusting effort (i.e. arithmetic-problem solving speed), competition will cause players to increase effort.

Hypothesis 3:

Competition x Effort Interaction: competition will have smaller effects on tiles revealed in the Effort conditions compared to the No-Effort conditions.



β = -0.02

95% HPDI: (-0.40, 0.35)

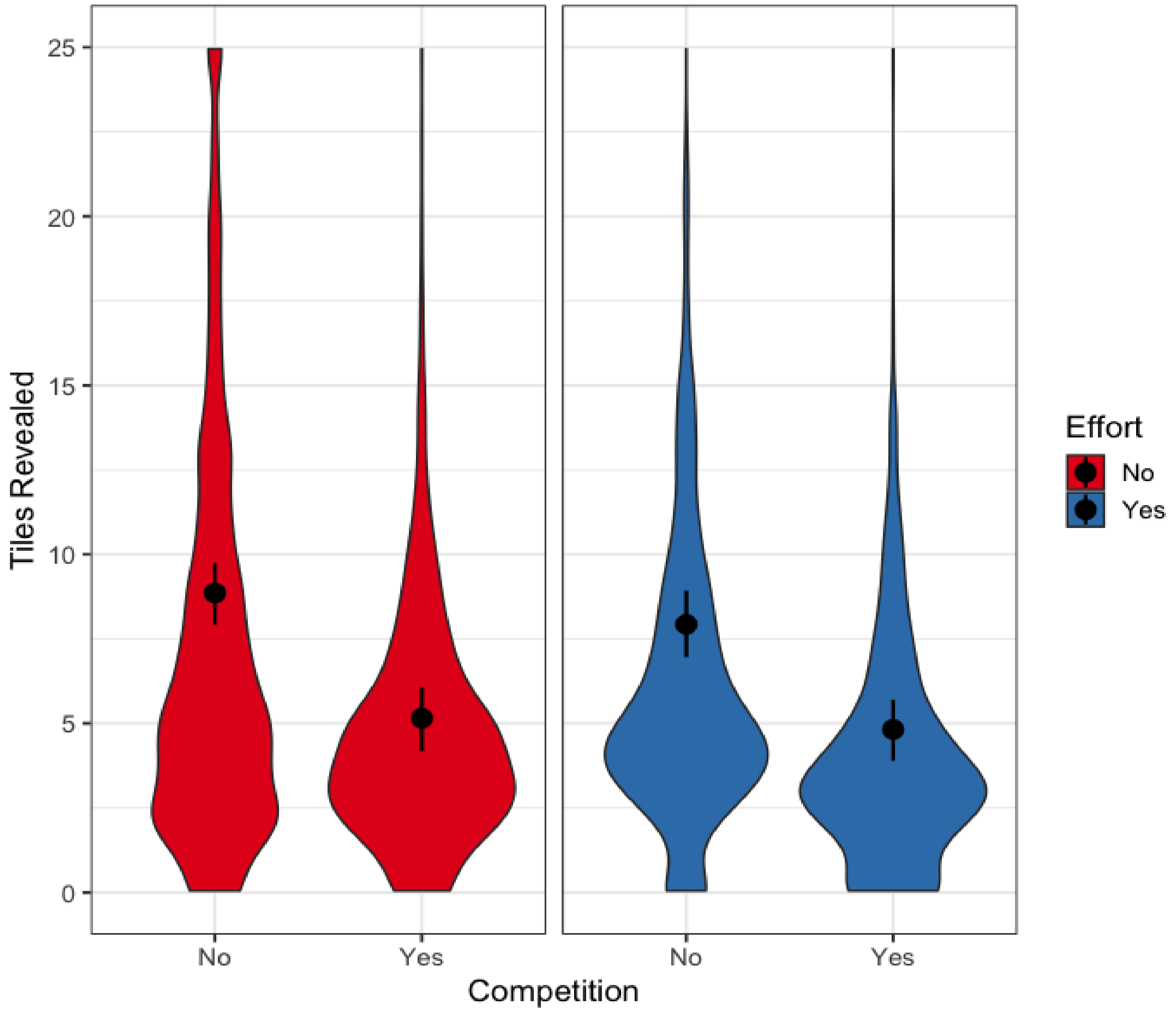
Model Comparison: Arithmetic Problems

EFFORT: TIME TO ACCURATELY SOLVE AN ARITHMETIC PROBLEM

	WAIC	dWAIC	weight	SE
Effort_C	78912.99	0.0	0.6	506.26
Effort_Intercept_Only	78913.79	0.8	0.4	506.25

Bayes Factors: Comp vs. No Comp

BIC Max N (20098 Observations) BF01 = 142
BIC Min N (130 Observations) BF01 = 11



NE Comp. β = -3.70

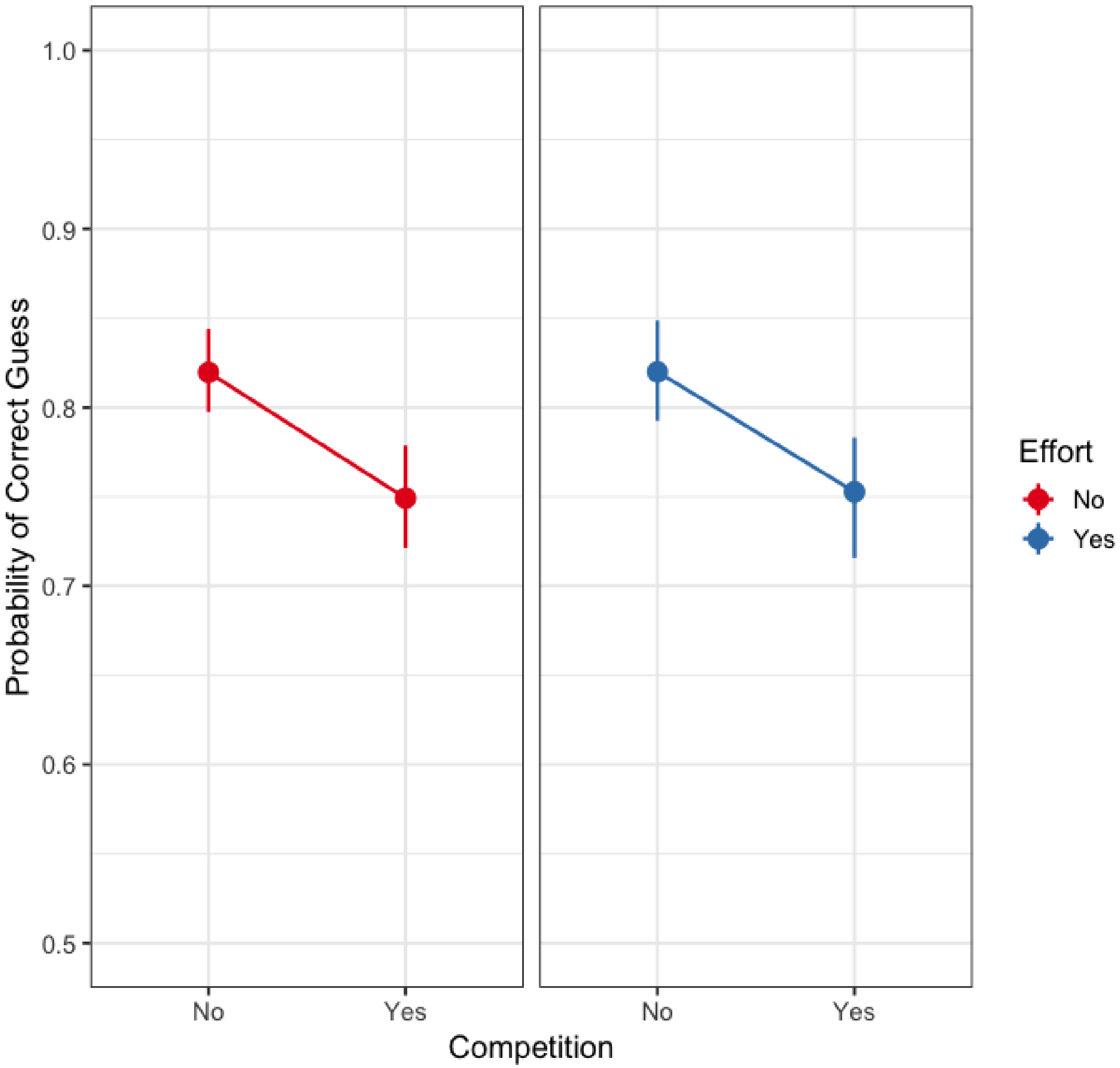
95% HPDI: (-5.03, -2.39)

E Comp β = -3.11

95% HPDI: (-4.47, -1.86)

Comp X E Interaction β = 0.60

95% HPDI: (-1.23, 2.54)



NE Comp. β = -0.07

95% HPDI: (-0.11, -0.03)

E Comp β = -0.07

95% HPDI: (-0.11, -0.02)

Comp X E Interaction β = 0.01

95% HPDI: (-0.07, 0.10)

Model Comparison: Tiles Revealed

TILES REVEALED

	WAIC	dWAIC	weight	SE
Tiles_E_C	69826.83	0.00	0.55	331.09
Tiles_C	69828.53	1.69	0.24	331.20
Tiles_E_C_EC	69828.74	1.91	0.21	331.06

Bayes Factors: Interaction vs. No Interaction

BIC Max N (14073 Observations) BF01 = 119

BIC Min N (260 Observations) BF01 = 16

Model Comparison: Accuracy

ACCURACY

	WAIC	dWAIC	weight	SE
Accuracy_C	14646.12	0.00	0.49	114.96
Accuracy_E_C	14647.03	0.91	0.31	114.98
Tiles_E_C_EC	14647.89	1.77	0.20	115.02

Bayes Factors: Interaction vs. No Interaction

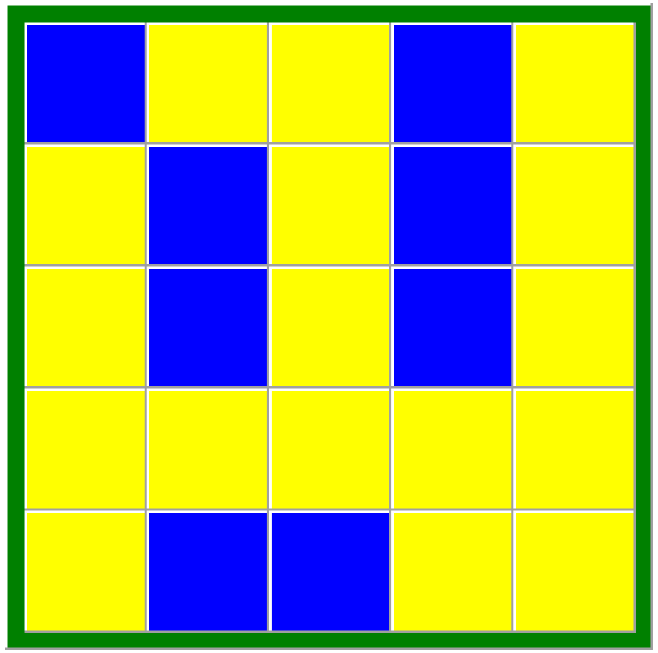
BIC Max N (14073 Observations) BF01 = 119

BIC Min N (260 Observations) BF01 = 16

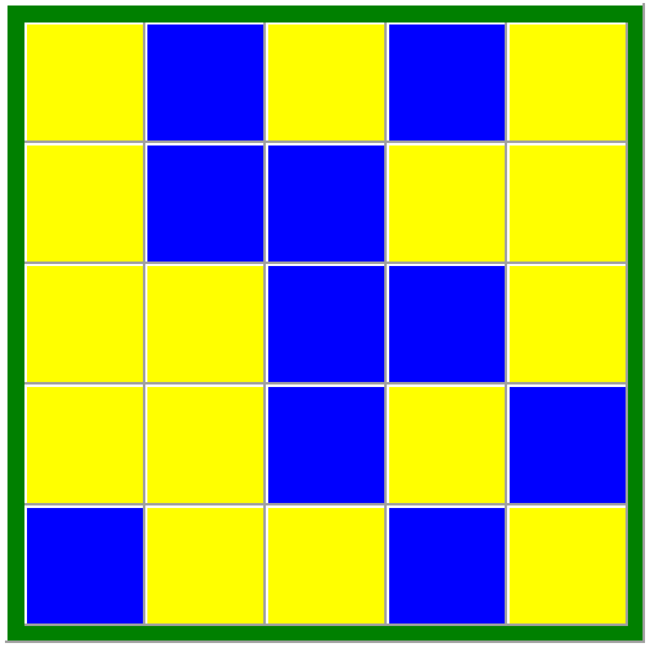
Takeaways:

Competition for novel results causes players to make guesses using less information and have reduced accuracy.

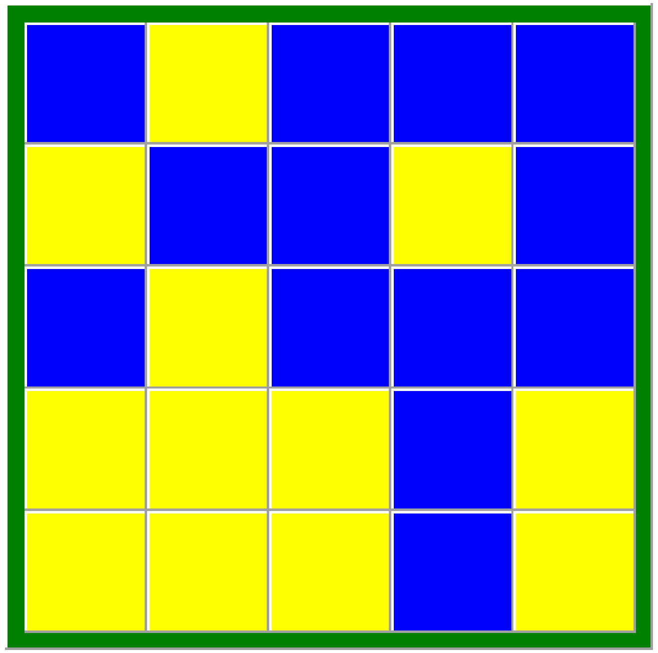
When players can potentially beat competitors by increasing effort, players do not increase effort but still reduce sample size*.



Large Effect

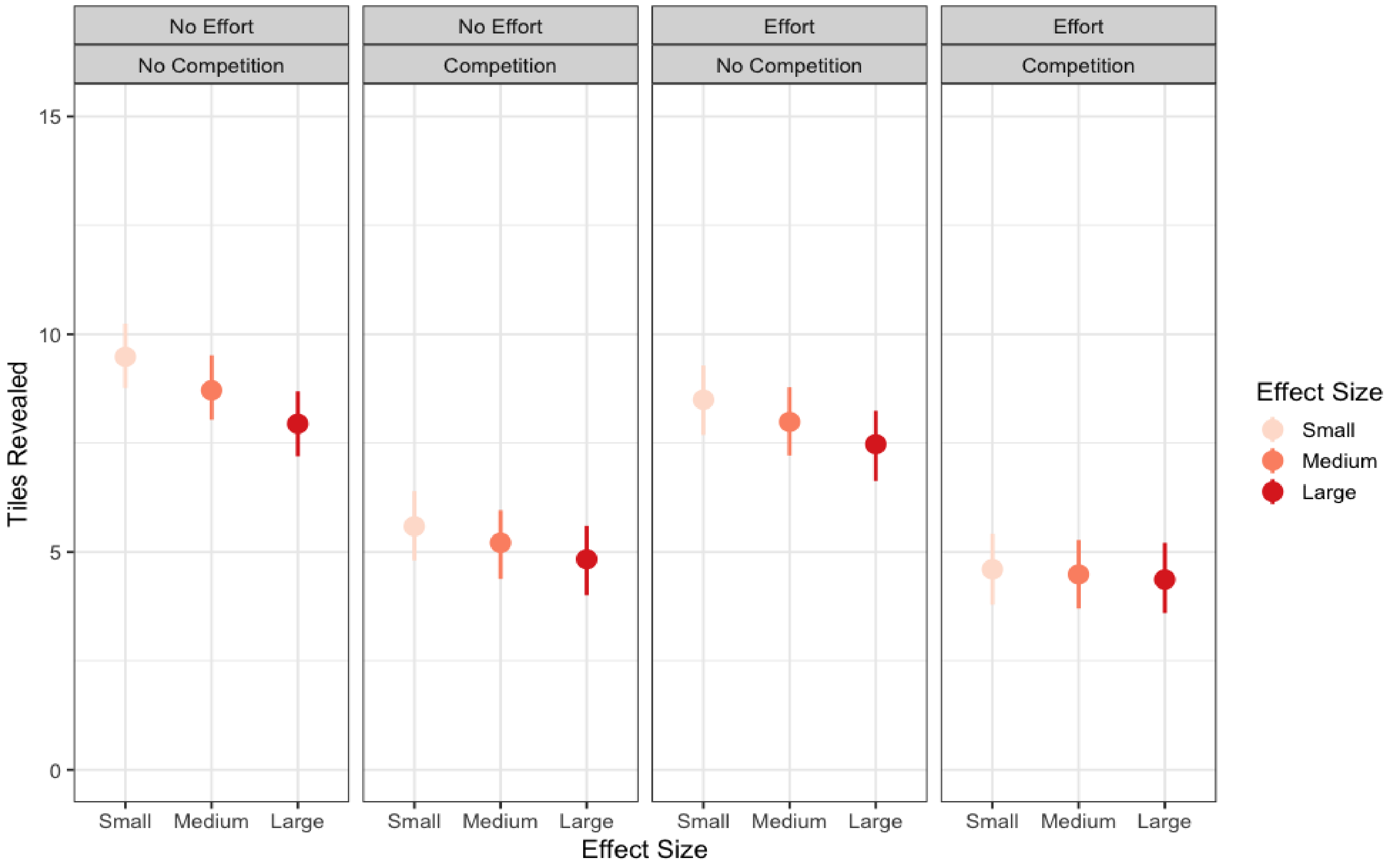


Medium Effect

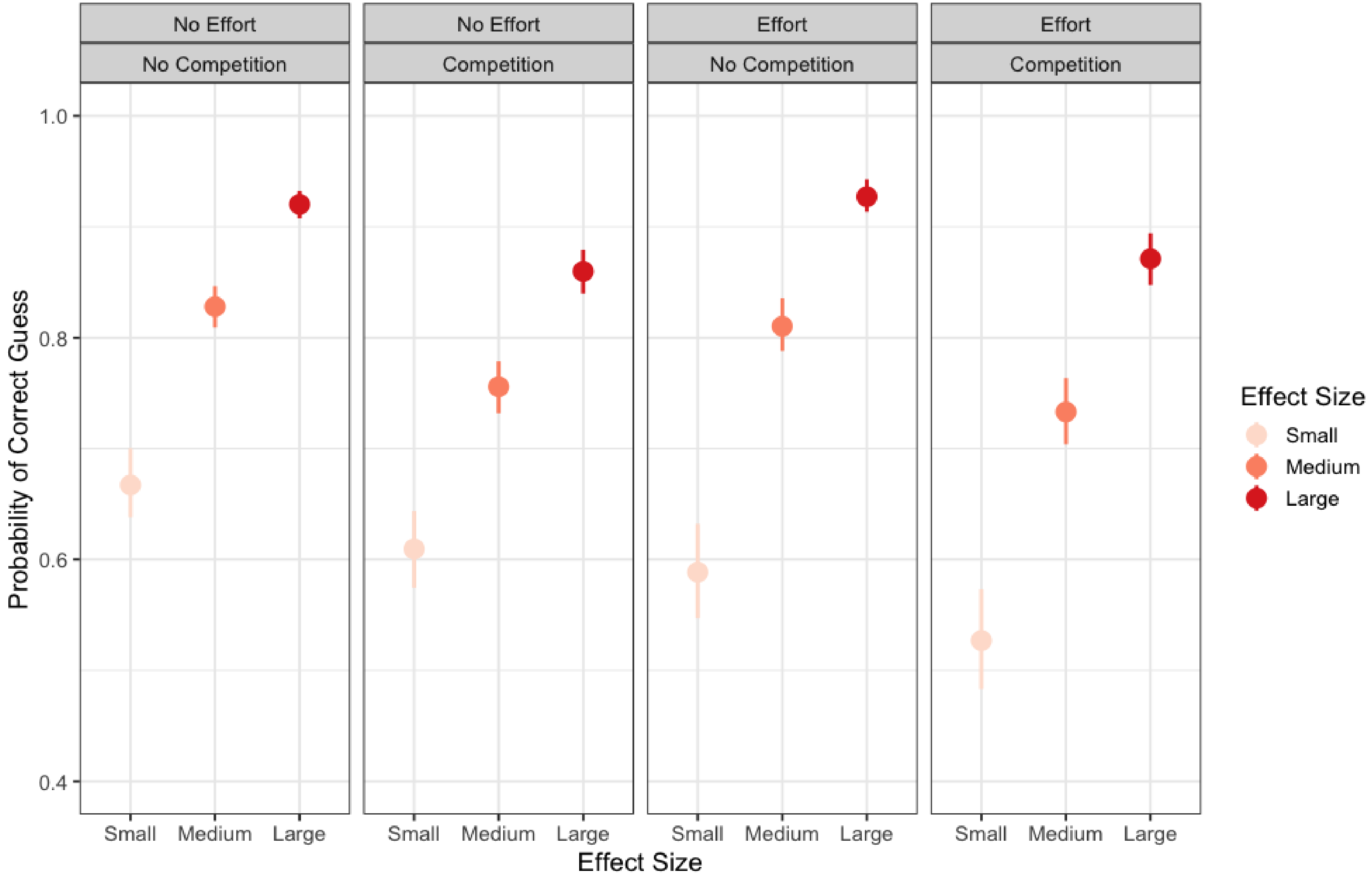


Small Effect

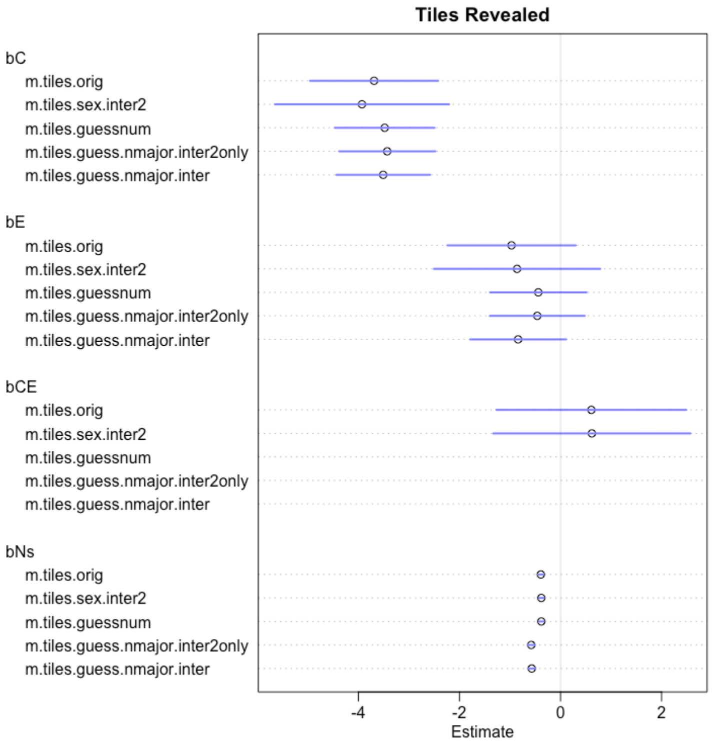
Tiles revealed as a function of effect size

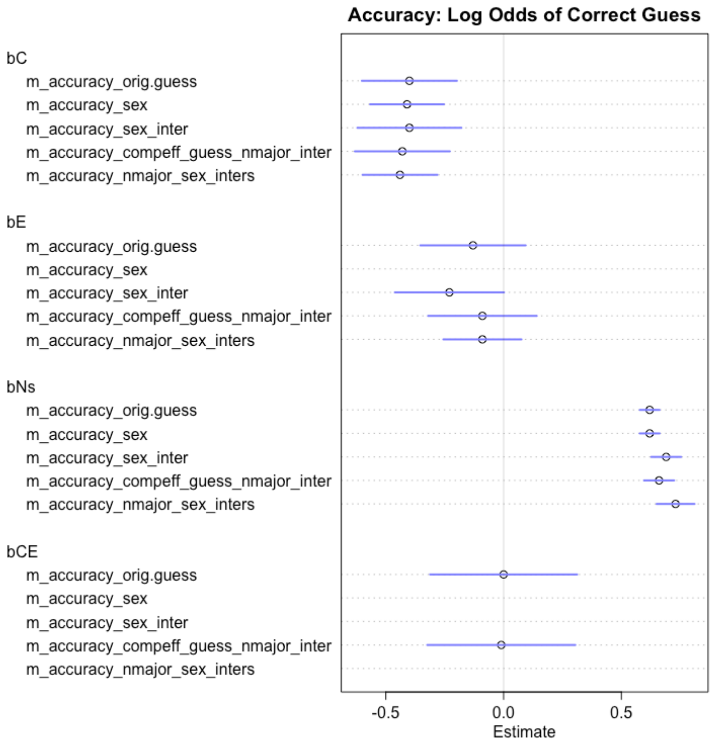


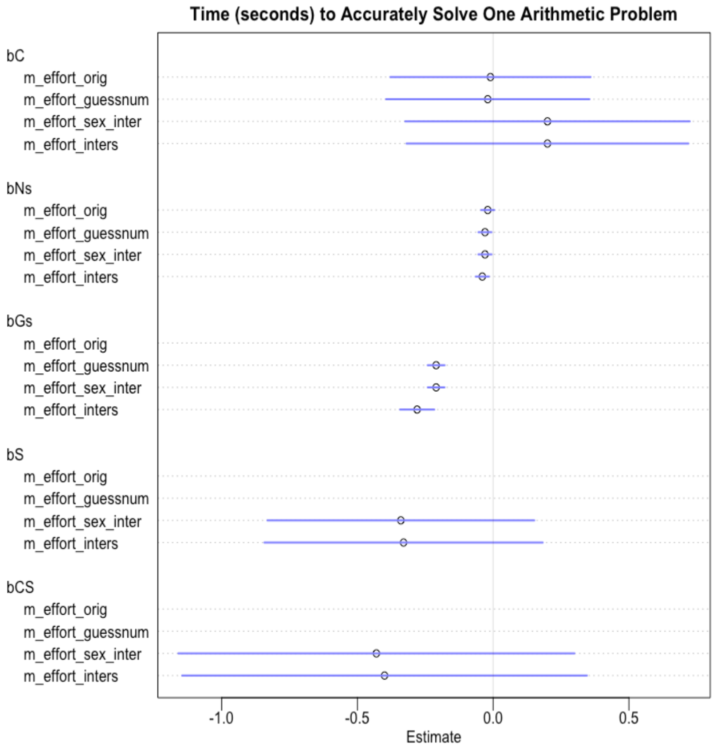
Accuracy as a function of effect size



Sensitivity Checks

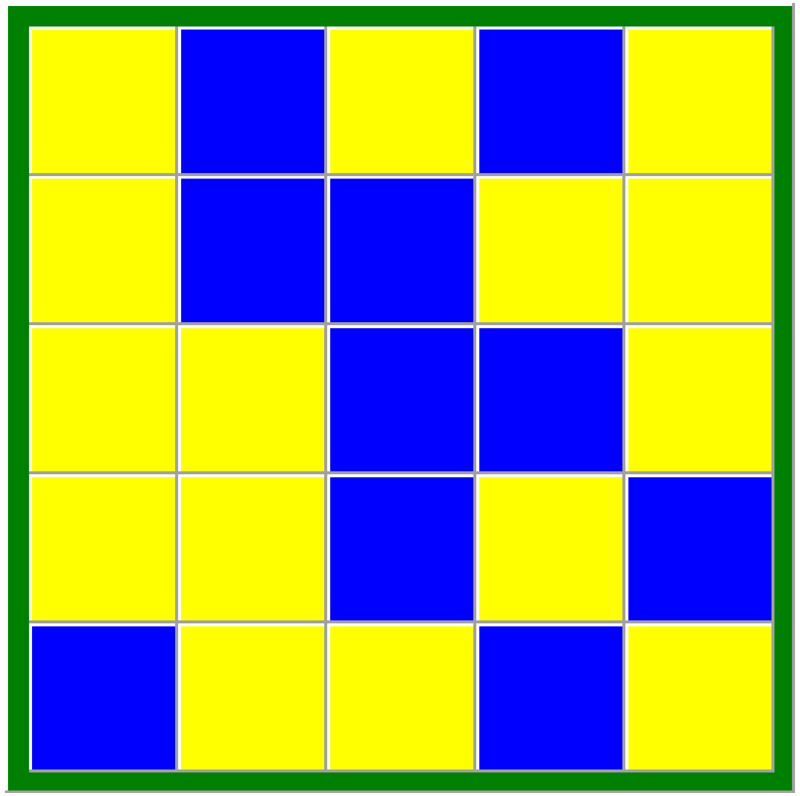






Limitations

- Lacks multi-sided strategic interaction.
- Effort null-effect generalizable?
- Models science as gathering information on independent, well-defined problems.



Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

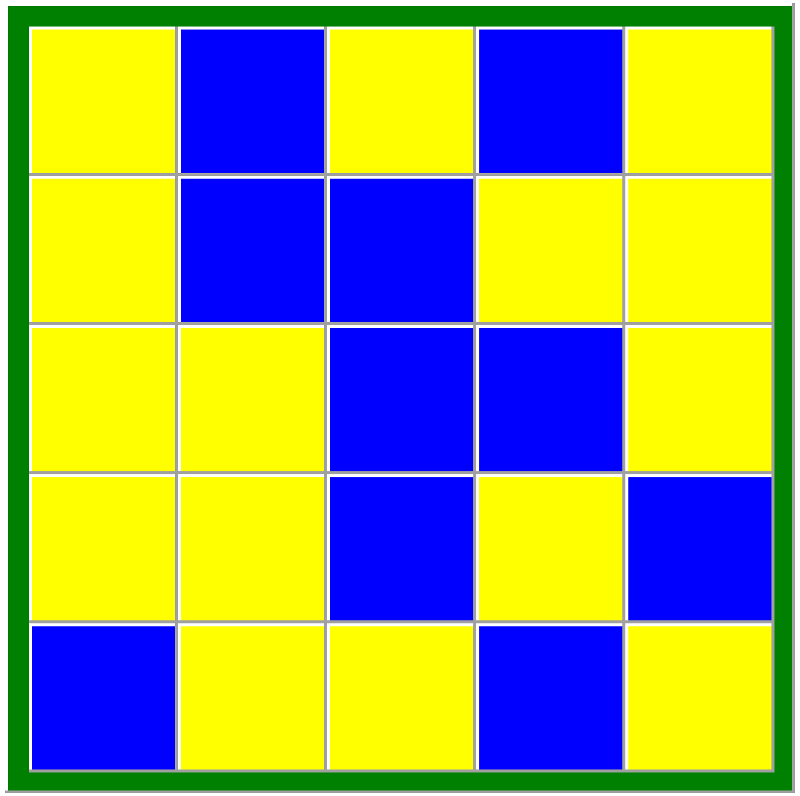
Rivals in the dark: How competition influences search in decisions under uncertainty

Nathaniel D. Phillips[a,*], Ralph Hertwig[a], Yaakov Kareev[b], Judith Avrahami[b]

[a] Max Planck Institute for Human Development, Berlin, Germany
[b] The Center for the Study of Rationality, The Hebrew University of Jerusalem, Jerusalem, Israel

Affordances and Inferences

- Proof of concept that rewarding novelty can incentivize people to acquire less information on research problems.
- Simple paradigm that can be used to test Metascientific hypotheses.



“Interventions to change the current system should not be accepted without proper scrutiny, even when they are reasonable and well intended.”

OPEN ACCESS Freely available online

PLOS | MEDICINE

Essay

How to Make More Published Research True

John P. A. Ioannidis[1,2,3,4*]

Thank You.

Contact

LeoTiokhin.com

LeoTiokhin@gmail.com

@LeonidTiokhin

Data, materials, code, and pre-registrations

All: https://osf.io/u97k3/

Code: https://github.com/ltiokhin

Registered Report: https://osf.io/tn2vu/

Thank You.

TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts

ASU School of Human Evolution & Social Change

Arizona State University

leibniz-psychology.org

Open Science 2019

March 12-14, 2019, Trier, Germany

Contact

LeoTiokhin.com

LeoTiokhin@gmail.com

@LeonidTiokhin

Data, materials, code, and pre-registrations

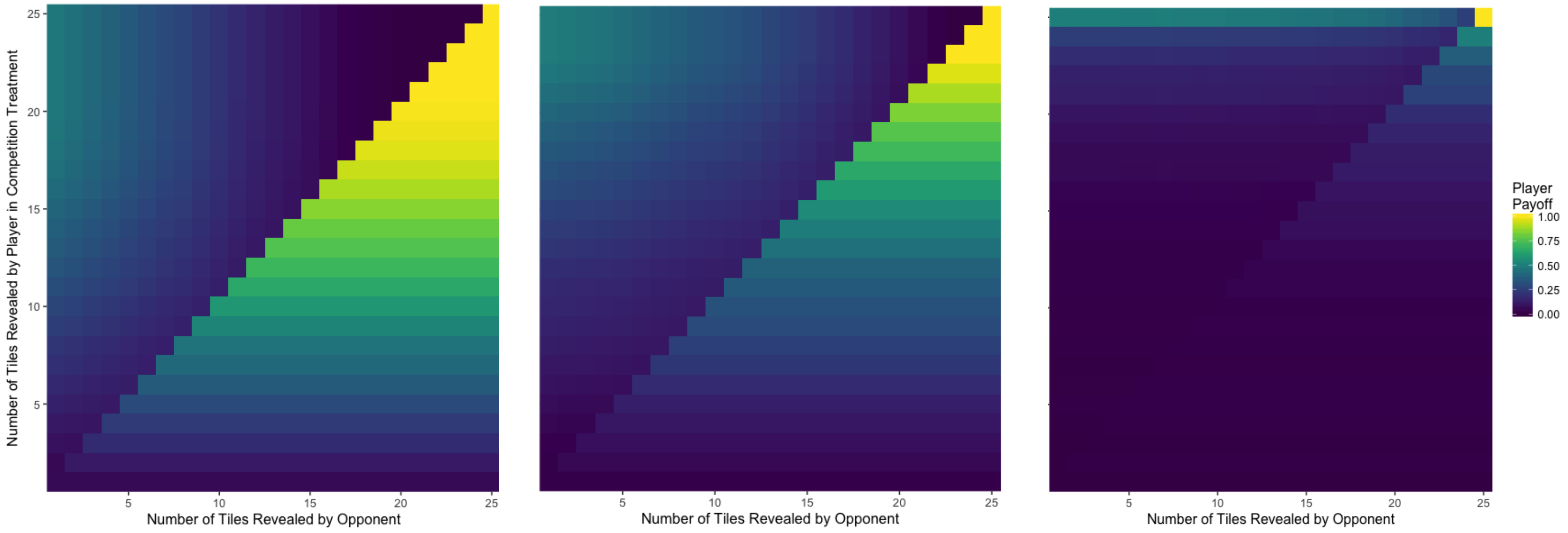
All: https://osf.io/u97k3/

Code: https://github.com/ltiokhin

Registered Report: https://osf.io/tn2vu/

	No Comp	Comp
No Effort	Time: 20 minutes Points: +1, -1 Tiles: 1 second	Time: As long as necessary Points: +1, -1 (if not scooped) Tiles: 1 second
Effort	Time: 20 minutes Points: +1, -1 Tiles: depends on effort	Time: As long as necessary Points: +1, -1 (if not scooped) Tiles: depends on effort

Expected Tiles Revealed: Competition



BIC Approximation: Bayes Factors (Wagenmakers, 2007)

tion: $\Pr_{\mathrm{BIC}}(D \mid H_i) = \exp[-\mathrm{BIC}(H_i)/2]$. In the case of two models, H_0 and H_1, the Bayes factor is defined as the ratio of the prior predictive probabilities; hence, the BIC approximation of the Bayes factor is given by

$$BF_{01} \approx \frac{\Pr_{\mathrm{BIC}}(D \mid H_0)}{\Pr_{\mathrm{BIC}}(D \mid H_1)} = \exp\left(\Delta\mathrm{BIC}_{10}/2\right), \qquad (10)$$

where $\Delta\mathrm{BIC}_{10} = \mathrm{BIC}(H_1) - \mathrm{BIC}(H_0)$. For instance, if

Models used for Bayes Factors

Time to accurately solve one arithmetic problem

Model 0: lmer(ElapsedTime_MathSolved ~ n_major.s + (1|ID_Player), data = d.math.agg.f, REML=FALSE)
Model 1: lmer(ElapsedTime_MathSolved ~ Competition + n_major.s + (1|ID_Player), data = d.math.agg.f, REML=FALSE)
Max N = 20098 observations. Min N = 130 observations.
BIC Max N. $BF_{01} = 142$
BIC Min N. $BF_{01} = 11$

Model 0: lmer(ElapsedTime_MathSolved ~ Sex + n_major.s + Guess_Number.s + (1|ID_Player), data = d.math.agg.f, REML=FALSE)
Model 1: lmer(ElapsedTime_MathSolved ~ Sex + Competition + n_major.s + Guess_Number.s + (1|ID_Player), data = d.math.agg.f, REML=FALSE)
Max N = 20098 observations. Min N = 130 observations.
BIC Max N. $BF_{01} = 142$
BIC Min N. $BF_{01} = 11$

Models used for Bayes Factors

Tiles revealed

Model 0: lmer(TilesRevealed ~ Competition + n_major.s + (1|ID_Player), data = d.math.agg.f, REML=FALSE)

*Model 1: lmer(TilesRevealed ~ Competition*Effort + n_major.s + (1|ID_Player), data = d.conf.agg.f, REML=FALSE)*

Max N = 14073 observations. Min N = 260 observations.

BIC Max N. $BF_{01} = 44$

BIC Min N. $BF_{01} = 6$

*Model 0: lmer(TilesRevealed ~ Sex*Effort + Competition + Effort + n_major.s +Guess_Number.s + (1|ID_Player), data = d.conf.agg.f, REML=FALSE)*

*Model 1: lmer(TilesRevealed ~ Sex*Effort + Competition*Effort + n_major.s +Guess_Number.s + (1|ID_Player), data = d.conf.agg.f, REML=FALSE)*

Max N = 14073 observations. Min N = 260 observations.

BIC Max N. $BF_{01} = 118$

BIC Min N. $BF_{01} = 16$

Models used for Bayes Factors

Accuracy

Model 0: glmer(Correct_Guess ~ Competition + Effort + n_major.s + (1|ID_Player), data = d.conf.agg.f, family = binomial)

*Model 1: glmer(Correct_Guess ~ Competition*Effort + n_major.s + (1|ID_Player), data = d.conf.agg.f, family = binomial)*

Max N = 14073 observations. Min N = 260 observations.

BIC Max N. $BF_{01} = 118$

BIC Min N. $BF_{01} = 16$

*Model 0: glmer(Correct_Guess ~ Sex*Effort + Competition + Effort + n_major.s + Guess_Number.s + (1|ID_Player), data = d.conf.agg.f, family = binomial)*

*Model 1: glmer(Correct_Guess ~ Sex*Effort + Competition*Effort + n_major.s + Guess_Number.s + (1|ID_Player), data = d.conf.agg.f, family = binomial)*

Max N = 14073 observations. Min N = 260 observations.

BIC Max N. $BF_{01} = 118$

BIC Min N. $BF_{01} = 16$

Bayesian Models used in RR

Model 1: To compare the number of tiles that players reveal before guessing the majority color in the Competition treatments versus the No-Competition treatments, we will use a multiple linear regression, with random effects for player, of the following form:

$Yi \sim \text{Normal}(\mu_i, \sigma)$

$\mu_i = \alpha + \alpha_{\text{PLAYER}[i]} + \beta_C * C_i + \beta_E * E_i + \beta_{CE} * C_i E_i + \beta_{Ns} * Ns_i$

Y_i: Number of tiles clicked before guessing. α: Intercept. $\alpha_{\text{PLAYER}[i]}$: Random intercept for each player. C: Competition Treatment (1 / 0). E: Effort Condition (1 / 0). $\beta_{CE} * C_i E_i$: Interaction between treatment and effort. β_{Ns}: Standardized number of tiles for the majority color (i.e. effect size).

Bayesian Models used in RR

Model 2: To assess the probability of a correct guess, we will use a logistic regression, with random effects for player, of the following form:

$S_i \sim \text{Binomial}(1, p_i)$

$\text{Logit}(p_i) = \alpha + \alpha_{\text{PLAYER}[i]} + \beta_C * C_i + \beta_E * E_i + \beta_{CE} * C_i E_i + \beta_{Ns} * Ns_i$

S_i: Successful guess. α: Intercept. $\alpha_{\text{PLAYER}[i]}$: Random intercept for each player. C: Competition Treatment (1 / 0). E: Effort Condition (1 / 0). $\beta_{CE} * C_i E_i$: Interaction between treatment and effort. β_{Ns}: Standardized number of tiles for the majority color (i.e. effect size).

Bayesian Models used in RR

Model 3: To test the effect of competition on the time between clicking one tile and being allowed to click the subsequent tile (i.e. time to accurately solve one arithmetic problem), we will use a multiple linear regression, with random effects for player, of the following form:

$Yi \sim \text{Normal}(\mu_i, \sigma)$

$\mu_i = \alpha + \alpha_{PLAYER[i]} + \beta_C * C_i + \beta_{Ns} * Ns_i$

Y_i: Time between clicking one tile and being allowed to click the subsequent tile (i.e. time to solve an arithmetic problem; seconds). α: Intercept. $\alpha_{PLAYER[i]}$: Random intercept for each player. C: Competition Treatment (1 / 0). β_{Ns}: Standardized number of tiles for the majority color (i.e. effect size).

Priors for Bayesian Models used in RR

Priors.

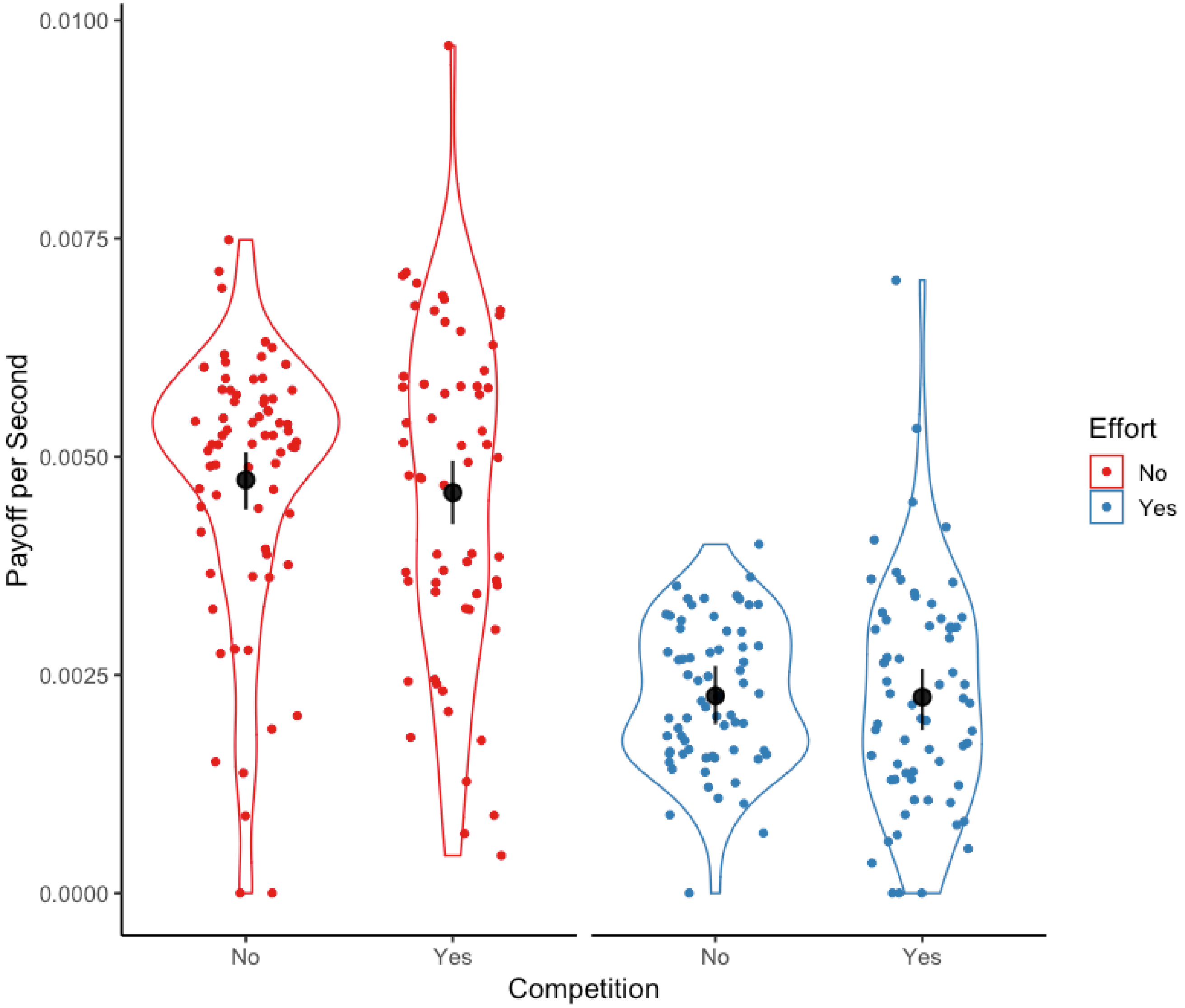
Parameter	Effort Manipulation Check	Competition Attention Check	Model 1 (Tiles)	Model 2 (Correct Guess)	Model 3 (Arithmetic Time)
σ	Gamma (2, 0.5)	NA	Gamma (2, 0.5)	NA	Gamma (2, 0.5)
α	Gamma (1.5, 0.05)	Normal (0, 10)	Uniform (0, 25)	Normal (0, 10)	Gamma (1, 0.05)
α_{PLAYER}	Normal (0, σ_{PLAYER})	NA	Normal (0, σ_{PLAYER})	Normal (0, σ_{PLAYER})	Normal (0, σ_{PLAYER})
σ_{PLAYER}	Gamma (1.5, 0.05)	NA	Gamma (1.5, 0.05)	Gamma (1.5, 0.05)	Gamma (1, 0.05)
B_C	Normal (0, 10)	Normal (0, 10)	Normal (0, 10)	Normal (0, 10)	Normal (0, 10)
β_E	Normal (0, 10)	NA	Normal (0, 10)	Normal (0, 10)	NA
β_{CE}	Normal (0, 10)	NA	Normal (0, 10)	Normal (0, 10)	NA
β_{Ns}	NA	NA	Normal (0, 10)	Normal (0, 10)	Normal (0, 10)

ROPES for Bayesian Models used in RR

Parameter	Effort Manipulation Check	Competition Attention Check	Model 1 (Tiles)	Model 2 (Correct Guess)	Model 3 (Arithmetic Time)
B_C	NA	(-0.8, 0.8)	(-1.22, 1.22)	(-0.19, 0.19)	(-0.33, 0.33)**
β_E	(-0.5, 0.5)	NA	NA	NA	NA
β_{CE}	NA	NA	(-0.10, 0.10)*	(-0.09, 0.09)	NA

Table 2 | Region of practical equivalence (ROPE) for quality checks and confirmatory analyses. ROPEs for quality checks are based on subjective assessment of what effect size would convincingly indicate a successful manipulation. ROPEs for confirmatory analyses (Models 1 - 3) are based on 95% statistical power, unless indicated otherwise. Model 1 tests the effect of the Competition treatment and Effort condition on number of tiles clicked before guessing, using a multiple linear regression with random effects for each player. Model 2 tests the effect of the Competition treatment and Effort condition on the probability of a correct guess, using a logistic regression with random effects for each player. Model 3 tests the effect of the Competition treatment on the time to accurately solve one arithmetic problem, using a multiple linear regression with random effects for each player. *ROPE based on 85% statistical power. **ROPE based on 99% statistical power.

Reward



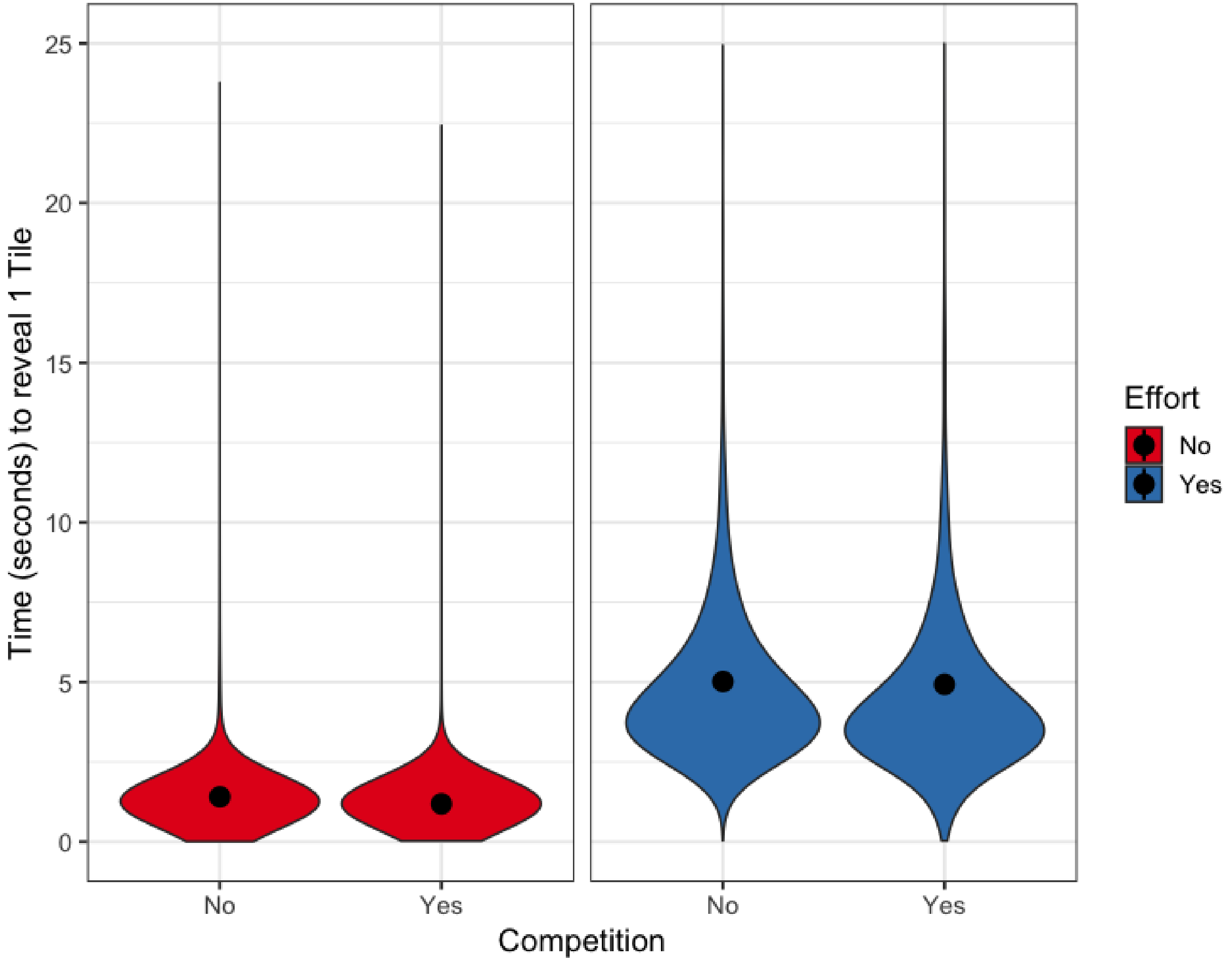
0.0100
0.0075
0.0050
0.0025
0.0000
Payoff per Second
No
Yes
No
Yes
Competition
Effort
No
Yes

WAIC: Reward Model Comparison

REWARD PER SECOND

	WAIC	pWAIC	dWAIC	weight
Reward_E	-2670.59	3.51	0.00	0.72
Reward_E_C	-2667.98	4.93	2.61	0.19
Reward_E_C_EC	-2666.40	5.76	4.19	0.09

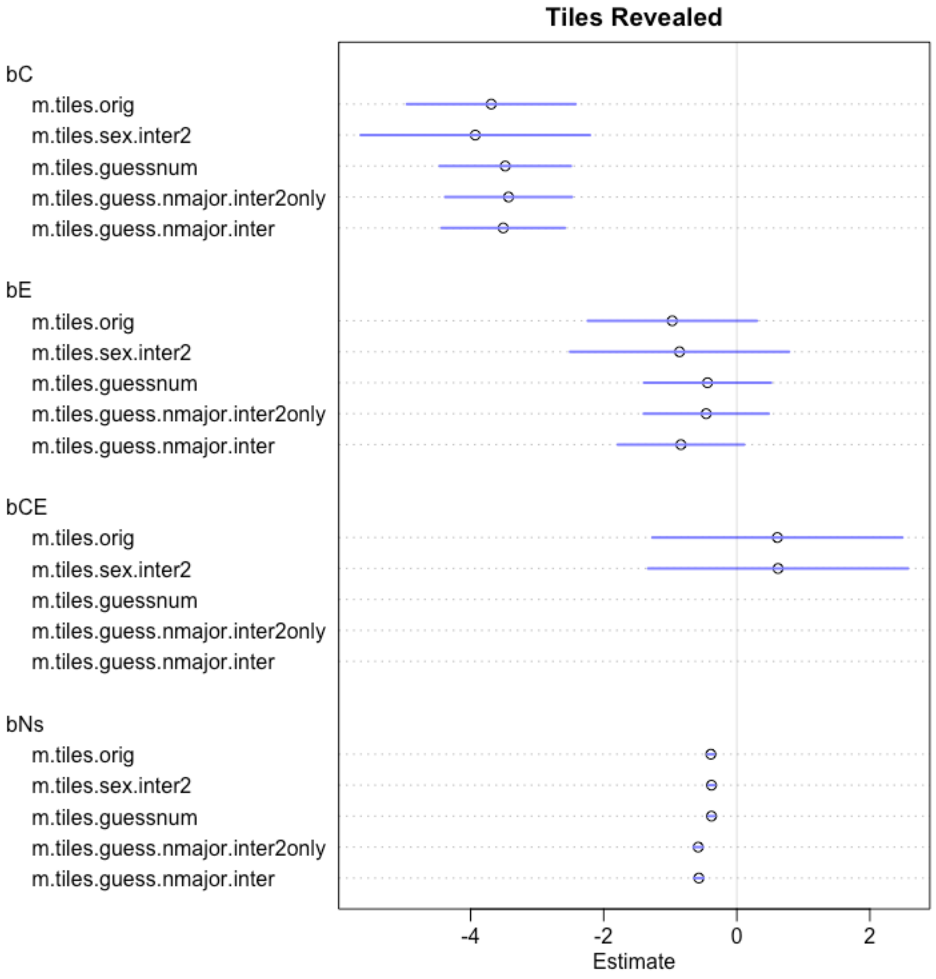
Time to reveal 1 tile

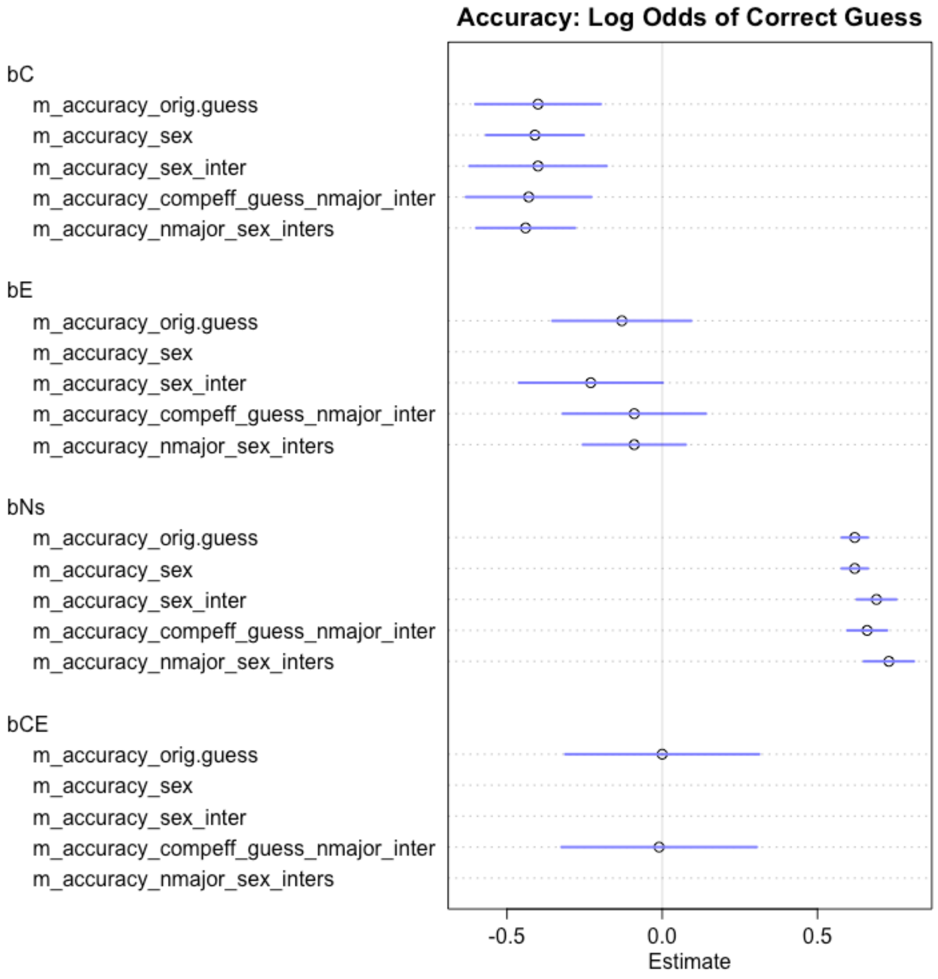


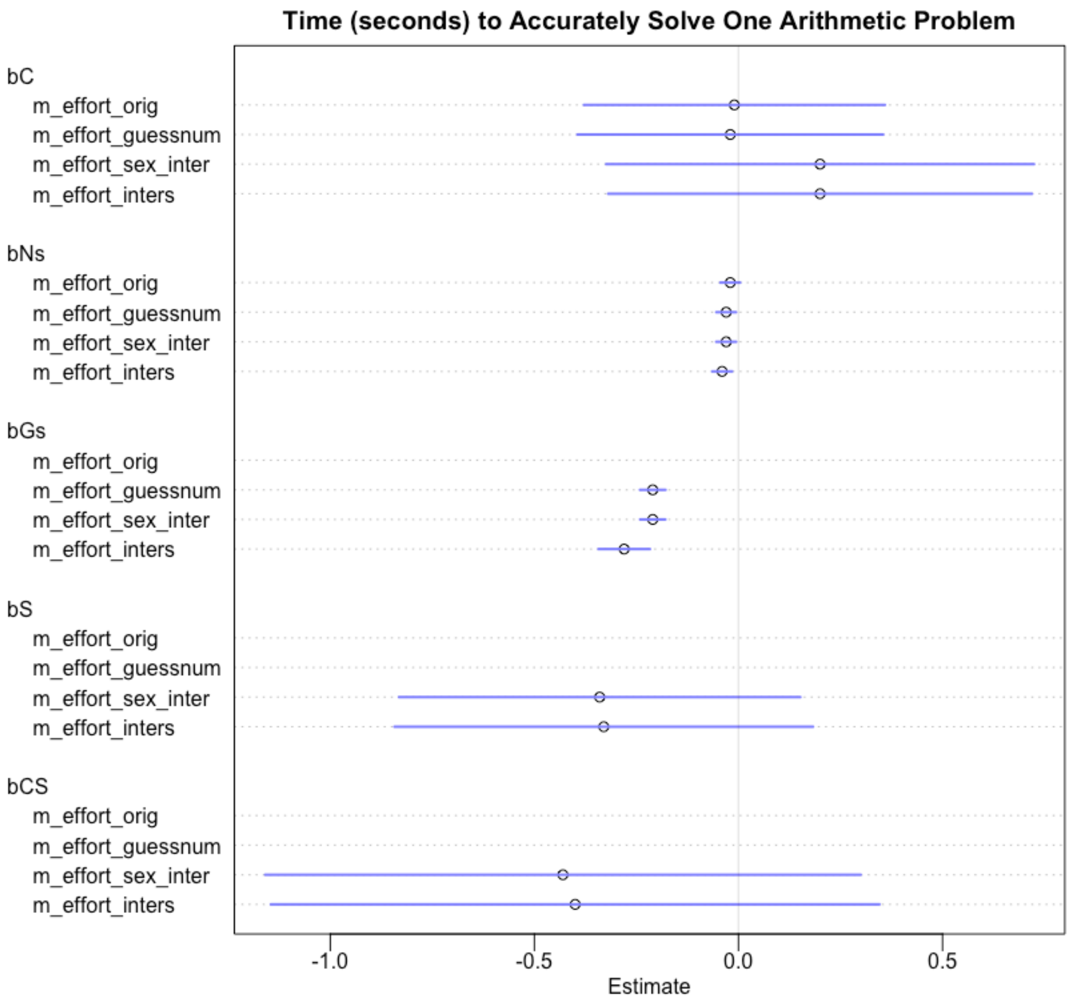
WAIC: Time 1 Tile Model Comparison

TIME TO REVEAL 1 TILE

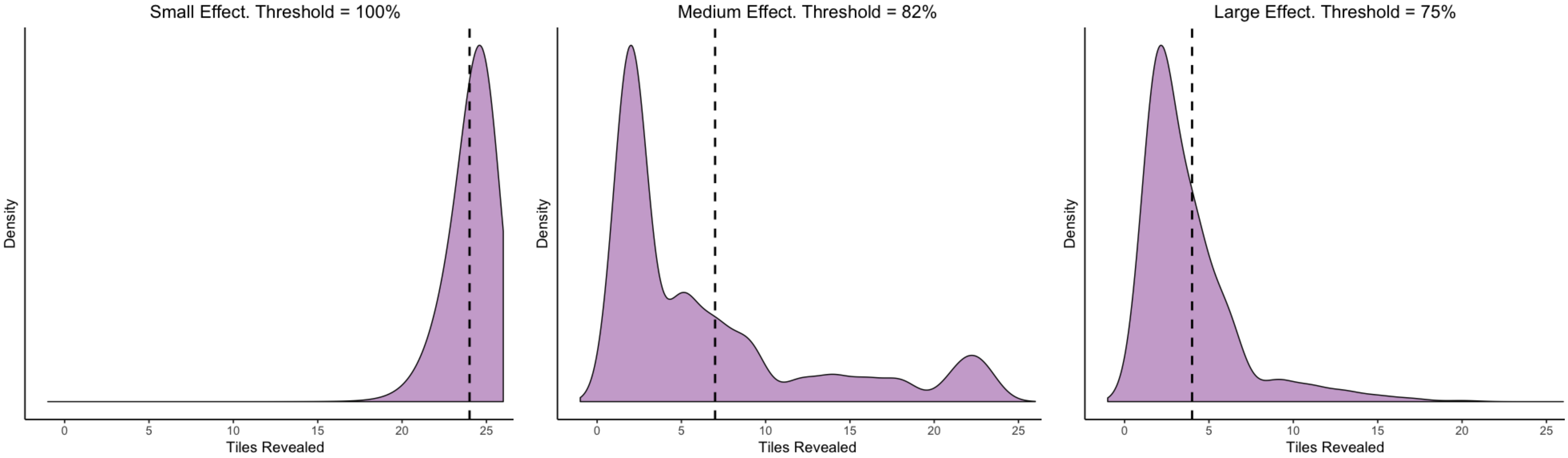
	WAIC	pWAIC	dWAIC	weight
Time_Tile_E	192698.0	475.96	0.00	0.40
Time_Tile_E_C_EC	192698.1	476.12	0.10	0.38
Time_Tile_E_C	192699.2	476.88	1.21	0.22

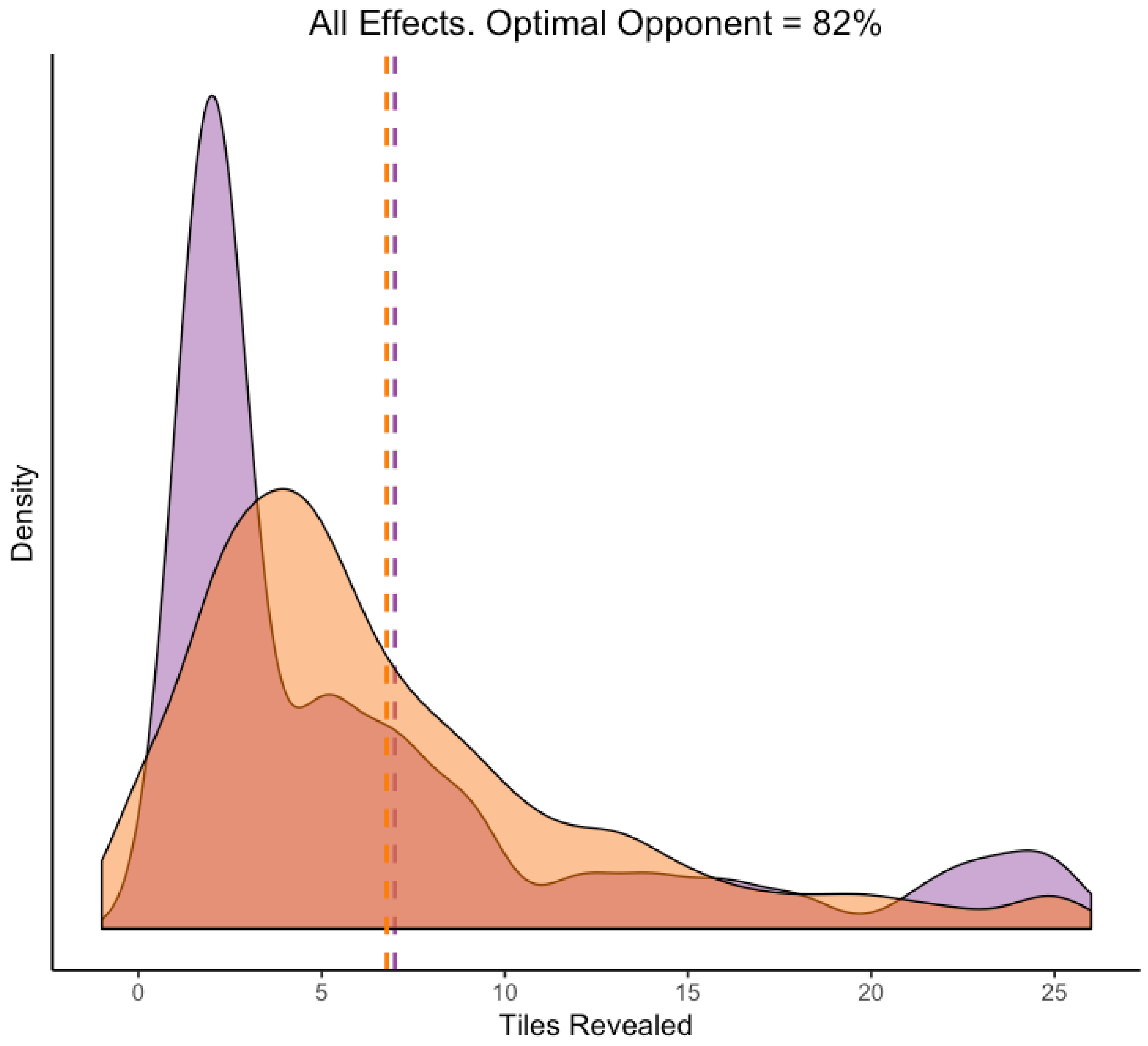

Tiles Revealed
bC
m.tiles.orig
m.tiles.sex.inter2
m.tiles.guessnum
m.tiles.guess.nmajor.inter2only
m.tiles.guess.nmajor.inter
bE
m.tiles.orig
m.tiles.sex.inter2
m.tiles.guessnum
m.tiles.guess.nmajor.inter2only
m.tiles.guess.nmajor.inter
bCE
m.tiles.orig
m.tiles.sex.inter2
m.tiles.guessnum
m.tiles.guess.nmajor.inter2only
m.tiles.guess.nmajor.inter
bNs
m.tiles.orig
m.tiles.sex.inter2
m.tiles.guessnum
m.tiles.guess.nmajor.inter2only
m.tiles.guess.nmajor.inter
-4
-2
0
2
Estimate


Accuracy: Log Odds of Correct Guess
bC
m_accuracy_orig.guess
m_accuracy_sex
m_accuracy_sex_inter
m_accuracy_compeff_guess_nmajor_inter
m_accuracy_nmajor_sex_inters
bE
m_accuracy_orig.guess
m_accuracy_sex
m_accuracy_sex_inter
m_accuracy_compeff_guess_nmajor_inter
m_accuracy_nmajor_sex_inters
bNs
m_accuracy_orig.guess
m_accuracy_sex
m_accuracy_sex_inter
m_accuracy_compeff_guess_nmajor_inter
m_accuracy_nmajor_sex_inters
bCE
m_accuracy_orig.guess
m_accuracy_sex
m_accuracy_sex_inter
m_accuracy_compeff_guess_nmajor_inter
m_accuracy_nmajor_sex_inters
-0.5
0.0
0.5
Estimate


Time (seconds) to Accurately Solve One Arithmetic Problem
bC
m_effort_orig
m_effort_guessnum
m_effort_sex_inter
m_effort_inters
bNs
m_effort_orig
m_effort_guessnum
m_effort_sex_inter
m_effort_inters
bGs
m_effort_orig
m_effort_guessnum
m_effort_sex_inter
m_effort_inters
bS
m_effort_orig
m_effort_guessnum
m_effort_sex_inter
m_effort_inters
bCS
m_effort_orig
m_effort_guessnum
m_effort_sex_inter
m_effort_inters
-1.0
-0.5
0.0
0.5
Estimate

Expected Tiles Revealed: No Competition





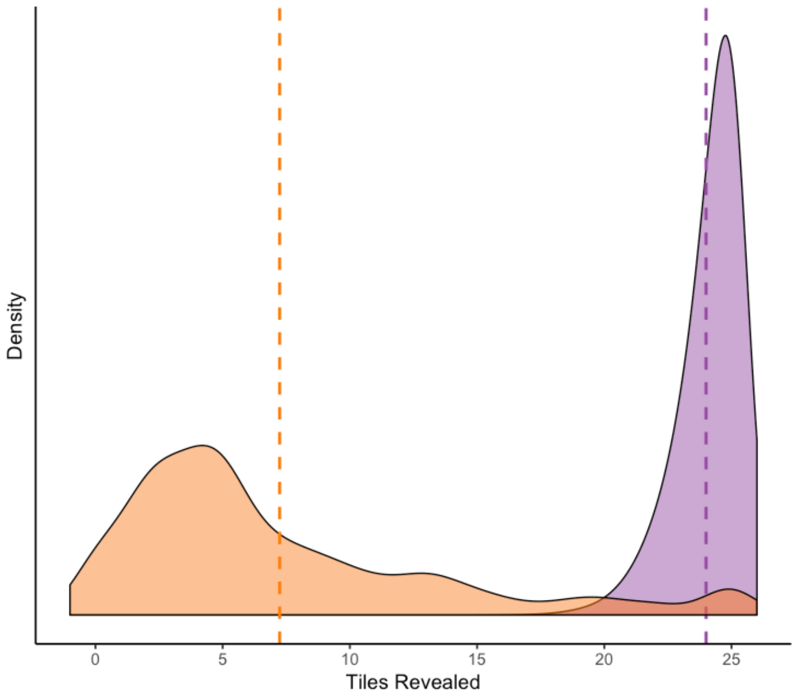
Experiment

Simulation

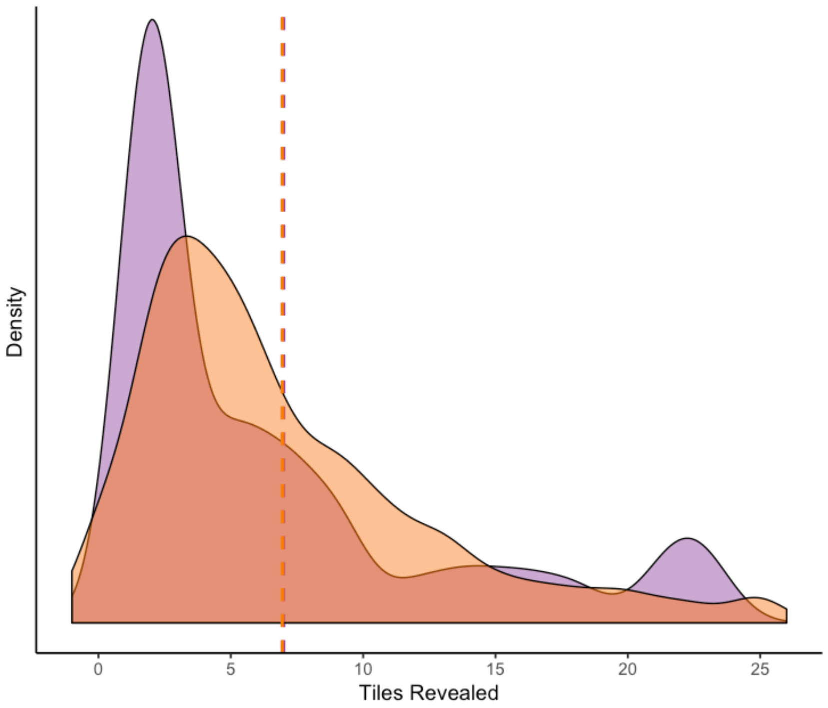
Experiment

Simulation

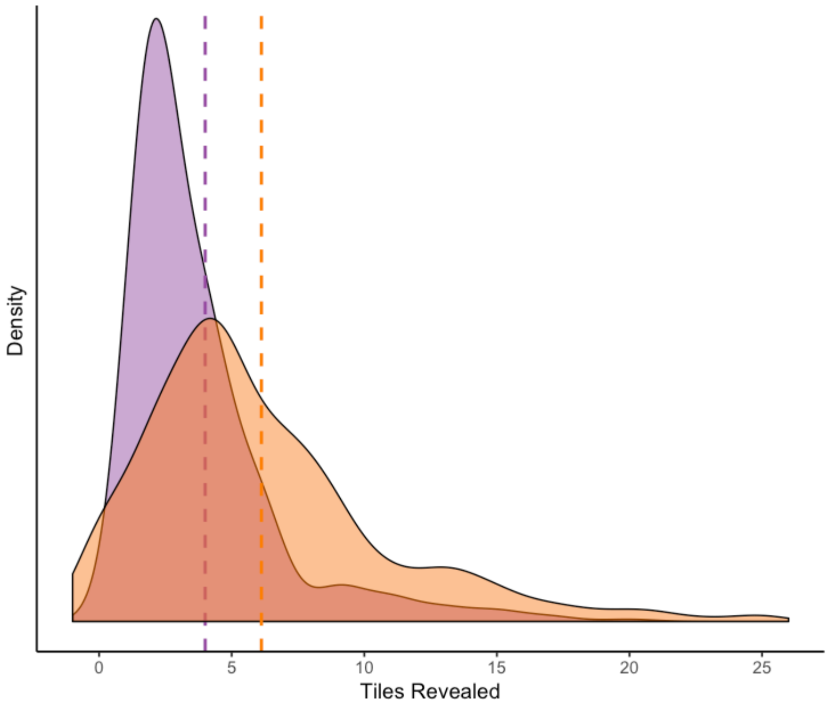
Optimal: 100%



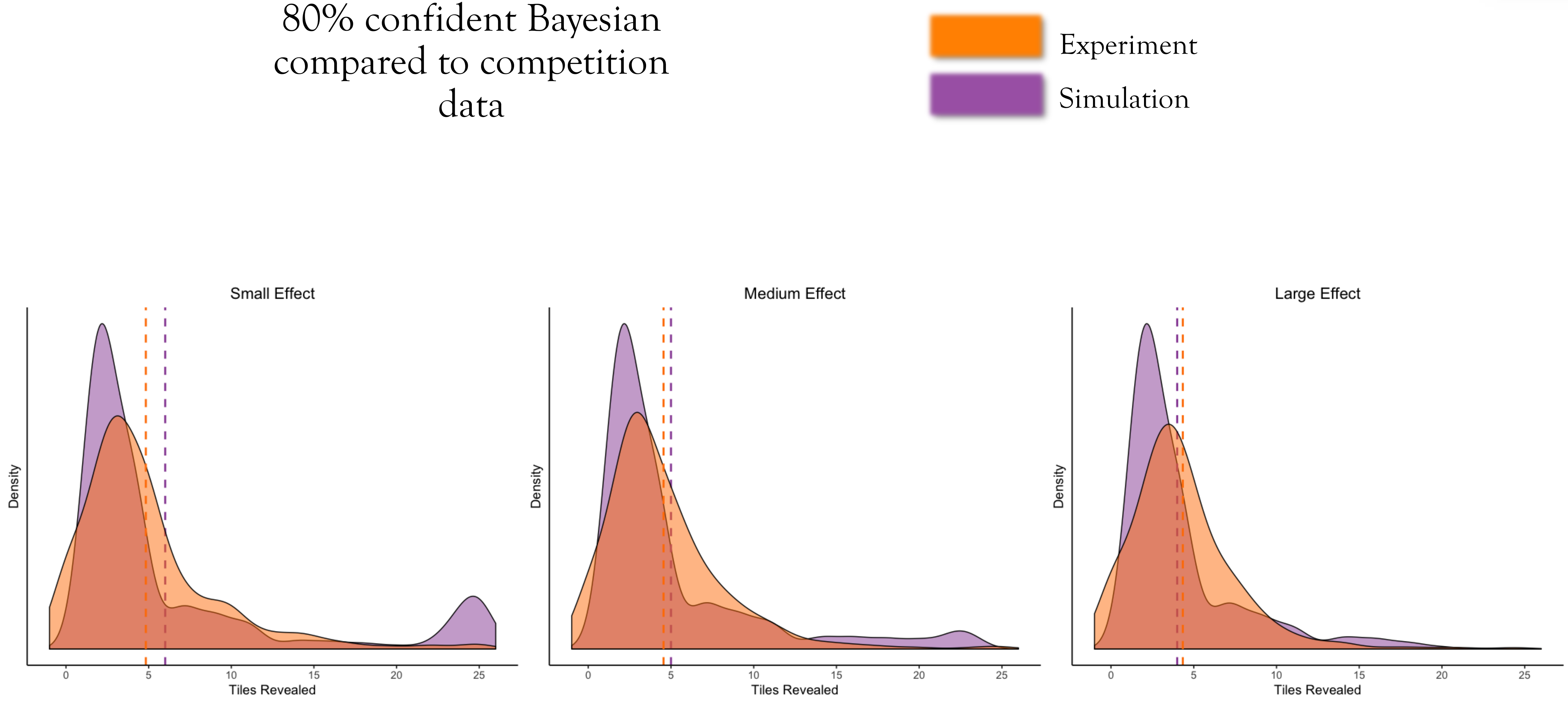
Optimal: 82%



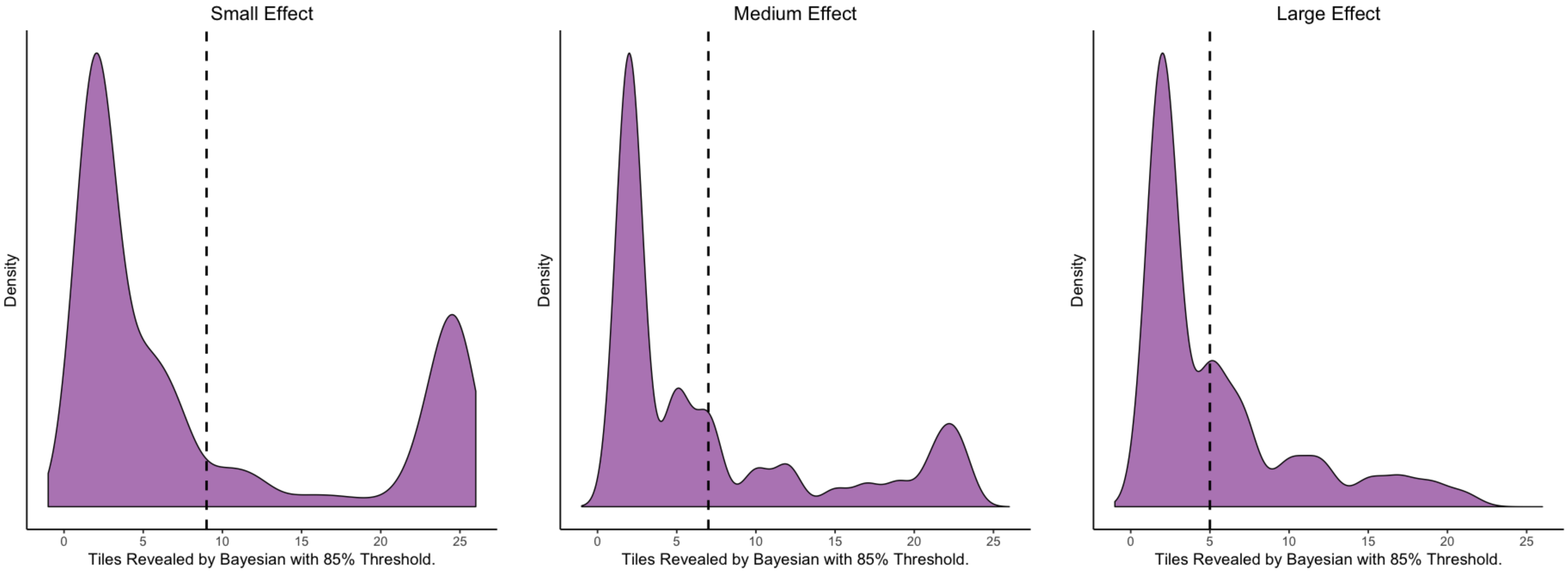
Optimal: 75%

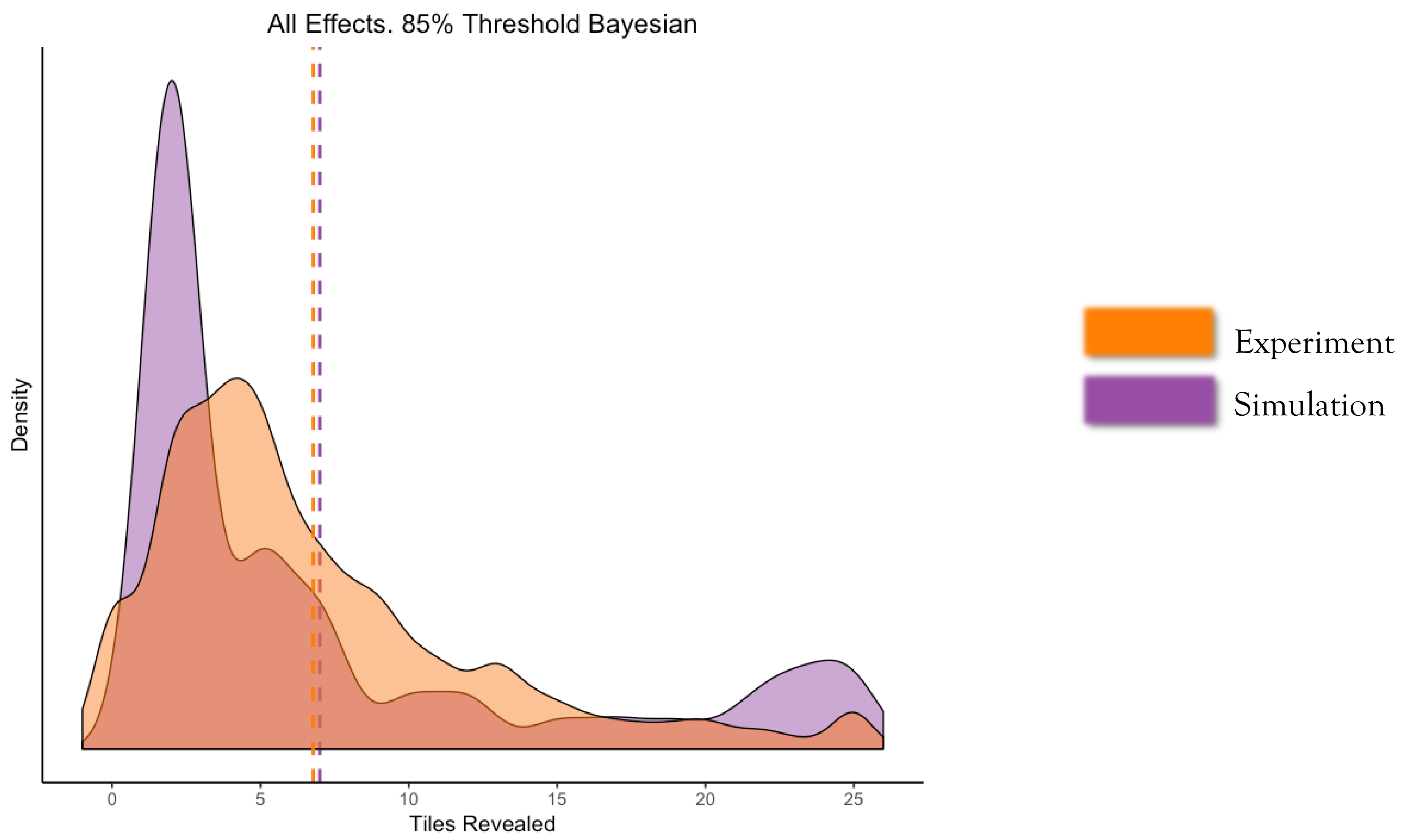


80% confident Bayesian compared to competition data



Tiles revealed by 85% Bayesian across effect sizes

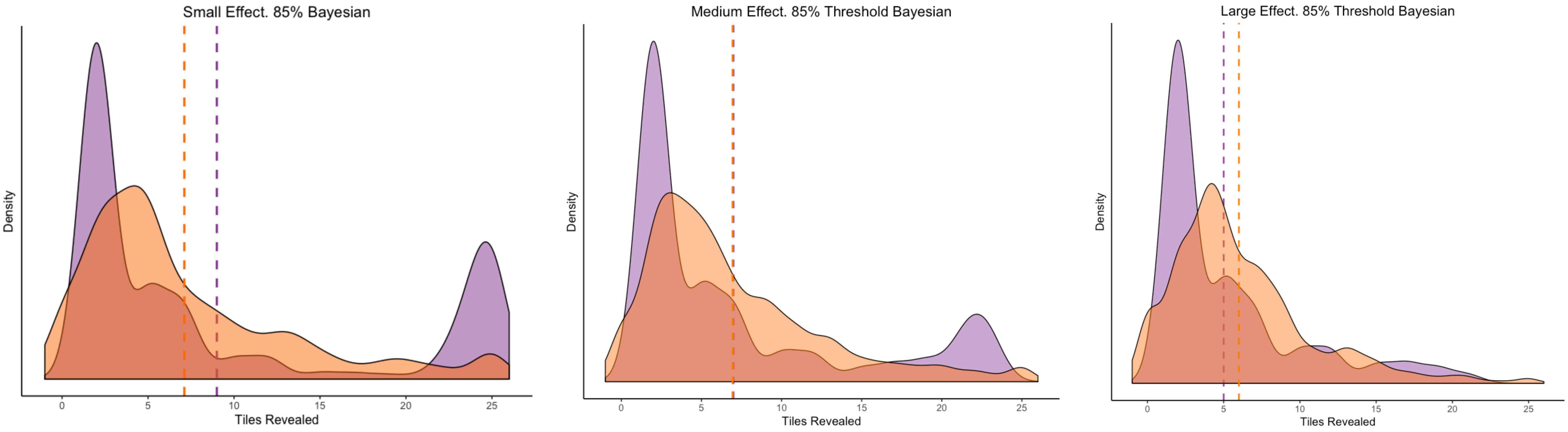




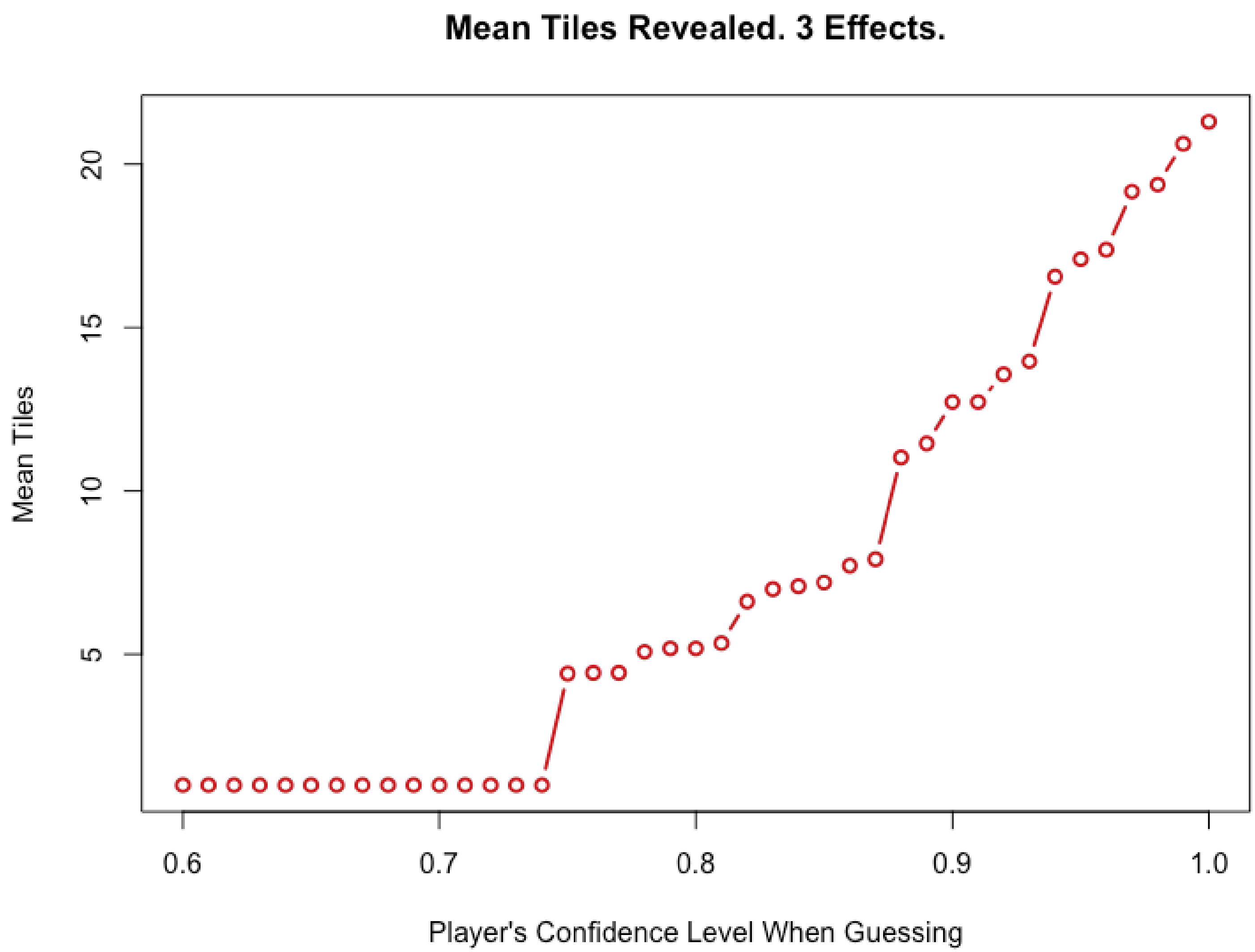
All Effects. 85% Threshold Bayesian
Density
0
5
10
15
20
25
Tiles Revealed
Experiment
Simulation

85% Bayesian compared to experimental data

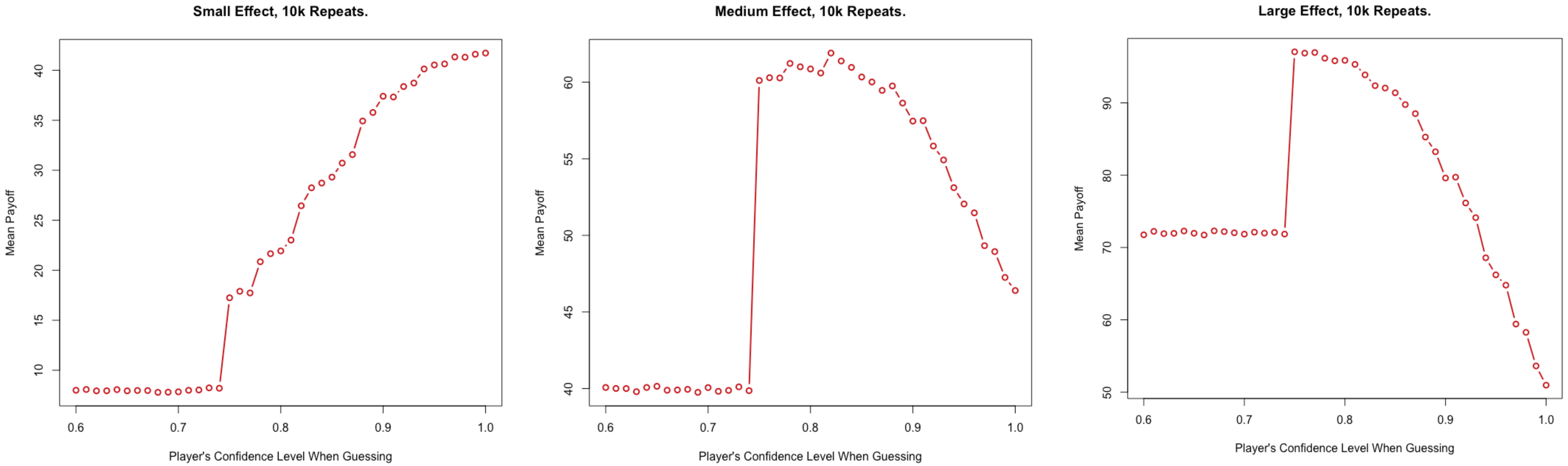
Experiment
Simulation



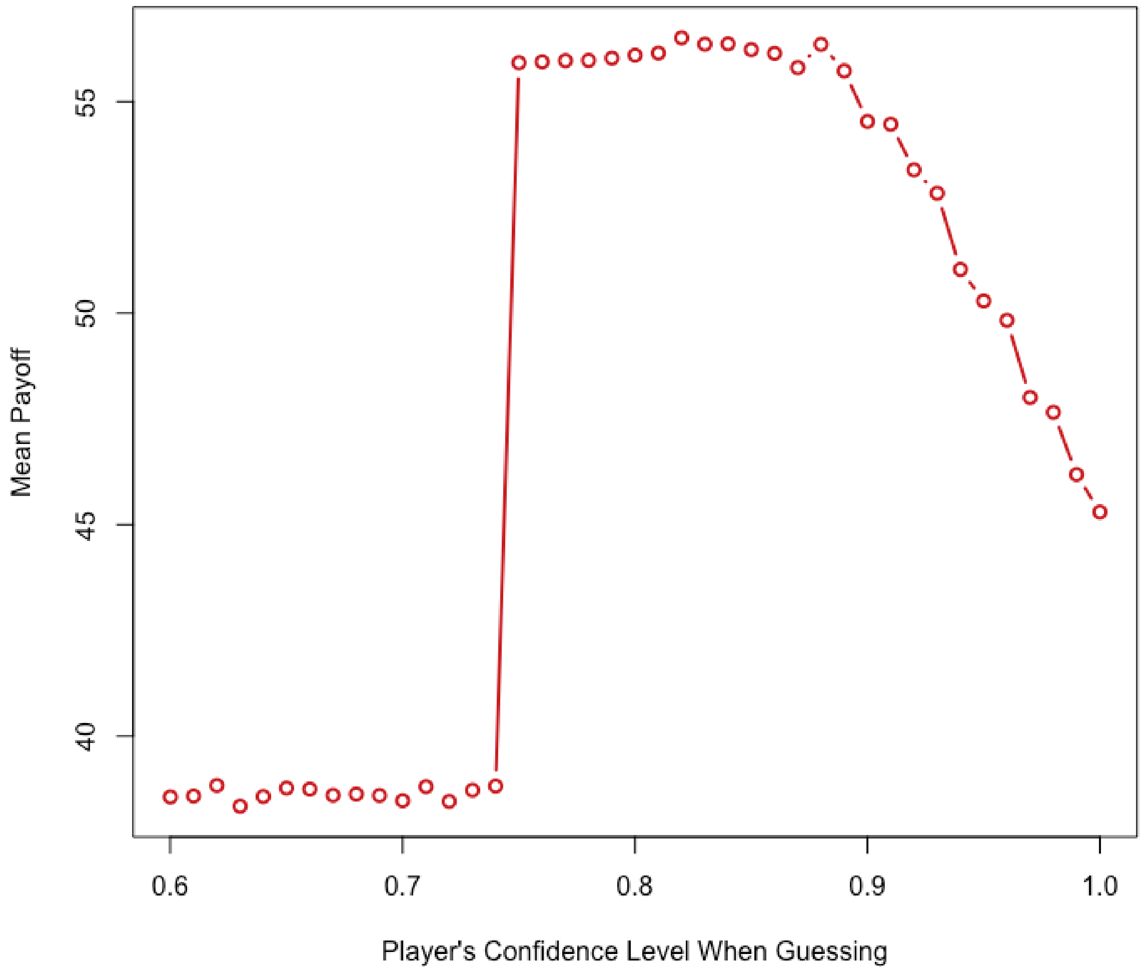
Expected Tiles Revealed



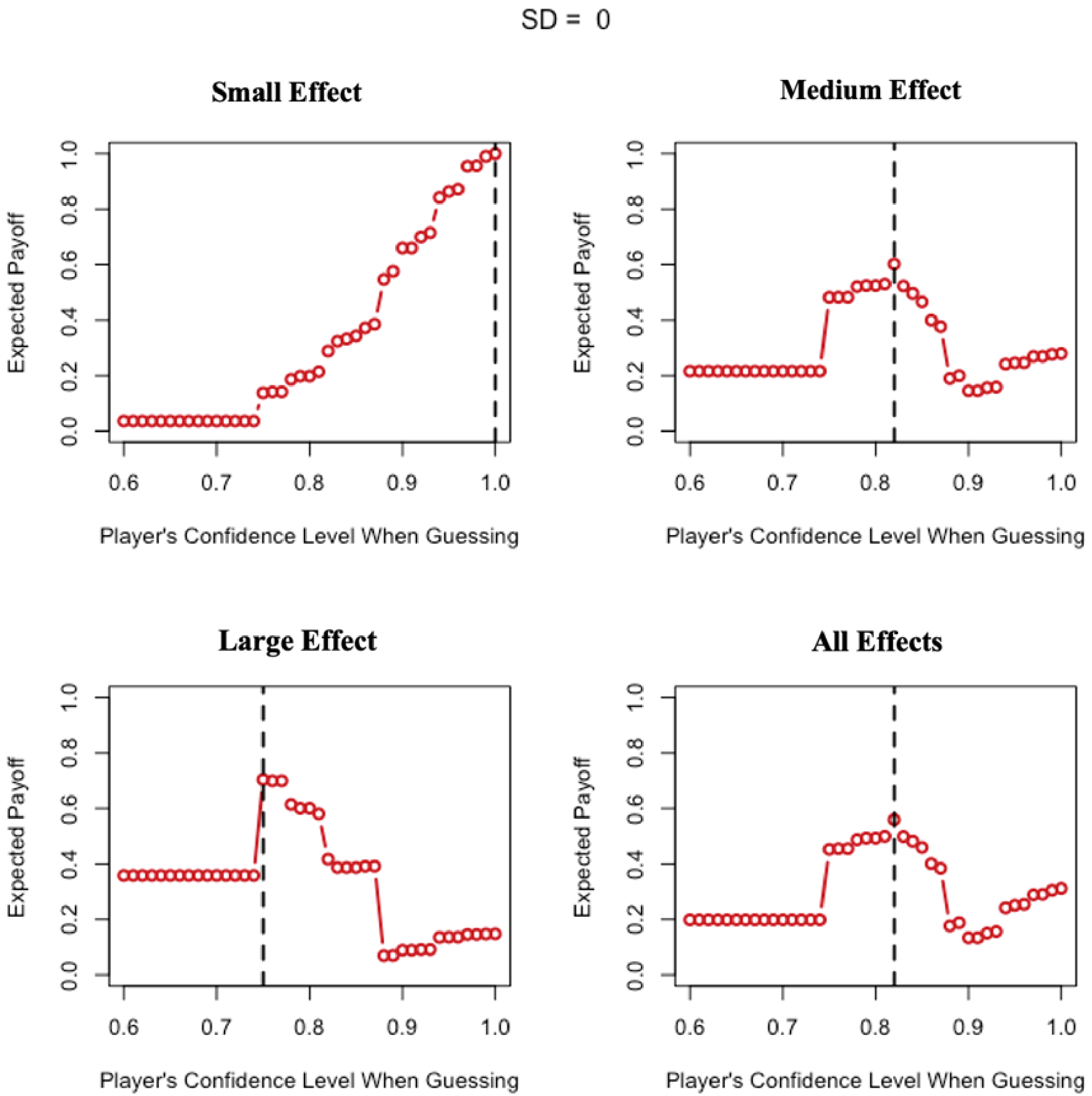
Payoff to Simulated Bayesian



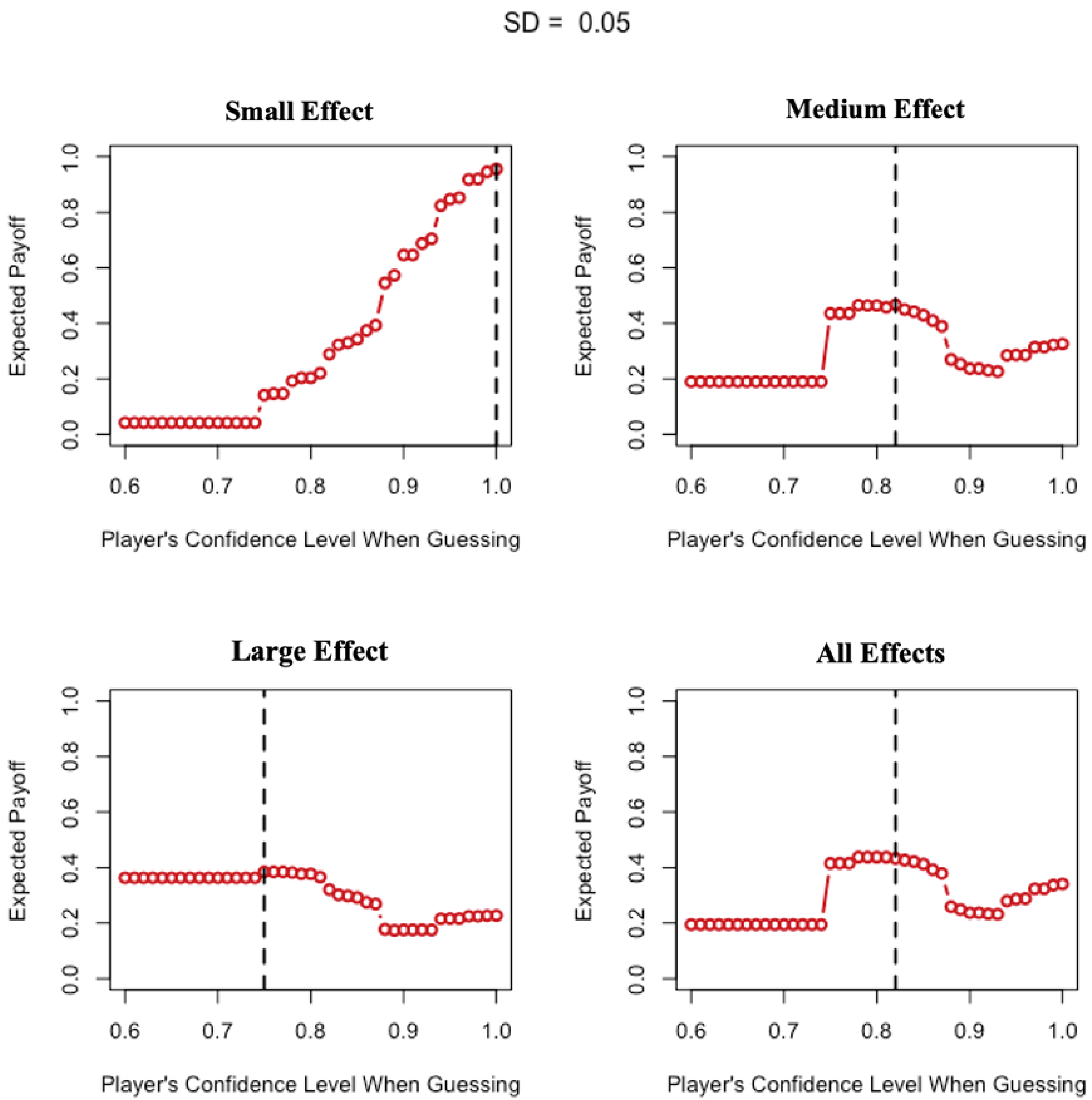
Payoff to Simulated Bayesian (All Effects)



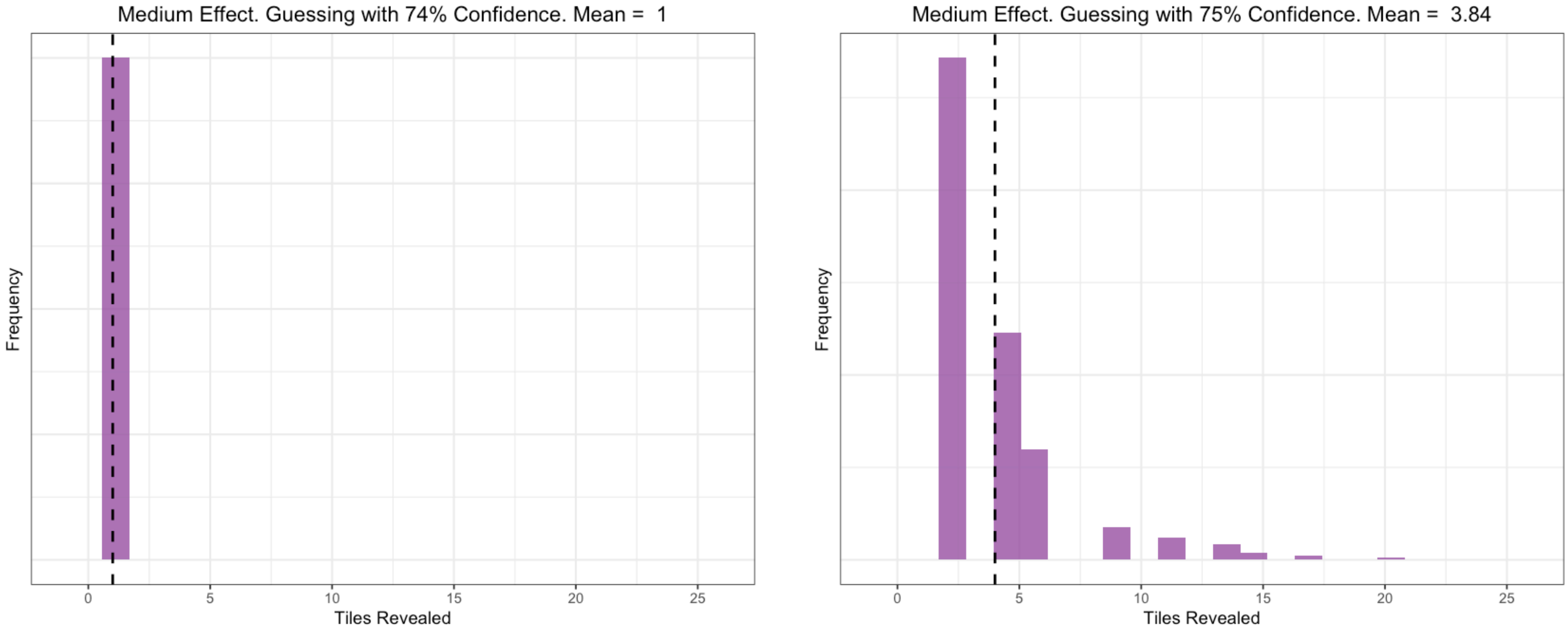
Payoff to Simulated Bayesian in Competition



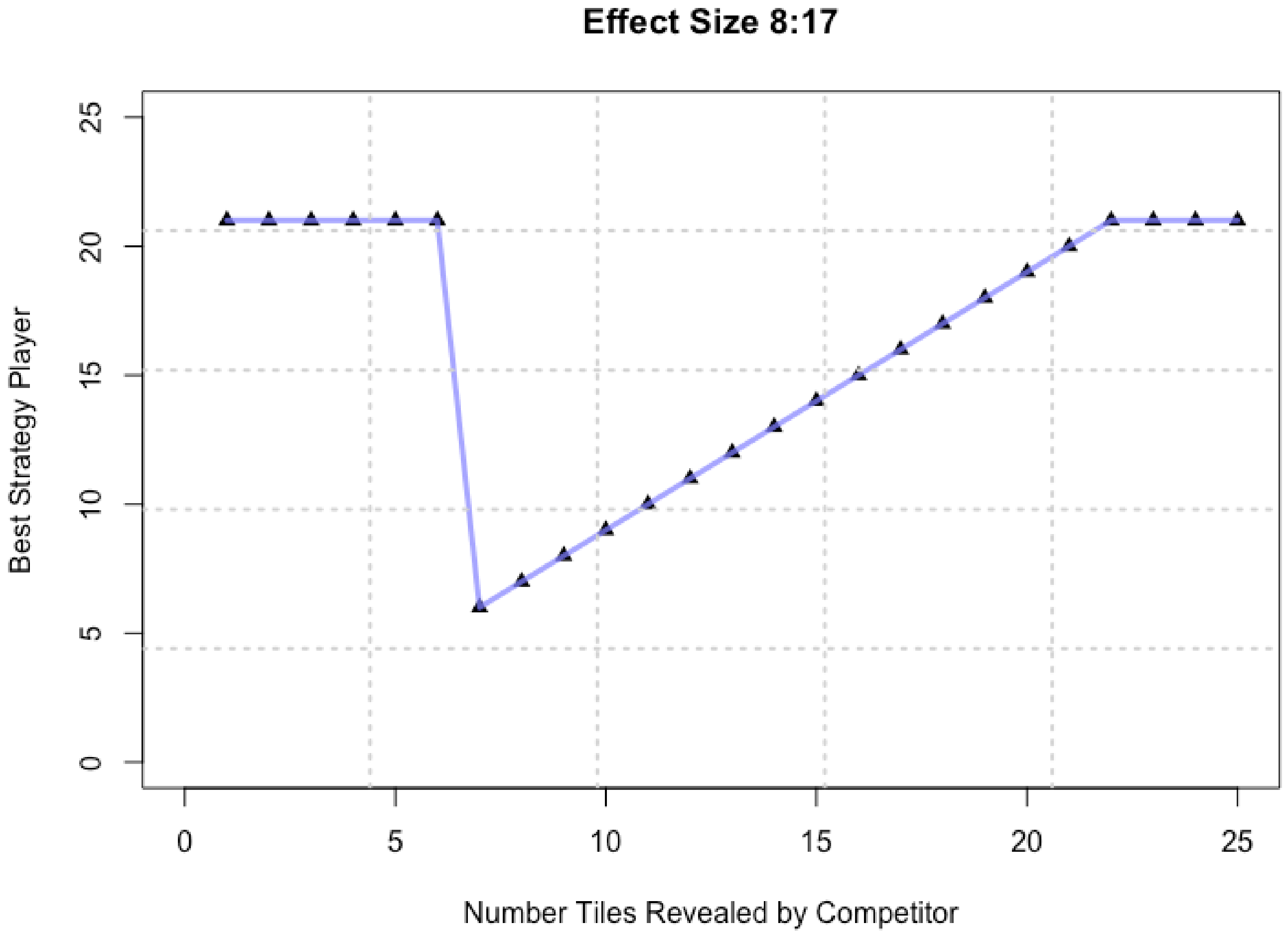
Payoff to Simulated Bayesian in Competition



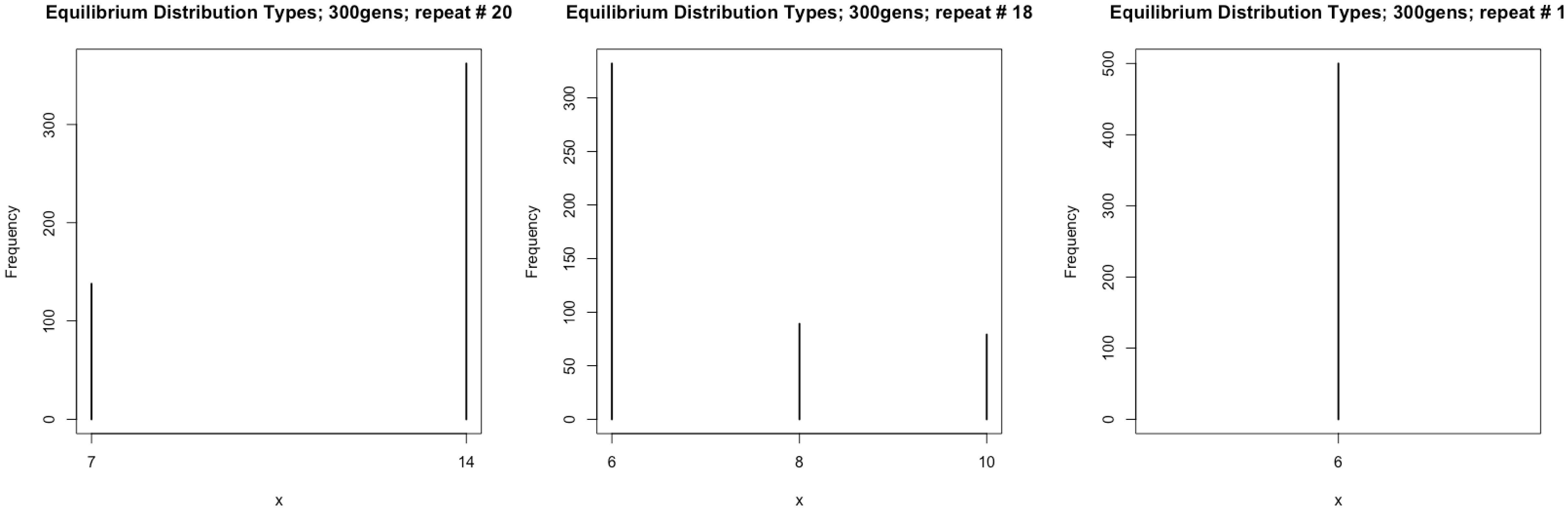
Why payoff jumps around 75%



No ESS w/ Fixed Strategies



No ESS w/ Fixed Strategies



Mixed Strategy ESS