The Streetlight Effect in Data-Driven Exploration

Online Appendix: Additional Results

Johannes Hoelzemann University of Vienna Gustavo Manso UC Berkeley

Abhishek Nagaraj UC Berkeley & NBER Matteo Tranchero The Wharton School

May 9, 2025

A	Exp	erimental Results: Additional Details	2
	A.1	Logistics of the Experiment	2
	A.2	Additional Figures and Tables	4
B	The	Genetic Roots of Human Diseases: Additional Details	9
	B .1	Scientific Background	9
	B.2	Data Description	9
	B.3	DisGeNET Score	10
С	Add	itional Figures and Tables	14
D	Exp	erimental Instructions and Interfaces	32
	D.1	No-Data Condition	32
	D.2	Data Condition	41
	D.3	Data Condition with Intermediate Rivalry	50
	D.4	No-Data Condition with Extreme Rivalry	60
	D.5	Questionnaire and Risk-Preferences Elicitation Task	69
	D.6	Payment Information	73

A Experimental Results: Additional Details

A.1 Logistics of the Experiment

Figure A.1 summarizes how our experimental sessions unfolded. When participants join, they are assigned either to a data or to a no-data condition.¹⁰ The experiment begins when a total of ten players are assigned to the same experimental set. Then, from each of these experimental sets, two groups of five people are randomly drawn to play the first five rounds (what we labeled as "block"). At the end of the block, the composition of the two groups is randomly reshuffled, and a second block of five rounds is played. This procedure is repeated a total of four times so that each player ends up playing exactly twenty rounds. The order of blocks seen by participants in different experimental sessions is random. The gem types change each round according to a pre-recorded script generated stochastically so that the actual gems and their values each round are effectively random for the player. Similarly, the payoffs and the specific order in which specific gems are revealed in the treatment condition is generated by a random script before the experiment begins.



Figure A.1: Flowchart of the Experimental Setup.

Note: This figure provides an overview of the experiment for one actual session that took place in December 2024.

¹⁰Not in every experimental session there was a no data condition, in which case the players would be randomly split (and then reshuffled) across two distinct data conditions.

Table A.1 presents the descriptive statistics for the experimental data of the main experiment, in which payoffs are non-rivalrous. Data are shown separately by treatment condition. Overall, the table already shows our main results in terms of payoffs, exploration, and discovery.

	Ν	Mean	SD	Median	Min	Max
Group Payoff (Share)						
Low	80	75.49	10.87	75.11	49	100
Medium	160	67.64	17.01	66.67	30	100
High	80	98.84	4.11	100.00	71	100
No Data	160	73.27	11.82	72.73	36	100
I(Group found max)						
Low	52	100.00	0.00	100.00	100	100
Medium	112	45.54	50.02	0.00	0	100
High	80	100.00	0.00	100.00	100	100
No Data	120	100.00	0.00	100.00	100	100
Mountains Explored (Share)						
Low	80	89.38	13.65	100.00	50	100
Medium	160	45.31	29.16	50.00	0	100
High	80	3.44	11.76	0.00	0	75
No Data	160	83.25	15.36	80.00	40	100

Table A.1: Descriptive Statistics of the Experimental Data.

Note: The table presents descriptive statistics on the 120 participants in the 480 rounds of the experiment with the non-rivalry condition. *Group payoff (Share)*= sum of payoffs as a share of the maximum possible payoff possible in each round; *I(Group found max)*:0/1=1 if the location of the maximum was found by at least one participant in the round; *Group Exploration (Share)*= share of unknown mountains explored in the round.

A.2 Additional Figures and Tables



Figure A.2: Outcomes Over Time by Experimental Condition.

Note: The figures show the impact of data on group outcomes as the experimental session progresses, separately for each block of five rounds. Figure (i) shows the average group payoffs divided by experimental condition. Payoffs are reported as a share of the maximum available in each round. Figure (ii) shows the share of rounds where the maximum was uncovered. Figure (iii) shows the average share of unmapped mountains chosen.

Panel A: Group Payoffs



Figure A.3: Outcomes over Time and by Period of the Game.

Note: Panel A reports the experimental results on group payoffs computed as a share of the maximum possible in each period. Figure (i) shows the average group payoffs achieved in period 1 by experimental condition and over time. Figure (ii) shows the average group payoffs achieved in period 2 by experimental condition and over time. Panel B reports the experimental results on the likelihood of a group breakthrough in each round. Figure (iii) shows the share of rounds in which the maximum was uncovered in period 1 by experimental condition and over time. Figure (iv) shows the shows the share of rounds in which the maximum was uncovered in period 2 by experimental condition and over time.

	Group	Payoff	Group Bre	eakthrough
	(1) (2)		(3)	(4)
	Period 1 Period		Period 1	Period 2
High	33.240***	11.281***	8.765**	-1.500
	(0.344)	(0.760)	(2.440)	(4.277)
Low	0.537	1.008	-1.818	-2.045
	(0.338)	(1.094)	(4.764)	(3.631)
Medium	4.837***	-7.969***	-54.948***	-56.338***
	(0.276)	(0.634)	(3.725)	(5.057)
Session FE	Yes	Yes	Yes	Yes
Block order FE	Yes	Yes	Yes	Yes
Payoff structure FE	Yes	Yes	Yes	Yes
Observations	480	480	364	364

Table A.2: Breaking Down Results by Experimental Period.

Note: † p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors clustered at the session level in parentheses . The sample is at the group-period level (480 rounds). In Columns 3 and 4, the sample only includes rounds that contained at least one diamond (364 rounds). *Group payoff* = group-level period payoffs in Euro; *Group Breakthrough*:0/1=1 if the maximum was found by at least one participant in the period. The excluded category is the control condition without data. See text for more details.

	I(Didn't Choose Medium in Period 1)			
	(1)	(2)	(3)	
Risk aversion (std)	-0.088** (0.023)			
Top quartile risk aversion		-0.133* (0.049)		
Bottom quartile risk aversion			0.169* (0.050)	
Session FE	Yes	Yes	Yes	
Block order FE	Yes	Yes	Yes	
Payoff structure FE	Yes	Yes	Yes	
Round order FE	Yes	Yes	Yes	
Observations	800	800	800	

Table A.3: Risk Aversion and Decision Not to Choose the Known Outcome in Period 1 when Medium Is Revealed.

Note: † p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors clustered at the session level in parentheses . Round-participant level observations, estimates from OLS models. The sample includes all the individual observations for the 160 rounds where the medium value was revealed. *I(Didn't Choose Medium in Period 1)*:0/1=1 if the player did not choose the medium value in period 1. *Risk aversion*= standardized measure of individual risk aversion (Holt and Laury, 2002); *Top quartile risk aversion*:0/1=1 if the participant is in the top quartile of the risk aversion distribution in our sample; *Bottom quartile risk aversion*:0/1=1 if the participant is in the bottom quartile of the risk aversion distribution in our sample.

	I(Didn't	I(Didn't Choose Medium in Period 1)					
	(1)	(2)	(3)	(4)			
English native	-0.064 (0.072)						
Wrong quizzes (std)		0.048 (0.030)					
Round number			-0.002 (0.006)				
Order of choice				0.009 (0.015)			
Session FE	Yes	Yes	Yes	Yes			
Block order FE	Yes	Yes	Yes	Yes			
Payoff structure FE	Yes	Yes	Yes	Yes			
Round order FE	Yes	Yes	Yes	No			
Observations	800	800	800	800			

Table A.4: Correlates of the Decision not to Choose the Known Outcome in Period 1 when Medium Is Revealed.

Note: † p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors clustered at the session level in parentheses . Round-participant level observations, estimates from OLS models. The sample includes all the individual observations for the 160 rounds where the medium value was revealed. *I(Didn't Choose Medium in Period 1)*:0/1=1 if the player did not choose the medium value in period 1. *English native*:0/1=1 if the participant is a native English speaker based on her reported nationality; *Wrong quizzes*= standardized number of wrong answers to the initial comprehension test; *Round number*= progressive order in which the rounds were played in the experimental session; *Order of choice*= random sequential order in which the player chose in that round.

B The Genetic Roots of Human Diseases: Additional Details

B.1 Scientific Background

Genetics is the branch of biology that studies genes, heredity, and variation in living organisms. Genes are segments of DNA (deoxyribonucleic acid) that contain the information necessary for living organisms' development, functioning, and reproduction. In practice, each gene is a portion of DNA that contains instructions for building one or more products, such as proteins, which are the fundamental constituents of an organism. Genes often acquire mutations (or variants) in their sequence, most of which are harmless. However, some mutations can lead the gene to alter its behavior and affect phenotypic traits, sometimes with significant consequences and the emergence of severe health conditions. Discovering which mutations are responsible for specific human diseases is thus a first-order priority since genes associated with a condition can often be used as drug targets (Nelson et al., 2015). When a drug molecule binds to its genetic target, it can modify its functioning, favorably affecting the outcome of a disease. Therefore, knowing the genetic roots of diseases has important practical consequences in the design of pharmaceutical drugs.

Diseases caused by single gene mutations are called Mendelian disorders, but such diseases are typically rare. Most common human diseases have a polygenic nature, meaning they are not due to a single genetic factor but rather by mutations in many genes. This class of diseases is called complex and genetic mutations may increase the risk of developing the condition without being either necessary or sufficient on their own. Despite often clustering in families, polygenic disorders do not have a predictable inheritance pattern because convoluted interactions between genes and environmental factors determine them. This means that scientists need to search through the over 19,000 protein-coding genes to find the mutations involved in each of the thousands of polygenic diseases (Tranchero, 2025).

Researchers have noted that even after the completion of the Human Genome Project, most scientists continue to investigate the same small number of genes (Stoeger et al., 2018). Gates et al. (2021) report that 1% of genes still receive 22% of all gene-related publications, helping to explain why current treatments exploit only around 10% of the potentially druggable targets. This situation is probably suboptimal since our chances of finding a cure for polygenic diseases would benefit from exploring a larger number of genes (Edwards et al., 2011) and several understudied genes showing high promise have been identified (Nguyen et al., 2017; Stoeger et al., 2018). Interestingly, despite much debate on this extreme concentration of attention on a small number of theoretically well-known genes, we still lack an explanation for its drivers. Some scholars have attributed it to scientists' preference for genes with past data that permit the formulation of functional hypotheses (Haynes et al., 2018), akin to what we characterized as a streetlight effect in this paper.

B.2 Data Description

DisGeNET. Our main data source is DisGeNET (v7.0), which is considered a complete repository of scientific results linking human diseases to their genetic causes (Piñero et al., 2020). This database

aggregates all novel gene-disease combinations studied by publications indexed in PubMed. The information is harvested from specialized sources, including curated datasets such as ClinVar, UniProt, and Orphanet.¹¹ In addition, DisGeNET complements these data with information extracted from the scientific literature indexed in PubMed using text-mining approaches. Our starting data are at the gene-disease-paper level, because for each gene-disease pair we observe both the publication that introduced it and the list of all follow-up articles that investigated it.

Genes. Each gene in the database is identified by a unique ID from Entrez Gene, a gene-centric resource maintained by the National Center for Biotechnology Information (NCBI). These identifiers are species-specific, meaning the ID assigned to a human gene differs from that of its homolog in another species. DisGeNET includes only data from studies on human genes and compiles the Entrez Gene ID for each gene examined in PubMed-indexed papers. We further restrict our sample to protein-coding genes, given their central role in the drug discovery process (Nelson et al., 2015).

Diseases. Disease entries in DisGeNET are annotated using vocabulary from the Unified Medical Language System (UMLS), a set of crosswalks that bring together many health and biomedical vocabularies and standards to enable interoperability between databases. DisGeNET compiles the UMLS ID of each disease studied by papers in PubMed. Since we focus on human diseases, we keep any entries that map to the following UMLS semantic types: disease or syndrome; neoplastic process; acquired abnormality; anatomical abnormality; congenital abnormality; and mental or behavioral dysfunction. Using the UMLS ID, we also obtain disease relations from Kehoe and Torvik (2019), which contains all pairwise relationships in the Medical Subject Headings vocabulary (MeSH) hierarchy.

B.3 DisGeNET Score

DisGeNET is designed to help researchers in both academia and industry prioritize promising genetic targets based on existing knowledge. To support this goal, it provides a synthetic DisGeNET Score for each gene–disease pair. The Score ranges from 0 to 1, with higher values indicating combinations that are more scientifically robust and therapeutically promising. It incorporates both the curation and reliability of the sources supporting a given association, as well as the number of publications that have studied it. In practice, the Score reflects how well-established a gene target is in the current literature. In the version used in this paper (v7.0), the score offers a parsimonious way to assess the scientific strength of any given gene–disease pair as of 2020.

In particular, the raw DisGeNET score is build with the following formula:

DisGeNET score of gene i for disease $j = C_{i,j} + M_{i,j} + I_{i,j} + L_{i,j}$

The first component $C_{i,j}$ summarizes the evidence from curated sources reporting gene-disease com-

¹¹For the complete list of sources aggregated by DisGeNET, see https://www.disgenet.org/dbinfo.

bination $\langle i, j \rangle$:

$$C_{i,j} = \begin{cases} 0.6 & \text{if } N_{sources_c} > 2\\ 0.5 & \text{if } N_{sources_c} = 2\\ 0.3 & \text{if } N_{sources_c} = 1\\ 0 & \text{otherwise} \end{cases}$$
(B.1)

where $N_{sources_c}$ is the number of curated sources supporting a gene-disease association, including CGI, ClinGen, Genomics England, CTD, PsyGeNET, Orphanet, and UniProt.

The second component $M_{i,j}$ summarizes the evidence from experiments using mice models reporting gene-disease combination $\langle i, j \rangle$:

$$M_{i,j} = \begin{cases} 0.2 & \text{if } N_{sources_m} > 0\\ 0 & \text{otherwise} \end{cases}$$
(B.2)

where $N_{sources_m}$ is the number of sources using the lab rat or lab mouse from RGD, MGD, and CTD.

The third component $I_{i,j}$ summarizes the evidence inferred from experiments on gene-disease combination $\langle i, j \rangle$:

$$I_{i,j} = \begin{cases} 0.1 & \text{if } N_{sources_i} > 1\\ 0 & \text{otherwise} \end{cases}$$
(B.3)

where $N_{sources_i}$ is the number of sources from HPO, CLINVAR, GWAS Catalog, and GWASDB.

Finally, the component $L_{i,j}$ summarizes the evidence mined from the literature about gene-disease combination $\langle i, j \rangle$:

$$L_{i,j} = \begin{cases} 0.1 & \text{if } N_{publications} > 9\\ \mathbf{N}_{publications} \cdot 0.01 & \text{if } N_{publications} \le 9 \end{cases}$$
(B.4)

where $N_{publications}$ is the number of publications supporting a gene-disease association as mined by LHGDN and BEFREE.

The DisGeNET Score has strong face validity and has been thoroughly validated in prior research (Piñero et al., 2020). Because it is designed to capture the biological importance of a gene-disease pair, we should expect higher-scoring associations to be linked to more downstream pharmaceutical development—such as clinical citations, granted patents, and approved drugs. To test this, we regress the raw DisGeNET score on each of these real-world outcomes. The results, presented in Appendix Table B.1, show that higher scores are associated with significantly greater levels of clinical citations, patenting activity, and drug development.

	Clinical Citations (1) Count (#) (2) Has Any (0/1)		Grant	Granted Patents		oved Drugs
			(3) Count (#)	(4) Has Any (0/1)	(5) Count (#)	(6) Has Any (0/1)
DisGeNET Score	51.428***	0.939***	1.669***	0.153***	0.085***	0.022***
	(2.302)	(0.005)	(0.070)	(0.003)	(0.007)	(0.001)
N	810,377	810,377	810,377	810,377	810,377	810,377

Table B.1: Associations Between DisGeNET Scores and Real-World Pharmaceutical Outcomes

* p < 0.05, ** p < 0.01, *** p < 0.001. Robust standard errors in parentheses. Estimates from OLS models. The sample is at the gene-disease level. We correlate the raw DisGeNET score with real-world measures of clinical impact. *Count Clinical* = total clinical articles based on a gene-disease pair. *Granted Patents* = count of USPTO granted patents for inventions leveraging a given gene as a drug target for a given disease. *Count Drugs* = count of FDA-approved drugs leveraging a given gene as a drug target for a given disease. Models (1), (3), and (5) use count variables, while Models (2), (4), and (6) use corresponding indicator versions.

In our main specification, we convert the raw DisGeNET scores into percentile ranks, and consider any score below the 60th percentile as a low payoff, between the 60th and 90th percentile as a medium payoff, and above the 90th percentile as a high payoff, respectively. We now verify that these score categories correspond to meaningful differences in real-world outcomes. In Appendix Figure B.1, we plot our three outcome metrics by score category. We find that clinical citations, granted patents, and approved drugs are all increasing in score category, suggesting that our score thresholds do capture substantive differences in impact.



Figure B.1: Relationship Between Score Categories and Real-World Outcomes.

Note: This figure shows the relationship between our categorization of raw DisGeNET scores (Low, Medium, High) and real-world innovation outcomes. Panel (i) shows the average number of clinical citations on a gene-disease pair (data as of 2024). Panel (ii) shows the average number of granted patents targeting a gene-disease pair (data as of 2023). Panel (iii) shows the average number of agproved drugs targeting a gene-disease pair (data as of 2023).

One potential limitation is that the DisGeNET score is time-invariant, since it is calculated ex post across the full sample period, incorporating all available evidence up to the time when our data were collected (2020). While this gives us the best available evidence of association strength at the time of analysis, it does raise concerns about the variability of the score over time. In particular, such a bias could arise if scientists initially pursued a genetic target thinking it was high value (H), only for it to be revised downward into the M range at a later moment. Alternatively, researchers may have believed an M-value gene-disease pair had the potential to become an H with further investigation

and behaved accordingly. To assess the validity of this concern, we compare the current DisGeNET scores to those from the first release in 2015 (version 1). As shown in Figure B.2, the ordinal rankings of gene-disease pairs are largely preserved over time, showing a remarkable stability of the DisGeNET scores. To further ease this concern, we replicate our main analysis using the 2015 scores. If the assumption of time-invariance is reasonable, then we should obtain the same results. As reported in Table B.2, the results are consistent with our main analysis.



Figure B.2: Stability of DisGeNET Scores Over Time

Note: This figure presents a binned scatterplot comparing the earliest available DisGeNET scores (version 1, released in 2015) to those from the release used in our main analysis (version 7, released in 2020).

	Group Breakthrough	Group Exploration	Group Delay
	(1) High-Value Gene (0/1)	(2) New Genes/Papers	(3) Years From 1980
Max Found: M	-0.049†	-0.151***	0.939*
	(0.029)	(0.029)	(0.441)
Max Found: H	0.739***	-0.273***	-23.229***
	(0.031)	(0.030)	(0.595)
Disease Class FE	Yes	Yes	Yes
Count of Publications	Yes	Yes	Yes
Ν	3261	3261	3261

Table B.2: Analysis Based on DisGeNET v1 Scores (2015)

Note: † p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors clustered at the disease-class level in parentheses.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification but uses the earliest version of DisGeNET scores (released in 2015). For each human disease, we determine the highest DisGeNET score for any gene identified during the exploration period (i.e., pre-2000). We classify maximum scores below the 60^{th} percentile as L, scores between the 60^{th} and 90^{th} percentile as M, and scores above the 90^{th} percentile as H. *High-Value Gene*: 0/1=1 if any H candidate was discovered for the disease. *New Genes/Papers*= the number of new genes explored per scientific publication in the years following the exploration period. *Years From 1980*= the number of years until the first H candidate is discovered. In all models, diseases in category L serve as the reference group. We include disease-class fixed effects and control for the number of publications post-2000. See text for more details.



C Additional Figures and Tables

Figure C.1: Illustration of the Empirical Setup.

Note: This figure presents a stylized depiction of our approach to translate the theoretical framework to the disease-level data in our sample. For each human disease, we record every gene identified during the early exploration period (i.e., pre-2000). We classify scores below the 60^{th} percentile as L (red), scores between the 60^{th} and 90^{th} percentile as M (yellow), and scores above the 90^{th} percentile as H (green). The highest-scoring genetic target for each disease is then used to classify the nature of early data available to scientists. In this stylized representation, scientists identified three gene-disease pairs with L scores for Disease 1, which means we classify its early data as L. For Disease 2, scientists found one L and one M, resulting in a classification of M. For Disease 3, two L scores and one H score were uncovered, leading to a classification of H. See text for further details.



Figure C.2: Illustration of the Instrumental Variable Strategy.

Note: This figure presents a stylized depiction of our IV approach, which relies on gene orthology (i.e., when genes in different species descend from a common ancestor, largely retaining the same biological function). For each disease, we only consider the gene candidates that have an ortholog in a mouse. We then measure the share of these orthologous genes that are classified as medium-value (M) candidates. In the example above, the M share for Disease 1 would be 75%, while for Disease 2 it would be 25%. See text for further details.



Panel B: Group Exploration



Figure C.3: Reduced-Form Evidence for the Streetlight Effect in the Search for Genetic Candidates.

Note: This figure shows binned scatterplots for each of our dependent variables against our instrumental variable, defined as the share of each disease's orthologous genes that fall into the M category. Panel (i) shows the impact of the instrument on the likelihood of finding any breakthrough during the sample period. Panel (ii) shows the impact of the instrument on the number of new genes explored per publication in the years following the exploration window. Panel (iii) shows the impact of the instrument on the delay in discovering a breakthrough, defined as the years elapsed from 1980 (the first year of our panel). We include controls for disease class and the number of publications post-2000. See text for more details.

Panel A: Keeping Sibling and Parent Diseases



Figure C.4: Considering only Diseases Genetically Related to a Disease with a Breakthrough.

Note: This figure replicates the event study of figure 5 but only considers diseases that are genetically related to a disease with a known breakthrough (genetic discoveries with scores above the 90th percentile of DisGeNET scores). We obtain genetic relations from the Medical Subject Headings vocabulary (MeSH). In Panel A, we keep both sibling diseases (i.e., diseases sharing the same parent MeSH code) and parent diseases (i.e., diseases one level up in the MeSH tree) of diseases with a breakthrough. In Panel B, we keep only sibling diseases (i.e., diseases sharing the same parent MeSH code) of diseases with a breakthrough. This figure plots OLS estimates and 95% confidence intervals from an event study design that explores how genetic exploration in each disease evolves in the years before and after the discovery of the first medium-value genetic association. Standard errors are clustered at the disease class level. See text for more details.

Table C.1: Association Between Genes Having Mouse Orthologs and Their Appearance in the Scientific Literature.

	(1) Publication Year	(2) Publication Year
Has Mouse Ortholog (0/1)	-0.798**	-2.559***
	(0.262)	(0.372)
Gene Group FE	No	Yes
Ν	16,000	10,344

Panel A: Gene Level

Panel B: Gene-Disease Level						
	(1) Publication Year	(2) Publication Year				
Has Mouse Ortholog (0/1)	-0.00385***	-0.0122***				
	(0.0000985)	(0.000169)				
Disease FE	Yes	Yes				
Gene Group FE	No	Yes				
Ν	339,136,000	257,764,556				

* p < 0.05, ** p < 0.01, *** p < 0.001. Robust standard errors in parentheses.

This table examines whether human genes with orthologous counterparts in the lab mice are explored earlier by scientists. In Panel A, the data is at the gene level. We assess whether genes with an ortholog appear in scientific studies earlier regardless of the disease. In Panel B, the data is at the gene-disease level. We assess whether genes with an ortholog appear in scientific studies earlier for a given disease. We impute a value of 2020 for gene-disease pairs with no recorded publications. In all models, the dependent variable is the first year of publication. In Column (2) of both Panel A and Panel B, we include controls for gene group classification.

	Mean	Median	Sd	Min	Max	Ν
Maximum Gene Score	39.68	0.00	45.56	0	100	220,760
Found Any High Gene (0/1)	0.28	0.00	0.45	0	1	220,760
Count of Publications	7.37	0.00	82.60	0	10,449	220,760
Count of Genes Discovered	3.25	0.00	15.46	0	685	220,760
New Genes per Paper	0.74	0.67	1.03	0	123	109,002

Table C.2: Descriptive Statistics at the Disease-Year Level.

Note: This table presents descriptive statistics for our disease-year panel. *Maximum Gene Score* = the highest DisGeNET score uncovered each year for a disease. *Found Any High Gene*:0/1=1 if any *H* was discovered a year for a given disease. *Count of Publications* = the number of publications on the disease in a given year. *Count of Genes Discovered* = the number of genes explored in relation to a disease each year. *New Genes Per Paper* = the number of genes explored per paper for a disease each year.

	Group Exploration	Group Exploration
	(1) New Genes/Papers	(2) New Genes/Papers
Post M Discovery	-0.180***	
	(0.030)	
Post H Discovery		-0.256***
-		(0.022)
Disease FE	Yes	Yes
Year FE	Yes	Yes
Count of Publications	Yes	Yes
Ν	98,547	97,956

Table C.3: Disease-Year-Level Analysis of Exploration Dynamics.

OLS estimates from differences-in-differences models. The sample is at the disease-year level. We examine how genetic exploration within each disease evolves following the discovery of the first medium-value (Column 1) and high-value (Column 2) genetic target. For each disease, we classify DisGeNET scores below the 60th percentile as "low," scores between the 60th and 90th percentiles as "medium," and scores above the 90th percentile as "high" (or breakthrough) discoveries. *Yearly Genes/Papers*= the number of new genes explored per scientific publication. All models include disease fixed effects and year fixed effects, and control for the annual number of publications. See text for more details.

	csdid		did_mu	did_multiplegt		outation
	(1) Genes/Papers	(2) Genes/Papers	(3) Genes/Papers	(4) Genes/Papers	(5) Genes/Papers	(6) Genes/Papers
Post M Discovery	-0.352***		-0.238**		-0.246***	
	(0.069)		(0.081)		(0.037)	
Post H Discovery		-0.223***		-0.187***		-0.281***
2		(0.031)		(0.031)		(0.029)
Disease FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Count of Publications	Yes	Yes	Yes	Yes	Yes	Yes
Ν	56870	49812	48619	32415	64503	59402

Table C.4:	Robustness to	TWFE	Weighting	Concerns.
			<i>L</i>) <i>L</i>)	

OLS estimates from differences-in-differences models. The sample is at the disease-year level. We replicate Table C.4 but use alternate estimators that avoid weighting problems associated with TWFE. We implement the csdid command from Callaway and Sant'Anna (2020) in Columns 1-2, the did_multiplegt_dn command from Chaisemartin and D'Haultfoeuille (2024) in Columns 3-4, and the did_imputation command from Borusyak et al. (2021) in Columns 5-6. For each disease, we classify DisGeNET scores below the 60th percentile as "low," scores between the 60th and 90th percentiles as "medium," and scores above the 90th percentile as "high" (or breakthrough) discoveries. *Yearly Genes/Papers*= the number of new genes explored per scientific publication. All models include disease fixed effects and year fixed effects, and control for the annual number of publications. See text for more details.

	Siblings and Parents	Siblings Only
	(1) Yearly Genes/Papers	(2) Yearly Genes/Papers
Post M Discovery	-0.251***	-0.287**
	(0.061)	(0.092)
Disease FE	Yes	Yes
Year FE	Yes	Yes
Count of Publications	Yes	Yes
Ν	40505	24416

Table C.5: Considering only Diseases Genetically Related to a Disease with a Breakthrough.

OLS estimates from differences-in-differences models. The sample is at the disease-year level. We replicate Table C.5, but only considers diseases in the sample that are genetically related to a disease with a known breakthrough (genetic discoveries with scores above the 90th percentile of DisGeNET score). We obtain genetic relations from the Medical Subject Headings vocabulary (MeSH). In Column 1, we keep both sibling diseases (i.e., diseases sharing the same parent MeSH code) and parent diseases (i.e., diseases one level up in the MeSH tree) of diseases with a breakthrough. In Column 2, we keep only sibling diseases (i.e., diseases sharing the same parent MeSH code) of diseases with a breakthrough. For each disease, we classify DisGeNET scores below the 60th percentile as "low," scores between the 60th and 90th percentiles as "medium," and scores above the 90th percentile as "high" (or breakthrough) discoveries. *Yearly Genes/Papers* = the number of new genes explored per scientific publication. All models include disease fixed effects and year fixed effects, and control for the annual number of publications. See text for more details.

	Group Breakthrough			Group Exploration			Group Delay		
	(1) >5 Pubs	(2) >15 Pubs	(3) >25 Pubs	(4) >5 Pubs	(5) >15 Pubs	(6) >25 Pubs	(7) >5 Pubs	(8) >15 Pubs	(9) >25 Pubs
Max Found: M	-0.043	-0.121**	-0.162***	-0.205***	-0.131***	-0.109***	0.883*	2.056***	2.773***
	(0.028)	(0.038)	(0.047)	(0.019)	(0.027)	(0.032)	(0.425)	(0.583)	(0.745)
Max Found: H	0.602***	0.477***	0.422***	-0.315***	-0.260***	-0.243***	-21.400***	-19.910***	-19.271***
	(0.037)	(0.045)	(0.056)	(0.029)	(0.033)	(0.038)	(0.594)	(0.729)	(0.909)
Disease Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Count of Publications	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ν	5641	4149	3355	5641	4149	3355	5641	4149	3355

Table C.6: Sensitivity to Definition of Marginally Explored Diseases.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification (which removes diseases with less than 10 publications over the sample window) and shows robustness when we keep only diseases with more than 5 publications (Columns (1), (4), and (7)), more than 15 publications (Columns (2), (5), and (8)), and more than 25 publications (Columns (3), (6), and (9)). For each human disease, we determine the highest DisGeNET score for any gene identified during the exploration period (i.e., pre-2000). We classify maximum scores below the 60^{th} percentile as L, scores between the 60^{th} and 90^{th} percentile as M, and scores above the 90^{th} percentile as H. Columns 1-3 show the impact of early discoveries on the likelihood of finding any breakthrough during the sample period. Columns 4-6 show the impact on the number of new genes explored per scientific publication. Columns 7-9 show the impact of the delay in discovering a breakthrough, defined as the years that elapsed from 1980 (the first year of our panel). In all models, diseases classified under L constitute the excluded category. We include disease-class fixed effects and control for the number of publications post-2000. See text for more details.

	Group Breakthrough	Group Exploration	Group Delay
	(1) High-Value Gene (0/1)	(2) New Genes/Papers	(3) Years From 1980
Max Found: M	-0.105**	-0.129***	1.878***
	(0.033)	(0.023)	(0.514)
Max Found: H	0.503***	-0.200***	-19.465***
	(0.042)	(0.036)	(0.642)
Disease Class FE	Yes	Yes	Yes
Count of Publications	Yes	Yes	Yes
Ν	4675	4675	4675

Table C.7: Sensitivity to the Exclusion of Outlier Diseases.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification but excluding outlier diseases (i.e., those in the top 1% by publications over the sample period). For each human disease, we determine the highest DisGeNET score for any gene identified during the exploration period (i.e., pre-2000). We classify maximum scores below the 60^{th} percentile as L, scores between the 60^{th} and 90^{th} percentile as M, and scores above the 90^{th} percentile as H. High-Value Gene: 0/1=1 if any H candidate was discovered for the disease. New Genes/Papers= the number of new genes explored per scientific publication in the years following the exploration period. Years From 1980= the number of years until the first H candidate is discovered. In all models, diseases in category L serve as the reference group. We include disease-class fixed effects and control for the number of publications post-2000. See text for more details.

	Gro	Group Breakthrough			Group Exploration			Group Delay		
	(1) 60 th P	(2) 70 th P	(3) 80 th P	(4) 60 th P	(5) 70 th P	(6) 80 th P	(7) 60 th P	(8) 70 th P	(9) 80 th P	
Max Found: M	-0.105**	-0.072**	-0.082***	-0.144***	-0.180***	-0.259***	1.743***	1.200**	1.303***	
	(0.033)	(0.025)	(0.024)	(0.023)	(0.020)	(0.019)	(0.519)	(0.425)	(0.391)	
Max Found: H	0.514***	0.549***	0.555***	-0.261***	-0.261***	-0.269***	-20.371***	-20.961***	-21.093***	
	(0.042)	(0.033)	(0.031)	(0.028)	(0.027)	(0.029)	(0.692)	(0.589)	(0.537)	
Disease Class FE	Yes									
Count of Publications	Yes									
Ν	4760	4760	4760	4760	4760	4760	4760	4760	4760	

Table C.8: Alternative Definitions of Low and Medium-Value Genes.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification but varies the cutoff between a low and medium-value genetic association. In our baseline, we adopt the 60^{th} percentile to separate medium and high scores. We test the baseline (Columns (1), (4), and (7)), the 70th percentile (Columns (2), (5), and (8)), and the 80th percentile (Columns (3), (6), and (9)) instead. For each model, we hold the cutoff between a medium gene score and a high gene score fixed at the 90th percentile. For each human disease, we determine the highest DisGeNET score for any gene identified during the exploration period (i.e., pre-2000). Columns 1-3 show the impact of early discoveries on the likelihood of finding any breakthrough during the sample period. Columns 4-6 show the impact on the number of new genes explored per scientific publication in the years following the exploration period. Columns 7-9 show the impact on the delay in discovering a breakthrough, defined as the years that elapsed from 1980 (the first year of our panel). In all models, diseases classified under *L* constitute the excluded category. We include disease-class fixed effects and control for the number of publications post-2000. See text for more details.

	Gr	Group Breakthrough			Group Exploration			Group Delay		
	$(1) 90^{th} P$	(2) 95th P	$(3) 100^{th} P$	$(4) 90^{th} P$	(5) 95 th P	(6) 100^{th} P	(7) 90 th P	(8) 95 th P	(9) 100 th P	
Max Found: M	-0.105**	-0.023	-0.053*	-0.144***	-0.110***	-0.142***	1.743***	0.588	1.048**	
	(0.033)	(0.026)	(0.022)	(0.023)	(0.023)	(0.023)	(0.519)	(0.388)	(0.350)	
Max Found: H	0.514***	0.667***	0.802***	-0.261***	-0.364***	-0.466***	-20.371***	-21.905***	-23.137***	
	(0.042)	(0.036)	(0.034)	(0.028)	(0.027)	(0.027)	(0.692)	(0.652)	(0.622)	
Disease Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Count of Publications	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
N	4760	4760	4760	4760	4760	4760	4760	4760	4760	

Table C.9: Alternative Definitions of Medium and High-Value Genes.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification but varies the cutoff between a medium and high-value genetic association. In our baseline, we adopt the 90th percentile to separate medium and high scores. We test the baseline (Columns (1), (4), and (7)), the 95th percentile (Columns (2), (5), and (8)), and the 100th percentile (Columns (3), (6), and (9)). For each model, we hold the cutoff between a low gene score and a medium gene score fixed at the 60th percentile (our baseline). Columns 1-3 show the impact of early discoveries on the likelihood of finding any breakthrough during the sample period. Columns 4-6 show the impact on the number of new genes explored per scientific publication in the years following the exploration period. Columns 7-9 show the impact on the delay in discovering a breakthrough, defined as the years that elapsed from 1980 (the first year of our panel). In all models, diseases classified under *L* constitute the excluded category. We include disease-class fixed effects and control for the number of publications post-2000. See text for more details.

	Group Breakthrough			Group Exploration			Group Delay		
	(1) <1995	(2) <2000	(3) <2005	(4) <1995	(5) <2000	(6)<2005	(7) <1995	(8) <2000	(9)<2005
Max Found: M	-0.024	-0.105**	-0.100**	-0.150***	-0.144***	-0.223***	0.646	1.743***	1.172**
	(0.036)	(0.033)	(0.036)	(0.021)	(0.023)	(0.028)	(0.663)	(0.519)	(0.435)
Max Found: H	0.435*** (0.045)	0.514*** (0.042)	0.656*** (0.037)	-0.223*** (0.027)	-0.261*** (0.028)	-0.337*** (0.034)	-20.104*** (0.863)	-20.371*** (0.692)	-22.368*** (0.473)
Disease Class FE Count of Publications N	Yes Yes 3385	Yes Yes 4760	Yes Yes 4756	Yes Yes 3385	Yes Yes 4760	Yes Yes 4756	Yes Yes 3385	Yes Yes 4760	Yes Yes 4756

Table C.10: Alternative Definitions of the Early Exploration Period.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification using alternative windows to define the period of early search. We report the results including all years before 1995 (Columns (1), (4), and (7)), the baseline (Columns (2), (5), and (8)), and before 2005 (Columns (3), (6), and (9)). For each human disease, we determine the highest DisGeNET score for any gene identified during the exploration period. We classify maximum scores below the 60^{th} percentile as L, scores between the 60^{th} and 90^{th} percentile as M, and scores above the 90^{th} percentile as H. Columns 1-3 show the impact of early discoveries on the likelihood of finding any breakthrough during the sample period. Columns 4-6 show the impact on the number of new genes explored per scientific publication in the years following the exploration period. Columns 7-9 show the impact on the delay in discovering a breakthrough, defined as the years that elapsed from 1980 (the first year of our panel). In all models, diseases classified under L constitute the excluded category. We include disease-class fixed effects and control for the number of publications in the post-exploration period. See text for more details.

	Group Breakthrough		Group Exploration			Group Delay			
	(1) 5%	(2) 10%	(3) 15%	(4) 5%	(5) 10%	(6) 15%	(7) 5%	(8) 10%	(9) 15%
Max Found: M	-0.086***	-0.117***	-0.101**	-0.103**	-0.149***	-0.137***	3.619***	3.203***	2.645***
	(0.026)	(0.032)	(0.033)	(0.031)	(0.034)	(0.031)	(0.671)	(0.686)	(0.719)
Max Found: H	0.332*** (0.035)	0.456*** (0.040)	0.549*** (0.037)	-0.168*** (0.022)	-0.243*** (0.025)	-0.245*** (0.025)	-14.057*** (0.937)	-17.297*** (0.893)	-18.972*** (0.842)
Disease Class FE Count of Publications N	Yes Yes 4761	Yes Yes 4761	Yes Yes 4761	Yes Yes 4761	Yes Yes 4761	Yes Yes 4761	Yes Yes 4761	Yes Yes 4761	Yes Yes 4761

Table C.11: Using Share of Publications on Disease to Define the Exploration Period.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification using a disease-specific definition of the early exploration period. We compute the share of total publications on a given disease that were published by a specific year. We then define the end of the exploration period as the year 5% of publications were published (Columns (1), (4), and (7)), the year 10% of publications were published (Columns (2), (5), and (8)), and the year 15% of publications were published (Columns (3), (6), and (9)). For each human disease, we determine the highest DisGeNET score for any gene identified during the exploration period (which varies by disease). We classify maximum scores below the 60^{th} percentile as L, scores between the 60^{th} percentile as M, and scores above the 90th percentile as H. Columns 1-3 show the impact of early discoveries on the likelihood of finding any breakthrough during the sample period. Columns 4-6 show the impact on the number of new genes explored per scientific publication in the years following the exploration period. Columns 7-9 show the impact on the delay in discovering a breakthrough, defined as the years that elapsed from 1980 (the first year of our panel). In all models, diseases classified under L constitute the excluded category. We include disease-class fixed effects and control for the number of publications in the post-exploration period. See text for more details.

		New Genes/Papers						
	All Years	5 Years	10 Years	Until H				
	(1)	(2)	(3)	(4)				
Max Found: M	-0.144*** (0.023)	-0.246*** (0.040)	-0.204*** (0.035)	-0.172*** (0.028)				
Max Found: <i>H</i>	-0.261*** (0.028)	-0.454*** (0.046)	-0.353*** (0.031)					
Disease Class FE Count of Publications N	Yes Yes 4760	Yes Yes 4495	Yes Yes 4715	Yes Yes 1778				

Table C.12: Using Alternative Windows to Examine Follow-on Exploration.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification using alternative windows to evaluate exploration dynamics after a genetic discovery. We report the results from the baseline (Column (1)), the 5 subsequent years after the year 2000 (Column (2)), the 10 subsequent years after the year 2000 (Column (3)), and until the first high gene score is found (Column (4)). For each human disease, we determine the highest DisGeNET score for any gene identified during the exploration period (i.e., pre-2000). We classify maximum scores below the 60^{th} percentile as L, scores between the 60^{th} and 90^{th} percentile as M, and scores above the 90^{th} percentile as H. Each model shows the impact on the number of new genes explored per scientific publication in the years following the exploration period. In all models, diseases classified under L constitute the excluded category. We include disease-class fixed effects and control for the number of publications post-2000. See text for more details.

	Group E	Group Delay	
	(1) 5 Years After	(2) All Years After	(3) Years From 1980
Max Found: M	-0.142*	-0.074^{\dagger}	0.747
	(0.062)	(0.040)	(0.468)
Max Found: H	-0.460***	-0.220***	-13.588***
	(0.062)	(0.036)	(0.457)
Disease Class FE	Yes	Yes	Yes
Count of Publications	Yes	Yes	Yes
Ν	3442	3581	3581

Table C.13: Considering only Diseases that Have a Breakthrough by the End of the Sample Period.

Estimates from OLS models. The sample is at the disease-level. This table replicates our baseline specification removing any diseases without a breakthrough during the sample period. For each human disease, we determine the highest DisGeNET score for any gene identified during the exploration period (i.e., pre-2000). We classify maximum scores below the 60^{th} percentile as L, scores between the 60^{th} and 90^{th} percentile as M, and scores above the 90^{th} percentile as H. Column 1 shows the impact on the number of new genes explored per publication in the 5 years following the exploration window, while Column 2 shows the impact on the number of new genes explored per publication in all years following the exploration window. Column 3 shows the impact on the delay in discovering a breakthrough, defined as the years that elapsed from 1980 (the first year of our panel). In all models, diseases classified under L constitute the excluded category. We include disease-class fixed effects and control for the number of publications post-2000. See text for more details.

Panel A: Pr	edicting H	Panel B: Pre	edicting M
	H Gene (0/1)		M Gene (0/1)
Max Sibling: L	0.011*** (0.002)	Max Sibling: L	0.013*** (0.002)
Max Sibling: M	0.025*** (0.002)	Max Sibling: M	0.144*** (0.002)
Max Sibling: <i>H</i>	0.231*** (0.003)	Max Sibling: <i>H</i>	0.034*** (0.003)
N	810,377	N	810,377

Table C.14: Predicting Scientific Value of Gene-Disease Pairs from Related Conditions

* p < 0.05, ** p < 0.01, *** p < 0.001. Robust standard errors in parentheses.

This table examines whether a gene-disease pair is more likely to be classified as H (Panel A) or M (Panel B) based on the strength of the gene's associations with related diseases (sharing the same parent disease in the MeSH taxonomy). The data is at the gene-disease level. For each gene-disease pair, we record the highest DisGeNET score between the gene in question and any disease classified as a sibling of the focal disease. We classify maximum scores below the 60^{th} percentile as L, scores between the 60^{th} and 90^{th} percentile as M, and scores above the 90^{th} percentile as H. H Gene: 0/1=1 if the pair is classified as H, and M Gene: 0/1=1 if the pair is classified as M. In all models, gene-disease associations for which no sibling score is found constitute the excluded category.

D Experimental Instructions and Interfaces

D.1 No-Data Condition

Instructions

General Information

Welcome. This is an experiment in the economics of decision-making. If you pay close attention to these instructions, you can earn a significant amount of money paid to you at the end of the experiment via bank transfer.

To participate in this online experiment, you will need to use your notebook or personal computer (mobile phones are not supported). If you are using a device that is not supported, please copy the experiment link, open a notebook or pc and paste the link into the address bar.

Your computer screen will display useful information. Remember that the information on your computer screen is PRIVATE. To ensure best results for yourself and accurate data for the experimenters, please DO NOT COMMUNICATE or interact with other people on other media at any point during the experiment. If you have any questions, or need assistance of any kind, please call +43-678-780-7284 or use Zoom anytime during the experiment and one of the experimenters will help you privately. We expect the entire experiment to take up to 60 minutes to complete.

Following these instructions, you will be asked to make some choices. There are no correct choices. Your choices depend on your preferences and beliefs, so different participants will usually make different choices. You will be paid according to your choices, so read these instructions carefully and think before you decide.

The Basic Idea

There are 5 mountains and each of them hides one type of gem, which can only be found by exploring the mountain.



V > V > W

You choose which mountains to explore and the value of the gems you find are your earnings in dollars.

How the Gems Are Distributed

You will not know where the gems are hidden from the outset. At the beginning of every round, a gem for each mountain will be randomly drawn, so any gem could be hidden in any mountain.

For each mountain, there is a:

- · 60% chance it contains a topaz
- · 20% chance it contains a ruby



20% chance it contains a diamond

These chances are the same for all five mountains. Hence, there is some chance that there could be more than one diamond, but there is also some chance that there could be no diamond. Further, even if, for example, the first two mountains happen to contain a diamond, the chance that the third mountain contains a diamond is still 20%.

How Participants Choose Mountains

In each round, participants choose which mountain to explore. The choice does not happen simultaneously, but participants choose sequentially, one after the other, according to a random order that changes every round. You can choose to explore any mountain you wish. If you choose the same mountain chosen by other participants, each of you will receive the gem's value uncovered. Similarly, if someone else chooses the same mountain that you previously chose, you will still receive the full gem's value (and so will the other participant(s) that chose it).

To repeat, no participant has any initial information in Stage 1 on the location of gems.

Each Round Has 2 Stages

A round consists of 2 stages. At the beginning of a new round, gems are redrawn for each of the five mountains. The position of gems will **not** be reset between the two stages in a round.

In Stage 1, all participants sequentially choose one mountain to explore. Before choosing a mountain, you will see which mountains have been selected by the other participants in your group who chose before you, and how many participants have selected each mountain. You can choose the same mountain or a different mountain.

At the end of Stage 1, the gems hidden in each mountain selected by all participants in Stage 1 are revealed, and you earn the value of the gem hidden in the mountain you chose.

In Stage 2, you can again choose any of the same five mountains; that is you can either choose the same mountain of Stage 1 or switch to another one. The position of gems remains the same as in Stage 1, but this time you will also see the gems located in the mountains revealed in Stage 1 in addition to the mapped mountain.

At the end of Stage 2, the gems hidden in each mountain selected by all participants in Stage 2 are revealed, and you earn the value of the gem hidden in the mountain you chose in Stage 2. You will also see your total earnings for the round which equals the sum of the value of the gem you found in Stage 1 and the value of the gem you found in Stage 2.

Game Structure

The game is divided into 4 blocks, each made of 5 rounds, with each round encompassing the two stages described above. At the beginning of each block, you will be randomly assigned to a new group of 5 participants, with whom you will play for the entire block (5 rounds in total). After the block is complete, you will be randomly assigned to a new group of 5 participants. Again, you will play for 5 rounds. This procedure will be repeated 4 times in total.

You will be reminded of this information in the top-right corner of your screen, as in the example below:





Payment

Fixed Participation Fee: You will earn a participation fee of \$5.00 for participating in this experiment.

Additional Payment and Random Round: One round will be randomly selected for payment at the end of the experiment. You will be paid and your earnings in that round as described above. Any of the 20 rounds (4 blocks with 5 rounds each) could be the one selected, so you should treat each round as if it will be the one determining your payment. This protocol of determining payments suggests that you should choose in each round as if it is the only round that determines your payment as the dollar value of the gems you select will directly translate into your earnings.

Survey and Payment: In addition to the participation fee and the payment for the randomly selected round, you will perform a small task at the very end of the experiment, and your earnings from this task will be paid to you.

You will be informed of your payment and the round chosen for payment at the end of the experiment. The \$ you have earned will be converted into Euros at an exchange rate of \$ 1 = € 0.67. Finally, after completing the experiment you will be paid electronically via bank transfer.

Frequently Asked Questions

Q1: Is this some kind of psychology experiment with an agenda you haven't told us?

A: No, it is an economics experiment. If we do anything deceptive or don't pay you as described, then you can complain to the University of Toronto Research Ethics Board and we will be in serious trouble. These instructions are meant to clarify how you earn money and our interest is in seeing how people make decisions.

Q2: Is there a "correct" or "wrong" choice of action? Is this kind of a test?A: No, your optimal choice depends on your preferences and beliefs and different people may hold different beliefs.

Next

This button will be activated after 281 seconds. Please take your time to read through the instructions.

Instructions Quiz Experiment Questionnaire

You have successfully finished reading the instructions.

The quiz, consisting of 8 questions in total, follows.

Next

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q1: In each round, you will select two mountains (one in Stage 1, and one in Stage 2) and collect the gem that they hide. You can choose the same mountain in both stages, or change after Stage 1." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q2: If more than one player selects the same mountain, they will all collect the full value of the gem." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q3: At the beginning of a new round, the gems are redrawn for each mountain." :

- Correct
- Incorrect







Next

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q4: No group member has any private initial information in Stage 1 on the location of gems." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q5: The position of gems will not be reset between the two stages of a round." :

Correct

Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q6: All group members select the mountains simultaneously." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q7: If another group member chose a mountain before you, you cannot choose it again." :

- Correct
- Incorrect

Read Instructions

Next

Next

Next

Next
Quiz Time!

Please mark the following statements as correct/incorrect:

"Q8: At the end of the experiment, one round will be randomly selected for payment." :

- Correct
- \bigcirc Incorrect



Next

You have successfully finished the quiz.

The experiment follows: When you are ready please click "Next" to start the experiment.

Next

Start of Block 1

This is Block 1 of 4 and each Block consists of 5 Rounds.

You have been randomly assigned to a **new** group of 5 participants.



Start of Round 1

You are now in Round 1 of 5 and each Round consists of 2 Stages.

The computer redrew the gems for each mountain.

No participant has any initial information on the location of gems.

In this round, for each mountain, there could be:

a topaz worth \$1.00 with 60% chance
a ruby worth \$6.00 with 20% chance

💎 : a diamond worth **\$11.00** with **20%** chance

Stage 1

_

Next



In this round, for each mountain, there could be:

😥 : a topaz worth **\$1.00** with **60%** chance

🕦 : a ruby worth **\$6.00** with **20%** chance

💎 : a diamond worth **\$11.00** with **20%** chance

The location of gems is random and no participant has any initial information where each gem is hidden.

It is NOT your turn yet, please wait.



Read Instructions

Stage 1

This is Block 1 of 4: You are in Round 1 of 5.



In this round, for each mountain, there could be:

😥 : a topaz worth **\$1.00** with **60%** chance

() : a ruby worth **\$6.00** with **20%** chance

💎 : a diamond worth **\$11.00** with **20%** chance

The location of gems is random and no participant has any initial information where each gem is hidden.

Now it is YOUR TURN, please select a mountain.



Stage 1: Earnings

You selected Mountain 3 and found a 💎. Thus, you earned \$11.00 from your choice.

All discovered gems and their locations are highlighted below. These will also be displayed in Stage 2 when you make your next choice.

Click "Next" to proceed to the next stage.



Read Instructions

Stage 2

This is Block 1 of 4: You are in Round 1 of 5.



In this round, for each mountain, there could be:

- 🚯 : a topaz worth **\$1.00** with **60%** chance
- 前 : a ruby worth **\$6.00** with **20%** chance
- 💎 : a diamond worth **\$11.00** with **20%** chance

Now it is YOUR TURN, please select a mountain.



Read Instructions

Confirm your mountain choice

Stage 2: Earnings

You selected Mountain 3 and found a 💎. Thus, you earned \$11.00 from your choice.

Your total earnings from both stages in this round are \$11.00 + \$11.00 = \$22.00

All discovered gems and their locations in both Stages are highlighted below.

Please click "Next" to proceed to the next round.













D.2 Data Condition

Instructions

General Information

Welcome. This is an experiment in the economics of decision-making. If you pay close attention to these instructions, you can earn a significant amount of money paid to you at the end of the experiment via bank transfer.

To participate in this online experiment, you will need to use your notebook or personal computer (mobile phones are not supported). If you are using a device that is not supported, please copy the experiment link, open a notebook or pc and paste the link into the address bar.

Your computer screen will display useful information. Remember that the information on your computer screen is PRIVATE. To ensure best results for yourself and accurate data for the experimenters, please DO NOT COMMUNICATE or interact with other people on other media at any point during the experiment. If you have any questions, or need assistance of any kind, please call +43-678-780-7284 or use Zoom anytime during the experiment and one of the experimenters will help you privately. We expect the entire experiment to take up to 60 minutes to complete.

Following these instructions, you will be asked to make some choices. There are no correct choices. Your choices depend on your preferences and beliefs, so different participants will usually make different choices. You will be paid according to your choices, so read these instructions carefully and think before you decide.

The Basic Idea

There are 5 mountains and each of them hides one type of gem, which can only be found by exploring the mountain.



rubies and the rubies are always worth more than the topazes:



You choose which mountains to explore and the value of the gems you find are your earnings in dollars.

How the Gems Are Distributed

You will not know where the gems are hidden from the outset. At the beginning of every round, a gem for each mountain will be randomly drawn, so any gem could be hidden in any mountain.

For each mountain, there is a:

60% chance it contains a topaz

20% chance it contains a ruby



20% chance it contains a diamond

These chances are the same for all five mountains. Hence, there is some chance that there could be more than one diamond, but there is also some chance that there could be no diamond. Further, even if, for example, the first two mountains happen to contain a diamond, the chance that the third mountain contains a diamond is still 20%.

The Map

At the beginning of each round, **one** mountain will be randomly selected to be mapped and its gem will be uncovered to all participants. Each participant will be able to see the same gem contained by the mountain. The mountain chosen for mapping is random and changes in each round. Besides the map, no participant has any other initial information in Stage 1 on the location of gems.

How Participants Choose Mountains

In each round, participants choose which mountain to explore. The choice does not happen simultaneously, but participants choose sequentially, one after the other, according to a random order that changes every round. You can choose to explore any mountain you wish or select the mapped mountain. If you choose the same mountain chosen by other participants, each of you will receive the gem's value uncovered. Similarly, if someone else chooses the same mountain that you previously chose, you will still receive the full gem's value (and so will the other participants) that chose it). To repeat, all participant have the same information in Stage 1 on the location of one of the gems.

Each Round Has 2 Stages

A round consists of 2 stages. At the beginning of a new round, gems are redrawn for each of the five mountains. The position of gems will **not** be reset between the two stages in a round.

Then, one of the mountains randomly selected for mapping and the gem it hides is revealed to all players.

In Stage 1, all participants sequentially choose one mountain to explore. Before choosing a mountain, you will see which mountains have been selected by the other participants in your group who chose before you, and how many participants have selected each mountain. You can choose the same mountain or a different mountain.

At the end of Stage 1, the gems hidden in each mountain selected by all participants in Stage 1 are revealed, and you earn the value of the gem hidden in the mountain you chose.

In Stage 2, you can again choose any of the same five mountains; that is you can either choose the same mountain of Stage 1 or switch to another one. The position of gems remains the same as in Stage 1, but this time you will also see the gems located in the mountains revealed in Stage 1 in addition to the mapped mountain.

At the end of Stage 2, the gems hidden in each mountain selected by all participants in Stage 2 are revealed, and you earn the value of the gem hidden in the mountain you chose in Stage 2. You will also see your total earnings for the round which equals the sum of the value of the gem you found in Stage 1 and the value of the gem you found in Stage 2.

Game Structure

The game is divided into 4 blocks, each made of 5 rounds, with each round encompassing the two stages described above. At the beginning of each block, you will be randomly assigned to a new group of 5 participants, with whom you will play for the entire block (5 rounds in total). After the block is complete, you will be randomly assigned to a new group of 5 participants. Again, you will play for 5 rounds. This procedure will be repeated 4 times in total.

You will be reminded of this information in the top-right corner of your screen, as in the example below:



Payment

Fixed Participation Fee: You will earn a participation fee of \$5.00 for participating in this experiment.

Additional Payment and Random Round: One round will be randomly selected for payment at the end of the experiment. You will be paid and your earnings in that round as described above. Any of the 20 rounds (4 blocks with 5 rounds each) could be the one selected, so you should treat each round as if it will be the one determining your payment. This protocol of determining payments suggests that you should choose in each round as if it is the only round that determines your payment as the dollar value of the gems you select will directly translate into your earnings.

Survey and Payment: In addition to the participation fee and the payment for the randomly selected round, you will perform a small task at the very end of the experiment, and your earnings from this task will be paid to you.

You will be informed of your payment and the round chosen for payment at the end of the experiment. The \$ you have earned will be converted into Euros at an exchange rate of \$ 1 = € 0.67. Finally, after completing the experiment you will be paid electronically via bank transfer.

Frequently Asked Questions

Q1: Is this some kind of psychology experiment with an agenda you haven't told us? A: No, it is an economics experiment. If we do anything deceptive or don't pay you as described, then you can complain to the University of Toronto Research Ethics Board and we will be in serious trouble. These instructions are meant to clarify how you earn money and our interest is in seeing how people make decisions.

Q2: Is there a "correct" or "wrong" choice of action? Is this kind of a test?A: No, your optimal choice depends on your preferences and beliefs and different people may hold different beliefs.

Next

This button will be activated after 281 seconds. Please take your time to read through the instructions.



You have successfully finished reading the instructions.

The quiz, consisting of 8 questions in total, follows.



Quiz Time!

Please mark the following statements as correct/incorrect:

"Q1: In each round, you will select two mountains (one in Stage 1, and one in Stage 2) and collect the gem that they hide. You can choose the same mountain in both stages, or change after Stage 1." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q2: If more than one player selects the same mountain, they will all collect the full value of the gem." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q3: At the beginning of a new round, the gems are redrawn for each mountain." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q4: No group member has any private initial information in Stage 1 on the location of gems." :

- Correct
- Incorrect

Read Instructions

Next

Next

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q5: The position of gems will not be reset between the two stages of a round." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q6: All group members select the mountains simultaneously." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q7: If another group member chose a mountain before you, you cannot choose it again." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q8: At the end of the experiment, one round will be randomly selected for payment." :

- Correct
- Incorrect

Read Instructions

Next

Next



You have successfully finished the quiz.

The experiment follows: When you are ready please click "Next" to start the experiment.

Next

Start of Block 1

This is Block 1 of 4 and each Block consists of 5 Rounds.

You have been randomly assigned to a **new** group of 5 participants.

Next

Start of Round 1

You are now in Round 1 of 5 and each Round consists of 2 Stages.

The computer redrew the gems for each mountain.

No participant has any initial information on the location of gems.

In this round, for each mountain, there could be:

😥 : a topaz worth **\$1.00** with **60%** chance

- 🏮 : a ruby worth **\$6.00** with **20%** chance
- 💎 : a diamond worth **\$11.00** with **20%** chance

Stage 1

This is Block 1 of 4: You are in Round 1 of 5.



In this round, for each mountain, there could be:

🚯 : a topaz worth **\$1.00** with **60%** chance

🌔 : a ruby worth \$6.00 with 20% chance

💎 : a diamond worth **\$11.00** with **20%** chance

The location of gems is random and no participant has any initial information where each gem is hidden.

It is NOT your turn yet, please wait.



Stage 1





In this round, for each mountain, there could be:

🚯 : a topaz worth **\$1.00** with **60%** chance

- 📋 : a ruby worth \$6.00 with 20% chance
- 💎 : a diamond worth **\$11.00** with **20%** chance

The location of gems is random and no participant has any initial information where each gem is hidden.

Now it is YOUR TURN, please select a mountain.



Stage 1: Earnings

You selected Mountain 2 and found a 🕦 . Thus, you earned \$6.00 from your choice.

All discovered gems in addition to the mapped mountain and their locations are highlighted below. These will also be displayed in Stage 2 when you make your next choice.

Click "Next" to proceed to the next stage.







Stage 2



In this round, for each mountain, there could be:

- 🚯 : a topaz worth **\$1.00** with **60%** chance
- 🚺 : a ruby worth \$6.00 with 20% chance
- 💎 : a diamond worth **\$11.00** with **20%** chance

Now it is YOUR TURN, please select a mountain.



Read Instructions

Confirm your mountain choice

Stage 2: Earnings

You selected Mountain 2 and found a 🏮 . Thus, you earned \$6.00 from your choice.

Your total earnings from both stages in this round are **\$6.00 + \$6.00 = \$12.00**

All discovered gems and their locations in both Stages are highlighted below.

Please click "Next" to proceed to the next round.





D.3 Data Condition with Intermediate Rivalry

Instructions

General Information

Welcome. This is an experiment in the economics of decision-making. If you pay close attention to these instructions, you can earn a significant amount of money paid to you at the end of the experiment via bank transfer.

To participate in this online experiment, you will need to use your notebook or personal computer (mobile phones are not supported). If you are using a device that is not supported, please copy the experiment link, open a notebook or pc and paste the link into the address bar.

Your computer screen will display useful information. Remember that the information on your computer screen is PRIVATE. To ensure best results for yourself and accurate data for the experimenters, please DO NOT COMMUNICATE or interact with other people on other media at any point during the experiment. If you have any questions, or need assistance of any kind, please call +43-678-780-7284 or use Zoom anytime during the experiment and one of the experimenters will help you privately. We expect the entire experiment to take up to 60 minutes to complete.

Following these instructions, you will be asked to make some choices. There are no correct choices. Your choices depend on your preferences and beliefs, so different participants will usually make different choices. You will be paid according to your choices, so read these instructions carefully and think before you decide.

The Basic Idea

There are 5 mountains and each of them hides one type of gem, which can only be found by exploring the mountain.



You choose which mountains to explore and the value of the gems you find are your earnings in dollars.

How the Gems Are Distributed

You will not know where the gems are hidden from the outset. At the beginning of every round, a gem for each mountain will be randomly drawn, so any gem could be hidden in any mountain.

For each mountain, there is a:

· 60% chance it contains a topaz

•	20%	chance	it	contains	а	ruby
---	-----	--------	----	----------	---	------

- 20% chance it contains a diamond

These chances are the same for all five mountains. Hence, there is some chance that there could be more than one diamond, but there is also some chance that there could be no diamond. Further, even if, for example, the first two mountains happen to contain a diamond, the chance that the third mountain contains a diamond is still 20%.

The Map

At the beginning of each round, **one** mountain will be randomly selected to be mapped and its gem will be uncovered to all participants. Each participant will be able to see the same gem contained by the mountain. The mountain chosen for mapping is random and changes in each round. Besides the map, no participant has any other initial information in Stage 1 on the location of gems.

How Participants Choose Mountains

In each round, participants choose which mountain to explore. The choice does not happen simultaneously, but participants choose sequentially, one after the other, according to a random order that changes every round. You can choose to explore any mountain you wish or select the mapped mountain. <u>If you choose the same mountain already chosen by three other participants, you will **not** receive the gem's value uncovered. Instead you will receive a value of zero. Similarly, if someone else chooses the same mountain that you previously chose and you were among the first three to do so, you will receive the full gem's value (and the other participants) that chose it will not receive the gem's value uncovered if they were not among the first three participants to select that mountain).</u>

To repeat, all participant have the same information in Stage 1 on the location of one of the gems.

Each Round Has 2 Stages

A round consists of 2 stages. At the beginning of a new round, gems are redrawn for each of the five mountains. The position of gems will **not** be reset between the two stages in a round.

In Stage 1, all participants sequentially choose one mountain to explore. Before choosing a mountain, you will see which mountains have been selected by the other participants in your group who chose before you, and how many participants have selected each mountain. You can choose the same mountain or a different mountain.

At the end of Stage 1, the gems hidden in each mountain selected by all participants in Stage 1 are revealed, and you earn the value of the gem hidden in the mountain you chose.

In Stage 2, you can again choose any of the same five mountains; that is you can either choose the same mountain of Stage 1 or switch to another one. The position of gems remains the same as in Stage 1, but this time you will also see the gems located in the mountains revealed in Stage 1 in addition to the mapped mountain.

At the end of Stage 2, the gems hidden in each mountain selected by all participants in Stage 2 are revealed, and you earn the value of the gem hidden in the mountain you chose in Stage 2. You will also see your total earnings for the round which equals the sum of the value of the gem you found in Stage 1 and the value of the gem you found in Stage 2.

Game Structure

The game is divided into 4 blocks, each made of 5 rounds, with each round encompassing the two stages described above. At the beginning of each block, you will be randomly assigned to a new group of 5 participants, with whom you will play for the entire block (5 rounds in total). After the block is complete, you will be randomly assigned to a new group of 5 participants. Again, you will play for 5 rounds. This procedure will be repeated 4 times in total.

You will be reminded of this information in the top-right corner of your screen, as in the example below:





Payment

Fixed Participation Fee: You will earn a participation fee of \$5.00 for participating in this experiment.

Additional Payment and Random Round: One round will be randomly selected for payment at the end of the experiment. You will be paid and your earnings in that round as described above. Any of the 20 rounds (4 blocks with 5 rounds each) could be the one selected, so you should treat each round as if it will be the one determining your payment. This protocol of determining payments suggests that you should choose in each round as if it is the only round that determines your payment as the dollar value of the gems you select will directly translate into your earnings.

Survey and Payment: In addition to the participation fee and the payment for the randomly selected round, you will perform a small task at the very end of the experiment, and your earnings from this task will be paid to you.

You will be informed of your payment and the round chosen for payment at the end of the experiment. The \$ you have earned will be converted into Euros at an exchange rate of \$ 1 = € 0.67. Finally, after completing the experiment you will be paid electronically via bank transfer.

Frequently Asked Questions

Q1: Is this some kind of psychology experiment with an agenda you haven't told us?

A: No, it is an economics experiment. If we do anything deceptive or don't pay you as described, then you can complain to the University of Toronto Research Ethics Board and we will be in serious trouble. These instructions are meant to clarify how you earn money and our interest is in seeing how people make decisions.

Q2: Is there a "correct" or "wrong" choice of action? Is this kind of a test?

A: No, your optimal choice depends on your preferences and beliefs and different people may hold different beliefs.

Next

This button will be activated after 281 seconds. Please take your time to read through the instructions.



You have successfully finished reading the instructions.

The quiz, consisting of 8 questions in total, follows.

Next

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q1: In each round, you will select two mountains (one in Stage 1, and one in Stage 2) and collect the gem that they hide. You can choose the same mountain in both stages, or change after Stage 1." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q2: If more than one player selects the same mountain, all players will always collect the full value of the gem." :

- Correct
- \bigcirc Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q3: At the beginning of a new round, the gems are redrawn for each mountain." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q4: No group member has any private initial information in Stage 1 on the location of gems." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q5: The position of gems will not be reset between the two stages of a round." :

- Correct
- Incorrect

Read Instructions

Next

Next

Next

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q6: All group members select the mountains simultaneously." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q7: If another group member chose a mountain before you, you cannot choose it again." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q8: At the end of the experiment, one round will be randomly selected for payment." :

- Correct
- Incorrect

Read Instructions

Next

Next



You have successfully finished the quiz.

The experiment follows: When you are ready please click "Next" to start the experiment.

Next

Start of Block 1

This is Block 1 of 4 and each Block consists of 5 Rounds.

You have been randomly assigned to a **new** group of 5 participants.

Next

Start of Round 1

You are now in Round 1 of 5 and each Round consists of 2 Stages.

The computer redrew the gems for each mountain.

No participant has any initial information on the location of gems.

In this round, for each mountain, there could be:

😥 : a topaz worth **\$1.00** with **60%** chance

- 🏮 : a ruby worth **\$6.00** with **20%** chance
- 💎 : a diamond worth **\$11.00** with **20%** chance

Stage 1

This is Block 1 of 4: You are in Round 1 of 5.



In this round, for each mountain, there could be:

😥 : a topaz worth **\$1.00** with **60%** chance

- 🕦 : a ruby worth \$6.00 with 20% chance
- 💎 : a diamond worth **\$11.00** with **20%** chance

The location of gems is random and no participant has any initial information where each gem is hidden.

It is NOT your turn yet, please wait.



Stage 1



Stage 2

In this round, for each mountain, there could be:

😥 : a topaz worth **\$1.00** with **60%** chance

前 : a ruby worth \$6.00 with 20% chance

: a diamond worth \$11.00 with 20% chance

The location of gems is random and no participant has any initial information where each gem is hidden.

Now it is YOUR TURN, please select a mountain.



Stage 1: Earnings

You selected Mountain 3 and found a 💎. Thus, you earned **\$0.00** from your choice because you were **not among the first three players** to select this mountain.

All discovered gems in addition to the mapped mountain and their locations are highlighted below. These will also be displayed in Stage 2 when you make your next choice.

Click "Next" to proceed to the next stage.



Stage 2: Earnings

You selected Mountain 2 and found a 🌔. Thus, you earned \$6.00 from your choice.

Your total earnings from both stages in this round are **\$0.00 + \$6.00 = \$6.00**

All discovered gems and their locations in both Stages are highlighted below.

Please click "Next" to proceed to the next round.





D.4 No-Data Condition with Extreme Rivalry

Instructions

General Information

Welcome. This is an experiment in the economics of decision-making. If you pay close attention to these instructions, you can earn a significant amount of money paid to you at the end of the experiment via bank transfer.

To participate in this online experiment, you will need to use your notebook or personal computer (mobile phones are not supported). If you are using a device that is not supported, please copy the experiment link, open a notebook or pc and paste the link into the address bar.

Your computer screen will display useful information. Remember that the information on your computer screen is PRIVATE. To ensure best results for yourself and accurate data for the experimenters, please DO NOT COMMUNICATE or interact with other people on other media at any point during the experiment. If you have any questions, or need assistance of any kind, please call +43-678-780-7284 or use Zoom anytime during the experiment and one of the experimenters will help you privately. We expect the entire experiment to take up to 60 minutes to complete.

Following these instructions, you will be asked to make some choices. There are no correct choices. Your choices depend on your preferences and beliefs, so different participants will usually make different choices. You will be paid according to your choices, so read these instructions carefully and think before you decide.

The Basic Idea

There are 5 mountains and each of them hides one type of gem, which can only be found by exploring the mountain.



You choose which mountains to explore and the value of the gems you find are your earnings in dollars.

How the Gems Are Distributed

You will not know where the gems are hidden from the outset. At the beginning of every round, a gem for each mountain will be randomly drawn, so any gem could be hidden in any mountain.

For each mountain, there is a:

· 60% chance it contains a topaz



• 20% chance it contains a diamond

These chances are the same for all five mountains. Hence, there is some chance that there could be more than one diamond, but there is also some chance that there could be no diamond. Further, even if, for example, the first two mountains happen to contain a diamond, the chance that the third mountain contains a diamond is still 20%.

How Participants Choose Mountains

In each round, participants choose which mountain to explore. The choice does not happen simultaneously, but participants choose sequentially, one after the other, according to a random order that changes every round. You can choose to explore any mountain you wish or select the mapped mountain. <u>If you choose the same mountain already chosen by other</u> <u>participants, you will **not** receive the gem's value uncovered. Instead you will receive a value of zero. Similarly, if someone else chooses the same mountain that you previously chose and you were the first to do so, you will receive the full gem's value (and the other participant(s) that chose it will not receive the gem's value uncovered). To repeat, no participant has any initial information in Stage 1 on the location of gems.</u>

Each Round Has 2 Stages

A round consists of 2 stages. At the beginning of a new round, gems are redrawn for each of the five mountains. The position of gems will **not** be reset between the two stages in a round.

In Stage 1, all participants sequentially choose one mountain to explore. Before choosing a mountain, you will see which mountains have been selected by the other participants in your group who chose before you, and how many participants have selected each mountain. You can choose the same mountain or a different mountain.

At the end of Stage 1, the gems hidden in each mountain selected by all participants in Stage 1 are revealed, and you earn the value of the gem hidden in the mountain you chose.

In Stage 2, you can again choose any of the same five mountains; that is you can either choose the same mountain of Stage 1 or switch to another one. The position of gems remains the same as in Stage 1, but this time you will also see the gems located in the mountains revealed in Stage 1 in addition to the mapped mountain.

At the end of Stage 2, the gems hidden in each mountain selected by all participants in Stage 2 are revealed, and you earn the value of the gem hidden in the mountain you chose in Stage 2. You will also see your total earnings for the round which equals the sum of the value of the gem you found in Stage 1 and the value of the gem you found in Stage 2.



Game Structure

The game is divided into 4 blocks, each made of 5 rounds, with each round encompassing the two stages described above. At the beginning of each block, you will be randomly assigned to a new group of 5 participants, with whom you will play for the entire block (5 rounds in total). After the block is complete, you will be randomly assigned to a new group of 5 participants. Again, you will play for 5 rounds. This procedure will be repeated 4 times in total.

You will be reminded of this information in the top-right corner of your screen, as in the example below:





Payment

Fixed Participation Fee: You will earn a participation fee of \$5.00 for participating in this experiment.

Additional Payment and Random Round: One round will be randomly selected for payment at the end of the experiment. You will be paid and your earnings in that round as described above. Any of the 20 rounds (4 blocks with 5 rounds each) could be the one selected, so you should treat each round as if it will be the one determining your payment. This protocol of determining payments suggests that you should choose in each round as if it is the only round that determines your payment as the dollar value of the gems you select will directly translate into your earnings.

Survey and Payment: In addition to the participation fee and the payment for the randomly selected round, you will perform a small task at the very end of the experiment, and your earnings from this task will be paid to you.

You will be informed of your payment and the round chosen for payment at the end of the experiment. The \$ you have earned will be converted into Euros at an exchange rate of \$ 1 = € 0.67. Finally, after completing the experiment you will be paid electronically via bank transfer.

Frequently Asked Questions

Q1: Is this some kind of psychology experiment with an agenda you haven't told us?

A: No, it is an economics experiment. If we do anything deceptive or don't pay you as described, then you can complain to the University of Toronto Research Ethics Board and we will be in serious trouble. These instructions are meant to clarify how you earn money and our interest is in seeing how people make decisions.

Q2: Is there a "correct" or "wrong" choice of action? Is this kind of a test?A: No, your optimal choice depends on your preferences and beliefs and different people may hold different beliefs.

Next

This button will be activated after 281 seconds. Please take your time to read through the instructions.



You have successfully finished reading the instructions.

The quiz, consisting of 8 questions in total, follows.

Next

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q1: In each round, you will select two mountains (one in Stage 1, and one in Stage 2) and collect the gem that they hide. You can choose the same mountain in both stages, or change after Stage 1." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q2: If more than one player selects the same mountain, all players will always collect the full value of the gem." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q3: At the beginning of a new round, the gems are redrawn for each mountain." :

- Correct
- Incorrect

Read Instructions

Next



Quiz Time!

Please mark the following statements as correct/incorrect:

"Q4: No group member has any private initial information in Stage 1 on the location of gems." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q5: The position of gems will not be reset between the two stages of a round." :

Correct

Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q6: All group members select the mountains simultaneously." :

- Correct
- Incorrect

Read Instructions

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q7: If another group member chose a mountain before you, you cannot choose it again." :

- Correct
- Incorrect

Read Instructions

Next

Next

Next

Quiz Time!

Please mark the following statements as correct/incorrect:

"Q8: At the end of the experiment, one round will be randomly selected for payment." :

- Correct
- \bigcirc Incorrect



Next

You have successfully finished the quiz.

The experiment follows: When you are ready please click "Next" to start the experiment.

Next

Start of Block 1

This is Block 1 of 4 and each Block consists of 5 Rounds.

You have been randomly assigned to a **new** group of 5 participants.



Start of Round 1

You are now in Round 1 of 5 and each Round consists of 2 Stages.

The computer redrew the gems for each mountain.

No participant has any initial information on the location of gems.

In this round, for each mountain, there could be:

a topaz worth \$1.00 with 60% chance
a ruby worth \$6.00 with 20% chance

💎 : a diamond worth **\$11.00** with **20%** chance

Next

Stage 1

This is Block 1 of 4: You are in Round 1 of 5.



In this round, for each mountain, there could be:

😥 : a topaz worth **\$1.00** with **60%** chance

前 : a ruby worth \$6.00 with 20% chance

💎 : a diamond worth **\$11.00** with **20%** chance

The location of gems is random and no participant has any initial information where each gem is hidden.

It is NOT your turn yet, please wait.







1 player selected this mountain Mountain 4





?

Read Instructions

Stage 1

This is Block 1 of 4: You are in Round 1 of 5.



In this round, for each mountain, there could be:

- 👸 : a topaz worth **\$1.00** with **60%** chance
- 🕦 : a ruby worth \$6.00 with 20% chance
- 💎 : a diamond worth **\$11.00** with **20%** chance

The location of gems is random and no participant has any initial information where each gem is hidden.

Now it is YOUR TURN, please select a mountain.



Stage 1: Earnings

You selected Mountain 3 and found a Intersection with the select this mountain.

All discovered gems in addition to the mapped mountain and their locations are highlighted below. These will also be displayed in Stage 2 when you make your next choice.

Click "Next" to proceed to the next stage.







Stage 2

This is Block 1 of 4: You are in Round 1 of 5.



In this round, for each mountain, there could be:

😥 : a topaz worth **\$1.00** with **60%** chance

- 前 : a ruby worth \$6.00 with 20% chance
- 💎 : a diamond worth **\$11.00** with **20%** chance

Now it is YOUR TURN, please select a mountain.







1 player selected this mountain Mountain 4

\$1.00

1 player selected this mountain



Read Instructions

Confirm your mountain choice

Stage 2: Earnings

You selected Mountain 1 and found a 🙆 . Thus, you earned \$1.00 from your choice.

Your total earnings from both stages in this round are \$0.00 + \$1.00 = \$1.00

All discovered gems and their locations in both Stages are highlighted below.

Please click "Next" to proceed to the next round.





D.5 Questionnaire and Risk-Preferences Elicitation Task



You have successfully finished the main part of the experiment.

A brief questionnaire together with a short task follows: When you are ready please click "Next".

Next

Please answer the following questions

Your answers will be kept confidential and will not affect your earnings for today's experiment.

Please state your age:

Please state your gender:

----- *****

Please state your student type:

----- **\$**

Please state your country of origin:

Please state your degree and field of study:

Please briefly explain, in your own words, the rules of today's experiment:



Please briefly describe how you reached your decisions in this experiment:

Please briefly describe how, in your opinion, other participants reached their decisions in this experiment:



Instructions

Thank you for your participation so far. In the last task of the experiment, you will earn an additional reward based on a set of 10 choice problems.

How does it work?

The Choice: You will be asked to choose between two options, "Option A" and "Option B" where:

- "Option A" always pays \$4.00 with probability p and \$3.20 otherwise.
- "Option B" always pays \$7.70 with probability p and \$0.20 otherwise.

Repeated Choices:

- You will be asked to make a choice between "Option A" and "Option B" not once, but **ten** times where p will increase from 10% to 100%, 10% at a time.
- For example, the first choice will have p=10% and you will choose whether you prefer "Option A" (\$4.00 with a 10% chance or \$3.20 otherwise) or "Option B" (\$7.70 with a 10% chance or \$0.20 otherwise).
- Each successive choice will increase p by 10 percentage points until the last choice where "Option A" will pay \$4.00 with certainty, and "Option B" will pay \$7.70 with certainty.

Note: Once you switch from choosing "Option A" to "Option B", it makes sense that you will continue to choose "Option B" in all consecutive choice problems. For example, if you prefer "Option B" when p=80%, then it makes sense to prefer "Option B" when p=90% and when p=100%, since "Option B" is even more attractive in these choice problems.

Therefore, we have designed the interface so that you must either (a) **always** choose "Option A" or "Option B" for all 10 choice problems or (b) if you **switch** to "Option B" for a given probability p, then you must choose "Option B" for all the following choices as well.

You can adjust your choices as many times as you wish. When you are ready to submit your choices, you can click on the "Next" button at the bottom of the page.

Payment

The computer will randomly select one of the 10 choice problems and pay you according to your choice in that problem where the computer will decide the outcome based on the value of p.

Please Choose Between "Option A" and "Option B" on Every Line

	Option A		Option B			
	\$4.00 with a chance of 10%, \$3.20 otherwise	0 0	\$7.70 with a chance of 10%, \$0.20 otherwise			
	\$4.00 with a chance of 20%, \$3.20 otherwise	0 0	\$7.70 with a chance of 20%, \$0.20 otherwise	6		
	\$4.00 with a chance of 30%, \$3.20 otherwise	0 0	\$7.70 with a chance of 30%, \$0.20 otherwise			
	\$4.00 with a chance of 40%, \$3.20 otherwise	0 0	\$7.70 with a chance of 40%, \$0.20 otherwise			
	\$4.00 with a chance of 50%, \$3.20 otherwise	0 0	\$7.70 with a chance of 50%, \$0.20 otherwise			
	\$4.00 with a chance of 60%, \$3.20 otherwise	0 0	\$7.70 with a chance of 60%, \$0.20 otherwise			
	\$4.00 with a chance of 70%, \$3.20 otherwise	0 0	\$7.70 with a chance of 70%, \$0.20 otherwise			
	\$4.00 with a chance of 80%, \$3.20 otherwise	0 0	\$7.70 with a chance of 80%, \$0.20 otherwise			
	\$4.00 with a chance of 90%, \$3.20 otherwise	0 0	\$7.70 with a chance of 90%, \$0.20 otherwise			
	\$4.00 with a chance of 100%, \$3.20 otherwise	0 0	\$7.70 with a chance of 100%, \$0.20 otherwise			
Read Instructions						
D.6 Payment Information

Thank you for participating in this experiment!

Your payoffs for this experiment are as follows:

Main Experiment:	 Round 1 of Block 1 was randomly selected for payment. In Stage 1, you found a and received \$11.00 and in Stage 2, you found a and received \$11.00 Thus, your total payoff is \$11.00 + \$11.00 = \$22.00 		
Last Task of Experiment:	• The following choice problem was randomly selected:		
	Option A		Option B
	\$4.00 with a probability of 10%, \$3.20 otherwise	• •	\$7.70 with a probability of 10%, \$0.20 otherwise
	 As indicated above, you chose Option A. The computer drew a random number to determine your payoff according to the chances specified. Your payoff is \$3.20 		
Participation Fee:	• You earned a fee of \$5.00		

In total, you earned **\$22.00 + \$3.20 + \$5.00 = \$30.20** from your choices.

In euros, this corresponds to: **€20.13**

Please click Next to upload your bank details.

Next