ECONOMIC EXPERIMENTS IN VIRTUAL WORLDS: FRAMING, RECIPROCITY & TRUST

A thesis
submitted by

Stephen A. Atlas

In partial fulfillment of the requirements
for the degree of

Master of Arts

In

Economics

TUFTS UNIVERSITY

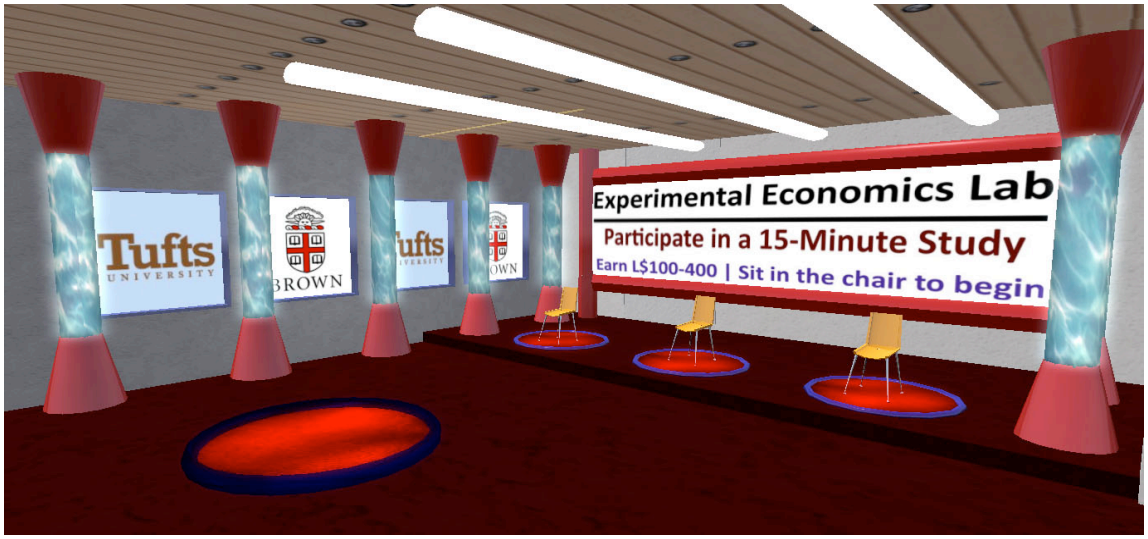
May 2008

Advisers:

Enrico Spolaore, PhD
Tufts University

Louis Putterman, PhD
Brown University

David Garman, PhD
Tufts University



ECONOMIC EXPERIMENTS IN VIRTUAL WORLDS: FRAMING, RECIPROCITY & TRUST

STEPHEN A. ATLAS
TUFTS UNVIERSITY

Abstract

This research investigates the viability of gaining insights about the real world through conducting economic experiments in virtual worlds. The core of this project was the replication of a classic economic experiment on trust and reciprocity and application of treatments to identify the impact of framing on these behaviors. To substantially reduce the cost of running the experiment, increase the sample size, and investigate a different pool of subjects than typical lab-based experiments studying college students, this study applied an innovative

method of recruiting and motivating participants that uses the virtual game currency, Second Life Lindens (L$). The experiment itself consists of two phases: first, an established trust experiment methodology, conducted in the real world, was replicated within the virtual world of Second Life to quantify differences in the features of this subject pool. Then framing effects were determined by infusing the subjects' instructions with language intended to elicit cooperative or selfish behavior. Ultimately over 1,200 observations were collected, and the research provides evidence substantiating oft-cited concerns about generalizing findings from college student data to the entire population. This paper concludes with a more general discussion about the key methodological considerations in and potential of running experiments in virtual worlds.

Acknowledgements

This project would not have been possible without the support of a number of people. Many thanks to the thesis committee, particularly for their flexibility and enthusiasm for exploring this new frontier of economic experimentation. This project was funded through a generous grant from Tufts' Graduate School Research Award program. There were also a number of individuals who helped me get my bearings in Second Life, including Sarah Robbins, Mark Bell, Robert Bloomfield, and the residents of Etopia Island. Additionally, University of Paris computer science doctoral candidate Stefan Bornhofen wrote outstanding code underlying the data collection apparatus. Thanks also to the continued support and encouragement of my family and fiancée, Maya Vadiveloo, for believing in me throughout this process.

Chapter 1: Worlds for Study

Virtual worlds provide a wealth of opportunities for economists and other researchers who seek to generalize insight into individual behavior and decision-making. Enabled by the rise of the Internet, widespread adoption of powerful personal computers, and continuously increasing bandwidth speeds, over the past decade a number of virtual worlds have emerged. Today, the most a number of immersive online environments exist, ranging from middle-earth fantasy (World of Warcraft, Everquest), intergalactic combat (Entropia, Eve Online) or postmodern daily life (The Sims, Second Life, There) (Bloomfield, 2007). Every day, millions of individuals make very real decisions about their virtual selves, known as avatars, and apply judgment about whether to cooperate with the individuals they encounter.

As virtual worlds become increasingly mainstream, researchers have begun to document the economic forces at work in virtual worlds as the field of "metanomics," the study of the economics of the metaverse. Accounting Professor Robert Bloomfield of Cornell University Johnson School of Management, divides the field of metanomics in to three areas. Immersionist research refers to research within virtual worlds from the residents' perspective. Augmentationist research explores how real-world enterprises and individuals use virtual worlds to achieve their strategic goals. Finally, experimentalist research, which encompasses this thesis project, includes carefully controlled tests conducted within virtual worlds ranging from subtle language adjustments

that induce framing effects to the tantalizing prospect of exogeneously adjusting macroeconomic variables at reduced risk (Bloomfield, 2008).

The classic example of immersionist metanomics research is Indiana Telecommunications Professor Edward Castronova's 1999 article "Virtual Worlds: A First-Hand Account of Market and Society on the Cyberian Frontier." A Wisconsin-trained economist, Castronova's article applies the exchange of virtual goods for real money to calculate the value of Norrath's economy despite its existence only as bytes on a server farm in San Diego, CA. Consequently, he was able to determine that in Norrath, the nominal hourly wage is $3.43/hr, the unit of currency is valued at $0.0107 and the GNP per capita lies somewhere between that of Russia and Bulgaria. As the first example of rigorous analysis of virtual economies, this article paved the way for economic analysis of virtual worlds. As virtual economies are storehouses of billions of dollars of value, analyzing the economic forces within these worlds is itself a relevant subject in its own right. However, aside from analysis of real-money-transfer (RMT), economists have been heavily underrepresented in contributing research about this space.

In contrast with immersionist research, augmentationist metanomics is more concerned with the applications virtual worlds can provide to real-world organizations individuals. The best example of this type of research is Robert Bloomfield's weekly virtual talk show, Metanomics, which explores the business and policy issues of virtual worlds. This series explores how individuals and organizations who have leveraged virtual worlds to meet their personal and

functional objectives. Initially, businesses developed what they hoped would become virtual profit centers and cheap advertising opportunities. However, as virtual worlds were (and in many ways, still are) in their nascent stages of development, the returns to such advertising initiatives were often ambiguous and virtual sales were often negligible in real terms. More recently, businesses have turned to virtual worlds to provide a medium to connect teams and to the public. Virtual office spaces over secure networks provide opportunities to collaborate with colleagues, while virtual showrooms enable companies to easily receive feedback from the public at early stages of product development. Similarly, academic institutions and educators have been exploring how to incorporate this new technology to help meet their educational and functional objectives. Arguably all individuals active in virtual worlds, even the most diehard role-players, have an augmentationist aspect of their decision. Indeed, though individuals flock to virtual worlds for a variety of reasons including security to express themselves, to reinvent themselves, or to roleplay, there is always a underlying aspect of virtual space that provides a real augmentationist purpose.

Bloomfield's final category of metanomics is experimentalist, which includes measuring the effects of the deliberate exogenous changes on agents' decisions. However, as experimentation is really a method rather than an approach, one could further divide such research as either immersionist experimentation or augmentationist experimentation. Immersionist experimental research is descriptive of behaviors within a specific virtual world and augmentationist experimental research seeks to leverage the benefits of virtual

world experimentation to provide insight about real-world behavior. Though immersionist experimental research is valuable in its own right to explore how individuals behave in virtual settings, greater scientific significance results when virtual worlds can contribute to the dialogue about real world behavior.

While several notable academics such as Robert Bloomfield and Ted Castronova have called for experimentation through virtual worlds, few true experiments have been conducted. One notable exception was a 2007 article by Chesney, Chuah and Hoffman, who chronicle their experience replicating classic economic experiments including the ultimatum, dictator, public goods, minimum effort, and guessing game. Using sample sizes between 4-30 pairs, their experiments are more concerned with replicating classic experiments rather than leveraging the technology to drastically increase the amount of data collected. Their results are generally consistent with findings from physical labs, attributing discrepancies to demographic differences of virtual and undergraduate populations.

Meanwhile, David Abrams and John List of the University of Chicago are reportedly currently conducting research quantifying the differences between experiments conducted virtually and in labs. According to the July 6, 2007 Chronicle of Higher Education,

> List, who often runs field experiments in microeconomics, says that if he sees people in some tests behaving the same in real-world and virtual environments, he will consider doing more studies in virtual worlds, because they are more cost-effective. 'For certain types of games, like bidding and auctions, I think that will generalize quite easily across the virtual world to the lab,' he says. Other studies, though, may work only in face-to-face laboratories. "Behaviors are influenced by whether people

> can link your identity to your behavior," says Mr. List. 'In the virtual world, I think, you're virtually free of these reputational concerns, so you might get people acting in a more self-interested way.'"

After scheduling a phone conversation with Abrams to understand their plans, their work involves 1) experiments such as testing the same set of subjects both in SL and in a lab to see their choices differ, 2) testing changes in the physical environment, such as color of walls, or wording of instructions influences agents' choices and 3) manipulating the subject pool within the virtual worlds to determine how that affects the outcome. Abrams reports that he is interested in creating contracts that offer proper incentives for completing transactions in virtual worlds in the absence of an effective authority to mediate disputes (Personal Interview, 2008).

While Bloomfield provides a useful framework under which to interpret the recent economics research, some research appears to fall outside the boundaries of the three categories. Simple demographic studies of virtual world residents or more subjective psychographics used by marketing researchers do not fit in Bloomfield's paradigm. Additionally, virtual worlds could provide insight into a host of social science research questions through econometric analysis of the terabytes (trillion bytes) of related data automatically collected by their servers. However, all of these research areas appear to fall outside the Metanomics framework unless we expand experimentalist metanomics to include non-experimental metanomics research studying real-world issues. Perhaps a concept such as "inductive metanomics" would suffice to refer to the application

of data obtained from virtual worlds to provide insight about the real world. In this context, experimental metanomics is but one method to provide such intuition.

In the process of conducting inductive metanomic research in the form of virtual experimentation, It is worth emphasizing the abundance of unanswered questions limiting the insights carried from the virtual world to the real world. While internal validity issues are paramount in developing meaningful research methods for both immersive and augmentationist, inductive metanomics, ecological validity is particularly important in assessing whether observations from virtual worlds are appropriate to provide any insight to reality.

Chapter 2: Constructing a Virtual Lab

Developing a virtual lab for conducting economic experiments presented a number of practical considerations not typically a factor in an experiment's design, including the selection of an appropriate virtual world, selecting a site, and developing the experimental apparatus. Once these issues were addressed, the study could turn to the more common methodological specifications such as recruiting participants, compensating them, and executing the substance of the experiment.

Virtual World Selection

A variety of available virtual worlds presented an important strategic decision in conducting this study. Virtual worlds vary widely by demographics, culture, structure, economic integration. While demographic and cultural considerations were not an important factor in the selection of the world, structure and economic integration were crucial factors in meeting IRB ethics requirements as well as developing lab infrastructure fully integrated with the virtual world.

"Structured" worlds, such as Blizzard Entertainment's World of Warcraft are driven by developer-created content. Conducting an automated experiment in a structured world required either the development company to create content specifically for the purpose of the experiment or for interactions to occur strictly

outside the virtual world in a website. As the World of Warcraft economy is officially closed, subjects' payments would be delivered through informal currency vendors operating against the wishes of Blizzard Entertainment. The IRB required that the research plan could not violate the Terms of Service (ToS) of the virtual world provider, which in effect prevented the use of World of Warcraft because it lacked a sanctioned real-money trading (RMT) mechanism. Consequently, if the experiment had been implemented in this space, it could have subjected participants to significant unnecessary risks such as account termination. Thus, World of Warcraft presented a unsatisfactory environment for conducting experiments through a virtual lab.

Alternatively, the "unstructured" world of Linden Labs' Second Life is defined by the content created by the users themselves. This allowed for the creation of a virtual lab with all subject-facing content existing inside the virtual world without the necessity of involving the game developers themselves. Additionally, the ethical concern present in WoW was eliminated by the fact that the Second Life economy was "open" through sanctioned RMT transfers of the virtual currency, Lindens (L$). In fact, the scripting language used in Second Life allowed for the automatic payment of subjects. The combination of these factors presented a much more promising environment for developing a virtual lab infrastructure, so Second Life was selected as the site of the Tufts University Virtual Experimental Economics Lab.

Selecting and Constructing a Site

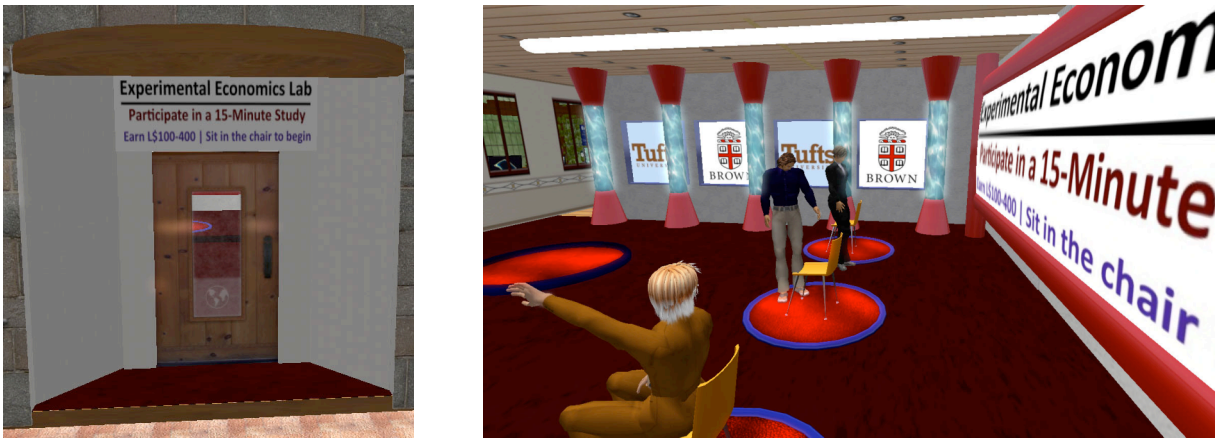
The diversity of environments available in Second Life provides an abundance of choices for the selection of a virtual lab site. Spanning over 65,000 acres and constantly expanding, the environment of Second Life is limited only by the imagination of its residents, including private beachfront residences, shopping malls of an urban metropolis, zombie-infested streets of a twisted future, or the hallowed lecture halls of academia.[i]

The ideal environment of a lab would be a dedicated island in which the terrain is built from the ground up and the entire space can be defined according to the needs of the researcher. However, such a construction would not be possible without a larger budget: after a 50% discount given to real-world academic institutions, islands are priced at $840 for 16 acres, plus $150 monthly land fees.[ii] In the absence of access to such great resources for what amounts to an exploratory study, we are left with options renting space within the established islands of Second Life.

For the purposes of conducting exploratory research in a foreign space, one of the most critical elements of conducting research was the availability of a local community who could assist with the practical tasks associated with constructing the lab. Etopia Island emerged as a suitable location with an appropriately-sized office space in which to construct the lab. The owner of Etopia Island, Williamthewise Goodman, describes Etopia as “Second Life’s premier environmental eco-village showcasing real-life examples of sustainable

development, renewable energy, organic living and authentic community." The particular draw to Etopia was the members of the community, most notably Jojogirl Bailey, who willingly answered the experimenter's "noob" questions about navigating Second Life's notoriously painful user interface. Additionally, Etopia was home to a labor force of talented "builders" such as Prim Chemistry, who helped convert the office space to include the décor of a virtual lab, complete with fluorescent lighting. Meanwhile, for the island's managers, the lab presented an important revenue stream and a way of drawing traffic to their community which they continually aspire to grow.

For the purposes of an exploratory study, the objective was developing a functioning lab. It is certainly not the case that Etopia Island was the only available option. One could rent a kiosk at a virtual shopping mall, which comes with a greater loss of control over the stimuli exposed to subjects. Additionally, researchers could rent a variety of other spaces according to their needs; most would probably find a more isolated location to be more appropriate to their needs once they become comfortable with Second Life development tools.

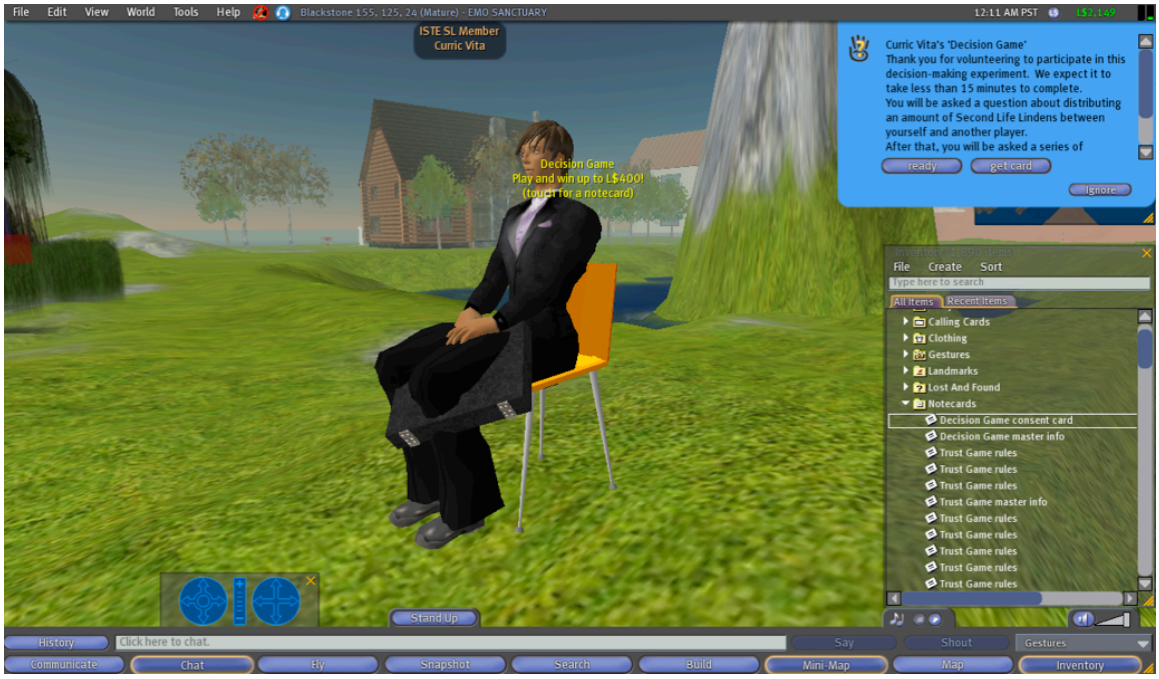


Above, images of the Tufts Experimental Economics Lab

Developing the Experimental Apparatus

In order to interact with Second Life residents and automatically collect data, a script was developed through the expertise of Stefan Bornhofen, a computer science doctoral student at the University of Paris. Using Linden Script Language (LSL), a C/Java-style language, we developed a "chair" with an internal script that automatically 1) confirms that participants are at least 18, 2) displays and captures consent to participate, 3) ensure avatars participate only once, 4) administers different versions of the instructions to comprise several treatments, 5) asks a series of follow-up demographic questions, 6) automatically compensates participants, and 7) sends the data to a centrally managed SQL database.

The ability of our apparatus to communicate with users was limited to either through text at the bottom of the screen or a blue-box at the upper right. While both options could be fashioned to collect the desired data, the blue box seemed to be a better choice. In contrast with the standard chat text used to communicate with other players, the blue box conveyed greater anonymity of players' selections and had user-friendly buttons instead of requiring a more complex syntax to "whisper" with the device. The picture below shows the typical appearance of this interaction method with users.



In the picture above, we see a screenshot of the apparatus interacting with the user through the blue box in the upper right corner.

Recruiting and Compensating Subjects

Subjects were recruited through the placement of classified advertisements in Second Life. Classified ads are browsed by users looking for new areas to explore in Second Life. The full announcement read:

> Earn L$100-400 in Academic Study
>
> Interested in earning some quick Lindens? Come to our economic research lab to participate in a 15-minute academic research study on virtual decisionmaking. If you are over 18, you can earn L$100-400. This project is sponsored by Tufts University and is part of a master's thesis - your responses will not be used for marketing purposes. Places are limited, so come now to ensure you do not miss this opportunity!
>
> (Some keywords: test subject, test subjects, experiments, free lindens, research, linden, earn, job, work, participate, win, clothes, clothing, furniture, skins, skin, pay, buy, sell, event, newbies, newbie, events, shop,

new year, cheap, deal, sale, easy job)

These keywords are included in the recruitment ad to help increase the frequency it is viewed. SL residents search the classifieds by keyword and the results are sorted by the amount of money paid by the advertiser. Bidding around L$5, 000 per week (around $18) ensured a prominent placement. Once subjects read the announcement, they followed a link that teleported their avatar directly to the lab.

Compensation occurred automatically based on decisions made by the subjects and their counterpart. Participants were exclusively paid in Second Life's virtual currency, which carried an exchange rate of roughly 280 Lindens per USD. It is worth emphasizing that the offer of L$100-400 has an explicit value between 35 cents and $1.40 for 15 minutes of time. Nevertheless, this is a relatively well-paying job in Second Life. Consequently, demand to participate in the experiment was exceptionally strong; a test of the apparatus with only one available chair exhibited such an abundance of willing participants fights broke out between avatars over who could participate in the experiment next, with subjects ejecting each other from the apparatus.

To keep up with demand, three chairs were included in the full implementation of the virtual lab. This raised a host of additional questions, most notably, how to ensure that participants do not participate more than once. Solving this problem required communication between the scripts and a central SQL database. The final version of the script confirmed that the avatar has not participated on another chair previously.

Chapter 3: Trust and Reciprocity in Virtual Worlds

Trust and Reciprocity in Physical Labs

Experimental economics poses the significant obstacle that researchers must develop their own data sets in order to explore individuals' economic behavior. One constraint is the high cost of providing test subjects with sufficient payment that not only compensates them for their time but also offers the additional incentives to drive their decisions in the experiment. This has impaired the pace at which economic theory has adapted the neoclassical economic assumptions driving many models with observations from other social science fields. This study will investigate a new method of recruiting and motivating experiment participants to leverage experimental funding in an investigation at the frontier of behavioral research in reciprocity, framing, and trust. The unique methodology of this study will take advantage of emerging aspects of the gaming world to attract and pay participants at a much lower cost per person than in conventional lab-based experiments and provide evidence about trust and reciprocity in a setting outside the typically sampled college student population.

Previous lab-based studies have explored the impact of trust and reciprocity at the expense of explicit self interest. The classic 1995 experiment by Berg, Dickhaut, and McCabe (BDM) explores trust and reciprocity in an investment setting. In this experiment, subjects in room 'A' decide how much of a $10 endowment to be tripled and sent to an anonymous counterpart in room

'B'. Then subjects in room 'B' decide how much to send back. Rational choice theory predicts a unique Nash equilibrium where subject 'A' sends $0, but in contrast 94% of subjects sent a positive amount. Subject 'A's sent $5.16, on average, and subject Bs reciprocated by sending back an average of $4.66. This evidence suggests that subjects' investment decisions were influenced by social norms.

Hoffman, McCabe and Smith (1998) note that more efficient social outcomes regularly occur in both small-group economic experiments, such as the prisoner's dilemma, and as large-group public goods experiments. They interpret this by applying the hypothesis of evolutionary psychologists such as Cosmides and Tooby that humans evolved mental algorithms for cooperation in social exchange.

Behavioral economists Fehr and Gächter (1998) adapted growing evidence that humans are motivated by both material payoffs and reciprocity into the concept of *Homo reciprocans*. Unlike the rational and self-interested *Homo economicus* of neoclassical economics, *Homo reciprocans* exhibits behavior characterized by both altruistic and negative reciprocity that enforces social norms. In recent years, studies have found that reciprocity and trust appear strengthened by communication (Ben-Ner, Putterman and Ren, forthcoming), others' cooperativeness (Fischbacher, Gachter, and Fehr, 2001), and gender (Ben-Ner, Kong, and Putterman, 2004). This investigation attempts to explore trust and reciprocity in the virtual world of Second Life and extend this literature

by more fully understand the responsiveness of trust and reciprocity to the effects of framing.

Experimental Motivation

Using traditional lab-based methods, even a short experiment is expensive to finance. At $15 per subject, a $1,000 budget can be expected to finance at most 66 participants that produce 33 observations. It is worth noting that many previous studies suffer from a small sample size due in part to cost constraints. In contrast, this study creates a 'virtual lab' infrastructure to collect data from nearly 1,500 participants. The key methodological element is the use of a novel recruitment and compensation strategy that exploits the fact that a large number of users in the virtual online game, Second Life, who are willing to participate in return for the virtual currency "Lindens" that carries little real-world value.

Acting in virtual worlds defined by player-created content, online 'gamers' explore a virtual environment embodied as an avatar, their in-game 'self'. As network technology evolves, online games can connect more users simultaneously while faster Internet transfer speeds allow more robust fantasy worlds. This provides an increasingly engrossing environment for gamers, who often work for hours to acquire money that exists outside the virtual world only as bytes in a computer server. Meanwhile, through established virtual-to-real currency exchange vendors, it is possible to buy a large amount of in-game

currency. By incentivizing participants using this inexpensive in-game currency, a research budget can be applied to answer behavioral questions by motivating gamers to react to experimental incentives much as they would in a lab setting, but at a fraction of the cost.

This research project evaluates the viability of running experiments with this subject pool by completing two phases: first this study replicates the BDM study to explore the different demographics and preferences of the subject pool. This will be accomplished by replicating the BDM trust experiment using the gaming population, followed by a series of demographic questions. Comparisons between these results and previous BDM experiments provide information with which to evaluate the efficacy of using this method compared with testing students in labs.

The second phase of this investigation applies the virtual world methodology to explore the relationship between framing, trust and reciprocity in the gaming subject pool. This will involve several variants of the instructions distributed to participants. While the content of each instructions set will be the same as BDM, the altered instructions include will be written in language designed to encourage participants to adopt a more or less cooperative strategy. One will be designed to encourage selfish decisions while the other will prompt cooperative choices. Statistical comparisons of the results with the neutral language of the first experimental phase allows an understanding of the effect of framing on trust and reciprocity.

Analysis of the resulting data allows researchers to assess the differences between decisions made in the physical labs and virtual worlds. Additionally, this provides an opportunity to assess the basis of oft-cited concerns about generalizing findings from experiments on college students.

It is worth noting that this methodological approach differs significantly from a typical controlled experiment. The automated data collection approach leaves openings for potential problems such as false demographic information by subjects and repeated enrollment with different avatars, a practice I will later refer to "experiment farming." Indeed, who exactly the subject pool is in Second Life and how it differs from the university student subject pool is an open research question. In deciding whether virtual worlds are an appropriate setting to explore new areas of economic behavior in the real world, these costs should be weighed against the benefit of having a significantly larger dataset, which allows a more robust analysis of the resulting experiment.

Running the Experiment

Once the core elements of the virtual lab structure were in place as outlined in Chapter 2, the experiment was ready to begin. Subjects learned about the experiment from the in-game classified ad, which transported them directly to the virtual experimental economics lab on Etopia Island of Second Life. From there, messages on the walls instructed participants to sit in one of the chairs to begin. The experiment began once the apparatus confirmed their

eligibility by 1) confirming that the subject's age was at least 18, 2) capturing the subject's consent to participate, and 3) comparing the name of the avatar against a database of avatars who had previously interacted with the apparatus, regardless of whether they successfully completed the experiment during the previous attempt. This prevented users from being able to restart the experiment to be able to have any control over their role or their partner.

Upon beginning the experiment, subjects were either assigned to role "A" or role "B." Each chair first collected data from five "A" subjects. After that, the chair collected five "B" subjects, matching each "B" in order to the five "A" subjects previously collected. This lag length was unknown to users, which helped prevent collusion between non-strangers. In each of the five treatment groups, both subject types were given the same instructions. The substance of all five experiments was the replication of the BDM Trust Game. As in the classic experiment, Player A was asked how much of a L$100 endowment to place in a fund that would be tripled and given to an unknown Player B. Then Player B received the tripled amount plus a L$100 show-up fee, and asked how much of the tripled amount to give to the unknown Player A.

Most treatments included follow-up questions including information about the subject's demographic and economic details in both their first- and second-lives as well as behavioral questions. Second Life questions included their avatar wealth, hourly wage in Lindens, and hours per week in Second Life. First-Life questions included their hourly wage in USD, monthly wage in USD, hours per week in paid work, educational attainment, employment/education status,

gender, race, number of siblings and younger siblings, marital status, number of children, birthplace, region of current residence. Additionally, some behavioral questions were asked including frequency of attending religious services, frequency of charitable giving, whether the subject was under the influence of drugs or alcohol, and where they found out about the experiment. We also asked the general trust question "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?" including several applications of this question to gamers including "In your opinion, are fellow gamers more deserving of reciprocal treatment, i.e. deserving of having a kindness returned, than people in the general population?" and "In your opinion, are fellow gamers more trustworthy, less trustworthy, or about the same as people in the general population?." Finally, in order to elucidate subjects' strategies, they were asked some hypothetical questions about what they would have done if they were B and were presented with various offers. The full text of these questions are included in Appendix II. These additional questions were "mandatory" in the sense that subjects received no payment from the experiment if they did not complete the additional questions. If a Player B dropped out of the experiment rather than participate in this portion, the next available person was matched to the Player A.

One treatment replicated, as closely as possible, the BDM experiment. This "BDM" treatment ended without additional questions. A second treatment, "BDMQ," used the same language as "BDM" but included the follow-up questions. The third treatment is referred to as "neutral" in that it replicated the

spirit of BDM but also included some additional instruction to help users interpret the range of possible outcomes available in the experiment. The final two treatments, “cooperative” and “competitive” inserted some additional language into the “neutral” instructions to encourage more (or respectively, less) cooperation by the subjects. The full language of the instructions provided to subjects can be found in Appendix I.

Description of Collected Data

In total, 1435 observations were collected using the automated data collection chairs. Of these, 244 subjects did not complete the survey, leaving 1,194 subjects who participated in the entire experiment. These partial data points were slightly dominated by “B” participants (which made up 55% of the dropped data points), though the 78 participants excluded because of the age restriction confounds that calculation because these users were assigned to A or B yet did not get far enough in the experiment to discover their status.

Of the 1,194 complete data points collected, 264 subjects participated in BDM replication without the followup questions, which provided 132 data points each on the decisions of player A and player B. The BDM replication with survey questions accounted for another 416 subjects, 208 of each type. Additionally, the neutral treatment applied to 100 of each player type (200 total), the competitive framing treatment was applied to 77 of each type (154 total), and the cooperative framing treatment was applied to 80 of each type (160 total).

Data Analysis and Interpretation

In contrast with neoclassic predictions, every treatment group's A subjects a nonzero number of subjects sent more than zero Lindens, as shown in the table below. In all five treatments, we observed that between 67% and 94% of A's sent non-zero values to B, with average contributions ranging from 40 to 57.

Table 1: A's Sending, by Treatment

Treatment	%A's Sent Nonzero	Avg. Sent by A	SD of Amt Sent by A
Neutral	.67	43.3	42.3
Competitive	.896	45.8	34.6
Cooperative	.938	57.0	34.6
BDM (no Q's)	.939	42.8	34.7
BDM (w Q's)	.875	39.9	35.9

Does trusting pay off for A? After merging A with the B counterpart to calculate A's final payoff, the neutral treatment group lost money by trusting while those in the cooperative treatment and the BDM (no questions) gained money by trusting. BDM (with questions) and Competitive treatment were not significantly different from the 100 Lindens participants would have received if they had sent zero to their counterpart.

Table 2: Does Trusting Pay Off for A, by Treatment

Treatment	Obs (nonzero)	Mean Payoff	SE	T-value (Ho: μ=100)	Pr(|T| > |t|)
Neut	67	83.52	8.42	-1.96	.055*

Comp	69	104.55	6.79	.67	.50
Coop	75	120.71	6.60	3.14	.0024***
BDM	124	108.23	5.03	1.64	.10*
BDMQ	182	95.75	3.63	-1.17	.24

*: Significant at 10%
**: Significant at 5%
***: Significant at 1%

As one might expect, the cooperative environment seems to encourage player A to adopt more socially efficient decisions compared with the SPNE. The neutral treatment seems to provide an environment where Player A's trust is not reciprocated, leaving trusting Player A's worse than if they had not trusted. In contrast, the cooperative and BDM treatments seem to have encouraged Player As to adopt a more trusting strategy that ultimately paid off through a higher final payment.

A more rigorous treatment analyzing the factors involved in Player A's decision can be determined through running a multivariate regression of the amount Player A placed in the trust fund on dummy variables for the treatment effects. This initial regression showed a significant (at 5%) result for the cooperative treatment. A second regression also included demographic dummy variables for gender and "college age", which both were significant at the 1% level and isolated a more highly significant result for the cooperative treatment (at 1%). The corresponding coefficients were interpreted as meaning that males contributed to the trust, on average, L$12.11 more than females and college age subjects contributed L$9.18 less than older subjects.

Table 3: Summary of Regression Output Explaining Amount A Sent

	1	2
Comp Trt.	2.45	3.28
Std.Err	5.5	5.45
Coop Trt.	13.7	16.0
Std.Err	5.5**	5.4***
BDM Trt.	-.527	7.46
Std.Err	4.8	5.14
BDMQ Trt.	-3.4	2.49
Std.Err	4.4	4.38
Male		12.11
Std.Err		3.35***
18≤Age≤22		-9.18
Std.Err		3.18***
Constant	43.33	39.03
Std.Err	3.64***	4.11***
N	597	597

*: Significant at 10%
**: Significant at 5%
***: Significant at 1%

Next, for each treatment group I calculated the percent of B's who received a nonzero amount who sent back a) anything, b) 1/3 or more of the tripled amount, c) more than 1/3 of the tripled amount, d) greater than or equal to 2/3 of the tripled amount. These results are given in this table:

Table 4: Breakdown of B's Decision

Treatment	%B's Sent Nonzero	SD	%B's Sent >1/3	SD	%B's Sent > 2/3	SD
BDM (no Q's)	.8226	.384	.6048	.501	.3548	.480
BDM (w Q's)	.7308	.445	.511	.490	.1923	.395
Competitive	.812	.394	.551	.504	.304	.464
Cooperative	.88	.327	.693	.483	.28	.452
Neutral	.478	.503	.418	.483	.149	.359

Here it is seen that the neutral treatment had the lowest frequency of B returns in each category. The other treatments clustered around 80% frequency of nonzero returns, 55% returned greater than a third of what they were sent and 44% returned greater than two thirds of what they were sent.

In analyzing B's responses when A sent nonzero, the results strongly suggest that B gifts to A were strongly dependent on A's previous decision to give. An initial simple regression of the amount B gave on the tripled amount B received from A returned a significant relationship, with a p value less than .001. The coefficient .282 is interpreted as meaning that for each additional Linden Dollar Player B received from the tripled amount Player A sent, Player B increased their gift by 28 Linden cents, which is slightly less than reimbursing Player A's initial gift. This significance persisted even after controlling for dummy treatment effects and interaction effects of treatment with the amount B was given.

Dummy variables for treatment effects were included in the second regression, with the neutral treatment as the base case. Perhaps not surprisingly, the cooperative treatment significantly increased B's giving rate by 36.6 Lindens when compared with the neutral case. What is harder to explain is that the competitive treatment also significantly increased B's giving rate by 18.8 Lindens when compared with the neutral treatment. The former finding can perhaps be accounted for by a decrease of subjects' willingness to trust and reciprocate when presented with a lack of explicit instruction about how to interpret the best course of action. However, this explanation is insufficient to

explain why participants reciprocated more in the competitive than the neutral treatment despite explicit explanation that “assuming that B wants to earn as much as possible, B will send none of what B receives to A.” A more compelling explanation is the possibility that individuals are more willing to reciprocate trusting behavior when the trust occurred despite experimenter warnings to the contrary; indeed trusting against the advice of the instructions could actually signal greater trust from Player A to Player B.

Additionally, the BDM case without follow-up questions significantly increased B’s giving rate by 21.6 Lindens, the BDM treatment with follow-up questions was not significantly different from the neutral case. One possible explanation for this discrepancy was that there was the largest incentive for subjects to participate multiple times under the guise of multiple “alts” in the BDM treatment without follow-up questions as this reduced the time to participate considerably, which could have distorted participants’ incentives.

Suppressing the regression constant in the first two regressions helps to clarify the relationship between Player B’s decision and the tripled amount he received from Player A. The downside of including a constant in the regression is that it “soaks up” a substantial part of the amount sent by the average B without a clear motivation. Consequently, it is useful to run this version of the first two regressions but unnecessary in the third when the constant term is relatively low. This methodological change most significantly alters the apparent likelihood that the amount player B received affects the amount he sent, more than

doubling the t-statistic. The full set of regression output is given by the following table:

Table 5: Summary of Regression Output Explaining Amount B Sent

	1	2	3	4	5
Amount Received	.282	.279	.278	.326	.270
Std.Err	.024***	.024***	.024***	.013***	.02***
Comp Trt.		18.8	19.5		14.3
Std.Err		9.6**	9.6**		7.34*
Coop Trt.		36.6	35.7		32.3
Std.Err		9.3***	9.35***		7.34***
BDM Trt.		21.6	16.7		16.9
Std.Err		8.5**	8.9*		5.67***
BDMQ Trt.		9.2	10.3		4.5
Std.Err		8.1	8.1		4.9
Male			-11.9		
Std.Err			-2.10**		
18≤Age≤22			-1.72		
Std.Err			5.18		
Constant	9.88	-6.1	-.101		
Std.Err	4.4**	8.2	8.74		
N	517	517	517	517	517

*: Significant at 10%
**: Significant at 5%
***: Significant at 1%

A more illuminating analysis of player B's decision in cases where A's sent nonzero is to regress the proportion of the amount B gives to A in terms of the tripled amount B received from A on the explanatory variables used in earlier regressions. This method allows insight into B's decision process without regard to endowment. The result was a statistically significant ($p<.001$) coefficient on the amount received from A both including and excluding the treatment effects. It is worth noting, however, that this coefficient is negative, specifically, -.0005.

This is interpreted to mean that for each L$10 increase from the tripled account received from player A, Player B returned .5% *smaller* proportion of the tripled fund. In the context of the earlier finding that Player B's absolute gift to A increases by 28 cents with each additional dollar B received, the data fits a story where Player B applied a strategy where the more Player A trusted, Player B responded by keeping an increasingly large portion of the joint gains.

Table 6: Summary of Regression Output Explaining Proportion B Sent

	1	2	3
Amount Received	-.0005	-.0005	-.00055
Std.Err	.00014***	.00014***	.0001***
Comp Trt.		.123	.127
Std.Err		.057**	.057**
Coop Trt.		.232	.225
Std.Err		.056***	.056***
BDM Trt.		.153	.113
Std.Err		.051***	.053**
BDMQ Trt.		.037	.045
Std.Err		.048	.050
Male			-.092
Std.Err			.034***
18≤Age≤22			-.007
Std.Err			.031
Constant	.462	.366	.410
Std.Err	.027***	.049***	.052***
N	517	517	517

*: Significant at 10%
**: Significant at 5%
***: Significant at 1%

This regression of Player B's proportion returned also provides greater insight into the treatment and demographic effects. Starting with the neutral treatment as a basis for comparison, the regression showed a significant

(p=.032, p<.001, p=.003) effect of the competitive, cooperative and BDM (without questions) treatments. The coefficients are thus interpreted to mean that, relative to the neutral treatment, the competitive treatment correlates with a 12% increase in the portion of B gives to A from the tripled fund while the cooperative and BDM (without questions) treatments each correlate with a 23% and 15% increase, respectively. Additionally, male gender is associated with a 9% reduction in Player B's proportion given, but being college age does not have a significant effect.

Discussion 1: Comparison with Outcomes from Physical Labs

Neoclassical economic theory predicts a subgame-perfect Nash equilibrium determined through backward induction where, knowing that Player B will return nothing, Player A will send nothing. However, the evidence from this experiment is consistent with offline laboratories in not supporting this prediction, even in the quasi-anonymous setting provided by virtual worlds inhabited by a population more attuned to strategic thinking. While both A and B's giving were somewhat consistently lower than BDM and in "Lavish Returns on Cheap Talk" (Ben-Ner, Putterman and Ren, 2008), these differences were not significant due to the heavy variation in individual strategies.

Table 7: Comparison with findings of similar studies: Berg, Dickhaut & McCabe (BDM) and Ben-Ner, Putterman, & Ren (BPR)

	Classic BDM	BPR	Virt. BDM-18-22	Virt. BDM-All Ages	Virt. BDMQ-18-22	Virt. BDMQ-All Ages	All Data
A Endow	$10	$1	L$100	L$100	L$100	L$100	L$100
A Giving Avg %	51.6%	63.6%	34.9%	42.8%	33.6%	39.9%	
B Return Avg %	46.6%	43.4%	40.2%	44.4%	35.4%	32.8%	
% B Returns on Amt. A Sends		Pos., Sig.	Neg., p=.37 (n/s)	Neg., p=.03**	Neg., p=.32 (n/s)	Neg., p=.14 (n/s)	Neg., Sig.
N	32	~ 26	53,58	132,134	55,49	208,182	

The finding on A's decision with respect to age is particularly interesting as it suggests that not only do college-age participants appear to behave differently from non-college age participants, but it is the reverse direction of what one might expect. The fact that differences exist between the behavior of college-age and older participants lends credence to the commonly stated concern that college students are not representative of the entire population and, therefore, experiments involving only college-age participants may not generalize across the entire population. Further, as one might think that college-age participants are more likely to naively trust than older subjects, the observed relationship is the reverse of this expectation. Consequently, as the data contains surprises from what one might assume to be the direction of the age bias, other trust experiments may suffer from an age bias that cannot be easily reasoned away.

Additionally, while "Lavish Returns" found a significantly positive coefficient when regressing the percent B returns on the amount A sent, the

findings from this study consistently produced findings in the opposite direction. In the case of BDM for all ages, B negative response to A's giving was significant. This difference could be attributed to a mindset online that is less bound by social norms compared with on college campuses.

Discussion II: Evaluation of Framing Effects

In order to validate treatment effects through non-parametric methods, Mann-Whitney Tests were applied to test the hypothesis that two independent samples are from populations with the same distribution. The following table shows the p-values from non-parametric MW tests on each treatment pair. The results confirm that player A's decisions under the cooperative treatment significantly differed from all other treatments, while none of the other treatment differences are significant. Running this test on Player B's absolute amount returned suggests nearly everything is significant, but this construction of the MW test is misleading as it does not factor for A's decision. Running the MW test on B's decision as a fraction of the tripled amount received from A yields a significant difference between the neutral treatment and all others as well as between the BDMQ treatment and all others.

Table 8: P-Values of Mann-Whitney Tests on Group Average A Contributions (Absolute) and B Contributions (Proportional) by Pair of Treatments

	Neut	Comp	Coop	BDM	BDMQ
Neut	---	.2522	.0094***	.2995	.6441
Comp	.0021***	---	.0402**	.5428	.1478
Coop	.0000***	.1132	---	.0045***	.0002***
BDM	.0001***	.4685	.3079	---	.3142
BDMQ	.0241**	.0949*	.0002***	.0035***	---

Note: Numbers above the diagonal are p-values for two-tailed tests comparing A absolute sending by treatment while numbers below the diagonal are corresponding p-values for B returning as a fraction of B's amount received, in cases when B received greater than zero Lindens.[iii]
*: Significant at 10%
**: Significant at 5%
***: Significant at 1%

However, several features of the previous regressions on Player B's decisions imply framing effects that are puzzling. First, why do both the competitive treatment and cooperative treatment correlate with an increase in the amount Player B's decision? Second, why is there a significant difference in the neutral treatment and the ostensibly "neutral" BDM treatment without the followup questions? Finally, given the difference between neutral and BDM (without questions), why does the BDM treatment with questions not share this difference?

In comparing the competitive and cooperative treatment effects with the neutral effect, it is surprising to find that each of these effects is *positive*, meaning that both the competitive and cooperative treatments appear to make Player B return a higher portion of the tripled amount they received from Player A. A further hypothesis test suggests that the cooperative treatment effect is

significantly higher than the competitive treatment effect ($p<.05$), which is interpreted as meaning that Player B gave roughly 11% more of the amount he received when subjected to the cooperative treatment rather than the competitive treatment. Broadly speaking, the significant differences between neutral and the other treatments seems very hard to explain by reference only to the difference in wording. Further analysis about the laboratory conditions could be run to explain these puzzling findings which may be ultimately attributable to experimenter error such as inconsistencies in laboratory testing conditions.

Additionally, while B's proportion returned is not significantly different from neutral under the BDM treatment with questions, when subjects received the BDM treatment without questions they returned a significantly higher portion than either of the other two groups. Indeed, an additional hypothesis test of the treatment effects associated with the two BDM treatments suggested that the BDM without questions was also significantly different from the BDM treatment with questions ($p<.003$). These differences can be explained by the fact that the BDM without questions treatment took significantly less time to complete than the other treatments, which exposed it disproportionately to the effects of "experiment farming," a previously undocumented form of experimenter bias unique to online experiments. Since this treatment took less time to complete and therefore carried a higher wage for participating in the study, there was an increased incentive for test subjects to use alternate avatars in order to reap the rewards of participating multiple times, an incentive less prevalent when treatments required participants to answer a series of follow-up questions.

However, taken at face value, this explanation remains unconvincing because it would imply a negative bias while the BDM-without-questions treatment expressed a positive effect. It would be explained, however if participants who were "experiment farming" felt a degree of guilt about their innovative method of cheating the experimenter while corrupting his data set, and relieving this guilt by being more generous with their return to Player A. Three other explanations include that 1) wealthier participants (who may be more likely to be generous with Player A) are less likely to drop-out of the experiment when not required to answer follow-up questions, 2) participants may be more generous when there were fewer conditions placed on the money they receive, specifically, since they were not expected to answer follow-up questions, they may have mentally accounted for the endowment more altruistically, and 3) B's in BDMQ felt entitled to more money to compensate themselves for the time of answering the questions. This final explanation may also account for the lower sending by A's in BDMQ than in BDM.

Chapter 4: Limitations & Future Research

Limitations

This experiment was rife with potential sources of experimenter bias because of its automated construction and lack of a present administrator. On many occasions test subjects spoke to each other or reported that they thought they knew who their partners were. While it was soon corrected by adding additional chairs, early neutral data points could have been biased toward noncooperative outcomes due to conflict over who was next to use a single available apparatus. In one known instance, a male subject was propositioned for sex while participating in the experiment, with ambiguous consequences. Further experiments could benefit greatly from starting with a more established virtual lab, by hiring a monitor to better document such incidents, and to run experiments in a larger space that allows participants greater privacy.

Theoretically, even in this context users are induced to be honest about their preferences regarding whether to trust and reciprocate because they are bearing the cost of those decisions. However, it is unclear how users valued the virtual currency relative to its real world value, which could distort their choices.

In contrast with the main experiment question, whether participants accurately providing their demographic information remains a mystery. As David Garman has commented, "I think there is information in the first life answers, but I don't know that it corresponds to the truth." This could be investigated, in part,

by providing a list of avatar names as well as their birth dates and gender to contacts at Linden Labs, the producer of Second Life, who may be willing to report aggregate levels of accuracy of these responses. However, this verification would require modification of the consent form, which is not possible after the fact.

Some demographic questions deliberately included extremely unlikely options to help flag users who may be providing inaccurate demographic data. While the data has not been cleaned for such unreliable demographic data, it would be important to remove the "junk data" before conducting a more rigorous analysis of the provided demographic data.

Additionally, there may be some selection bias present due to which users did not complete the experiment or completed it multiple times. The fact that over 80% of users who started the 15-minute experiment continued to participate through the end is a promising sign about the former concern. As for the latter, users were prevented from participating more than once with the same avatar, but as their identity was not verified, they could not be prevented from participating again under the guise of another avatar. However, while the large sample size most likely corrects for internal validity issues, overrepresentation in the group who did not complete the experiment could limit external validity of the experiment in generalizing to the broader Second Life population.

In sum, there is a lot more research that can be completed into establishing best practices in virtual experimentation and incentivizing truth-telling, though this study attempts to help build that foundation.

Answering the Critics

Prior to this research being presented, published or discussed openly, debate has emerged over the use of this methodology. Most notably, John Duffy of the University of Pittsburgh Economics Department has written a working paper criticizing the reliability of data collected from experiments in virtual worlds, citing heavily the lab developed for this research as his case study.[iv] Certainly, there are some unresolved methodological concerns with gathering data in online settings. However, it is important to weigh these costs against the benefits of online research: these methods allow us to test the external validity of general principles by experimenting on a different subject pool than the usual undergraduates; additionally, the combination of automated data gathering scripts and a population who are willing to participate at a fraction of the normal cost allows samples to be dramatically larger. The end result was that online methods allowed the collection of over 1,200 data points over two months on a graduate student budget, realizing a 95% cost savings compared with more traditional laboratory methods. This allowed evaluation of five treatments on this subject pool and tease out more subtle factors that influence behavior that might not be detected in a smaller sample.

Duffy's objections to the virtual experimental methods used in this study appear to be:

1. Data accuracy. Subjects can be dishonest about their demographic information and are more likely to do so compared with real life

2. Selection Bias. There is little control over who shows up to participate, their knowledge of economics, and the low stakes nature of the experiment may result in some subjects dropping out prematurely

3. Identity Mapping. "There is little control over whether the same individual is logged in on multiple machines, under different identities, perhaps playing a two-person game with himself."

It is important to note that even if the demographic data may not be perfectly accurate, the substance of the experiment was about subject's behavior the trust game. On this issue subjects were making decisions with real (in virtual terms) stakes about which they would be truthful. For example, while Duffy did not feel compelled to be accurate about his age and gender, he indeed answered the core experiment question on trust with what he truly believed to be the "best" course of action. Other participants selections provided results that were consistent with the trust and reciprocity effects observed in the 1995 Berg, Dickhaut, and McCabe experiment, in contrast with the subgame-perfect equilibrium expected by neoclassical economic assumptions.

What remains of questionable data integrity, however, are the 28

demographic and background questions that followed the experiment. Indeed some (17%) of participants did not choose to complete the experiment and followup questions. These "partial" data points were dropped from the resulting analysis. Whether there were patterns in peoples' decisions to drop out prematurely does indeed affect the outcome is a matter that could be tested by further experiments and data analysis.

While the followup questions were indeed not for "extra payment," it is not accurate to say that there was no incentive to complete the survey aspect of the experiment because subjects were required to complete the questions in order to receive any earnings from the prior question. Nevertheless, Duffy's concerns about the accuracy of subjects' responses is noted and is a real practical consideration in the design of experiments in any setting. In online experiments (both in virtual worlds and in web-based experiments), the absence of an authority could result in users providing inaccurate information on their demographics. One possible way to assess the accuracy of the demographic data would be to verify aspects of the data with previous data provided to Linden Labs (though this was not possible in this study due to the limitations of the confidentiality agreement provided to subjects. Hence I suggest that future experimenters insert a clause into such agreements that would allow them to share the data with the virtual world management company for validation purposes). However, this does beg for further research into mechanisms to elicit truth-telling in anonymous online settings.

Duffy's final concern is that "there is little control over whether the same

individual is logged in on multiple machines, under different identities, perhaps playing a two-person game with himself." In anticipation of this, our script prevented individual avatars from participating in the experiment more than once. We also used a delay mechanism between matched players so subjects would not know the identity of their counterpart. These two features made a two-player game with oneself practically impossible.

Underlying this concern, however, is a legitimate issue about players' use of alternate characters, known as "alts" to participate in the experiment multiple times using different avatars, a practice described earlier as *experiment farming*. This can sometimes be manually cleaned by noticing obviously duplicated avatars with names such as "Po Potez," "Po1 Potez," etc. My experience is that such experiment farming is most prevalent when participants are offered a large reward for participating in a relatively short experiment.

In closing, Duffy has identified some very real concerns to be addressed in designing effective virtual experiments. In truth, I think he is just scratching the surface about the issues that virtual experimentation needs to overcome. However, to invalidate these methods while in such a nascent state would be an overreaction. I believe the solution is to expand academic inquiry into experimentation in virtual worlds and develop better tools for collecting online data. In the mean time such confounding issues should certainly be addressed by researchers, and the field is wide open for the design of experiments to demonstrate the dimensions along which subjects behave differently in virtual worlds than the real world.

Opportunities for Further Research

Based on the experience gained during this thesis project, a number of open research threads are currently available. These can be broadly categorized into 1) improving small-scale virtual lab infrastructure to eliminate confounds and reduce bias, 2) determine dimensions virtual decision-making departs from decisions in physical labs by replicating established experiments in virtual settings and comparing outcomes, and in this context, 3) applying new technology to answer questions impossible or unethical to investigate in physical labs.

The first area of open research in virtual worlds is in improving the virtual lab infrastructure to support higher quality experiment design. It is suggested that data collected through virtual lab infrastructure embody the following improvements on the design executed in this research project:

1. Randomize between treatments on each chair, or at least cycle between them in a manner to increase consistency
2. Hire "Observer" / "Monitor" to greet participants and document potential incidents
3. Explicitly assure users of anonymity & confidentiality of responses. Ask separate questions for "Second Life" and "Real Life" demographic info.
4. Use only 3 treatments: BDMQ, Coop, Comp

5. Clean up question language to provide unambiguous framing effects and to reduce confusion in instructions
6. Provide full assurance that participants are *not* partnered with anyone currently in the lab. Also tell subjects that B will be paid immediately afterward and A will be paid after a couple of hours in order to accommodate the delay mechanism and B's decision.
7. Repeat their answers at the end of the experiment for subjects to confirm their answers

Additionally, further virtual experiments could draw on the best practices of the marketing community. Market Truths Ltd. Managing Director, Dr. Mary Gordon, mitigates the effects of those who provide inaccurate answers online through a variety of methods, with the most applicable practices including:

1. Require that people had been in Second Life for at least 30 days
2. Require respondents to have a verified account
3. Apply quality control checks, such as consistency checks across multiple small samples and algorithms to identify and remove points of likely "junk" data, as the majority of inaccurate data is provided by a small portion of survey participants.

While developing more consistent methods of conducting virtual experiments in the context of inductive metanomics, one topic that deserves special mention is a virtual behavior bias. This includes the facts that there is a

different population in virtual worlds than the real world, as well as the possibility that individuals behave differently online. Consequently, it is yet unclear how individuals' judgment and decision-making processes differ when mediated through technology. For example, there is some initial evidence that social distance is negatively associated with reciprocity (e.g. Charness, Haruvy and Sonsino, 2006) but there remain other aspects of virtual association that are yet unexplored. Better understanding the dimensions of these differences have immediate payoffs in modeling consumer behavior, and also provide enhanced context for additional research to address the heavy external validity issues when engaging in inductive metanomics.

The final area of open inquiry, applied inductive metanomics, is the most intriguing and unbounded. This includes the development and application of new technology to answer questions impossible or unethical to investigate in physical labs. Nick Yee's research involving perceived body size and aggressiveness utilizes avatars to dissociate own and others' perceptions of individual size, which is impossible without the use of virtual environments. Another example is that while ethical considerations preclude a trust game experimental design with partners known and present in physical labs due to the threat of fights, virtual worlds could help investigate this effect while mitigating the risk to participants. The above examples would utilize the avatar as the key technological enabler, but other elements of virtual worlds could be similarly utilized for scientific inquiry. Perhaps the most exciting applications will arise from the array of political economic and macroeconomic questions that could be explored through the

abundance of willing test participants motivated cost-effectively through the virtual currency premium.

It is too early to determine wither virtual world experiments emerge as a revolutionary tool or an interesting diversion. Certainly, there are benefits to be gained from the fact that virtual worlds reduce barriers to conducting experiments. At the very least, virtual labs can provide a pre-test to investigate novel theories without a large expense. Similarly, with the proper guidance, students could design and execute experiments of their own as a class exercise, through the use of virtual world methods. However, the key issue of establishing data collected in virtual worlds as a valid laboratory for testing real-world theories will take a significant amount more research into developing virtual research methods, understanding virtual behavior, and crafting novel research questions to best utilize virtual experiments.

Appendix I: Instructions with Differences by Treatment Group

The language differences are highlighted in bold for the purposes of this explanation but were not emphasized to the participants. Additionally, comments inside [brackets] were not supplied to subjects and are also included here for clarity.

Instructions Read by Treatments 1-3: Neutral, Cooperative, Competitive

Thank you for volunteering to participate in this decision-making experiment. We expect it to take less than 15 minutes to complete.
You will be asked a question about distributing an amount of Second Life Lindens between yourself and another player.
After that, you will be asked a series of demographic and background questions.
Please take a moment to review the consent card obtained from this object.
|ready|get card

Do you give your consent to participate in this study?
AUTHORIZATION OF CONSENT
I have read the consent form. I confirm that the purpose of the research, the study procedures, the possible risks and discomforts as well as the benefits have been explained to me.|yes|no

"Before we begin, please confirm your real-life age."

You have been randomly paired with another individual to engage in a two-stage decision-making process in which your earnings will depend on both your own and the other person's decision.

[The next line depended on the treatment.]

[Treatment I: Neutral]
You will be assigned the role of either Person A or Person B. Before telling you which role is yours, we'll describe the process. You and your **counterpart** are given the identical instructions.|ok

[Treatment II: Cooperative]
You will be assigned the role of either Person A or Person B. Before telling you which role is yours, we'll describe the process. You and your **teammate** are given the identical instructions.|ok

[Treatment III: Competitive]
You will be assigned the role of either Person A or Person B. Before telling you which role is yours, we'll describe the process. You and your **rival** are given the identical instructions.|ok

If you follow the instructions and complete the required decisions, you can guarantee yourself earnings of L$100.
Person A, who makes the first decision, can try to earn more than 100, but at the risk of ending up with less—in the worst case for A, nothing.
Together, A and B can earn up to a maximum of L$400 that can be split between them in various ways.|ok

The procedure is as follows. A and B each begin with their L$100.
A's decision:
A moves first. A decides whether to send some of his/her L$100 to B. A can send 0, 1, 2, …, or 100 Lindens.
This is A's only decision. Whatever A sends to B is tripled. If A sends thirty to B, B receives ninety; if A sends zero to B, B receives zero.
The maximum combined earnings are L$400 because if A sends L$100 to B (the maximum possible), B gets L$300.
Added to the L$100 B already has, B would have L$400.|ok

B's decision:
The second decision is made by B. If A sends to B any amount other than zero, B can send some of the money received back to A.
For example, if A sends fifty so that B receives one hundred fifty, B may send 0, 1, 2, …, 149 or 150 to A. |ok

Notice that B is free to send nothing, and that B cannot send more than the one hundred fifty that B received from A.
In other words, B always ends up holding B's original L$100, plus any of the amount received that B chooses not to send back.|ok

Both of you can earn more if A sends more to B and if B sends back at least a little more than the amount that A sent.
For example, if A sends 100 and B sends back 120, A earns 120 and B earns 280, whereas if A sends nothing, A and B each earn only 100.|ok

[The next line depended on the treatment.]

[Treatment I: Neutral]
[No additional line was present]

[Treatment II: Cooperative]
By working together and acting fairly, A and B can both double their money (A can send 100, B can return 200, and they therefore each earn 200 instead of each earning 100). |ok

[Treatment III: Competitive]
However, assuming that B wants to earn as much as possible, B will send

none of what B receives to A.
So if A sends B ten and B does what leaves B with the largest amount of money, B will earn 400 and A will earn zero.|ok

We recommend that you read the instructions a second time and think about the problem that you face before moving to the decision stage. Would you like to repeat the instructions?|yes|no

A) "You have been assigned to role A. Please decide what number of Lindens (if any) from your L$100 endowment to send to your counterpart, Person B. This amount will be tripled, and then Player B will decide how much of this amount to return."
B) "You have been assigned to role B. In addition to your L$100 show-up fee, Person A decided to send you L$X. Due to tripling of this money, the amount you have received is L$(3*X). Please decide how much (if anything) you want to send to person A."

Please remember that you must answer the following demographic questions to receive your earnings. Your answers will not be used for any purpose other than for academic research, and your honesty is essential for useful results.|ok

Instructions Read by Treatments 4-5: BDM, BDM with Demographic Questions

[The first line differed by treatment.]

[Treatment 4: Original BDM w/o Demographic Questions]

Thank you for volunteering to participate in this economics experiment. You will be asked a series of **questions**. We expect it to take less than 15 minutes to complete.

[Treatment 5: BDM with Demographic Questions]

Thank you for volunteering to participate in this economics experiment. You will be asked a series **of experiment and background questions**. We expect it to take less than 15 minutes to complete.

Do you give your consent to participate in this study?
AUTHORIZATION OF CONSENT
I have read the consent form. I confirm that the purpose of the research, the study procedures, the possible risks and discomforts as well as the benefits have been explained to me.|yes|no

"Before we begin, please confirm your real-life age."

You have been asked to participate in an economics experiment. The instructions you are about to read are self explanatory. If you have any questions, you should read back through these instructions.

In this experiment you will be paired with a different person whose identity will not be made known to you. Your identity will not be made known to that person either.

Both you and the other person are receiving L$100 as a show up fee for participating in this experiment. One of you will have the opportunity to send some, all, or none of their show up fee to the other. Each Linden sent by this first decision-maker will be tripled. For example, if you are the first decision-maker and you send L$20, your counterpart will receive $60. If you send L$9, your counterpart will receive L$27. The second person will then decide how much money to send back to the first person and how much to keep.

Would you like to repeat the instructions?|yes|no

A) "You are the first decision maker. Please decide what number of Lindens (if any) from your L$100 endowment to send to your counterpart, Person B. This amount will be tripled, and then Player B will decide how much of this amount to return."
B) "You have been assigned to be the second decision-maker. You were sent L$X by your counterpart, which means you have received X in addition to your L$100 show-up
fee. You can send any part, including nothing, of the money you received, but you keep your original L$100 plus any portion of the amount you choose to keep. Unlike with your counterpart, any money you send goes to that individual as it is, without being tripled.

[One additional instruction was supplied to Treatment 5: BDM w/ Questions:]

Please remember that you must answer the following demographic questions to receive your earnings. Your answers will not be used for any purpose other than for academic research, and your honesty is essential for useful results.|ok

Appendix II: Demographic/Attitudinal Questions

Additionally, these 28 follow-up questions were asked of all participants except the "BDM" treatment.

Question 1 of 28
How much money (in $L) does your SL account currently contain?

below 1K
1K-10K
10K-100K
100K-1M
above 1M

Question 2 of 28
What is your typical earnings (in $L) during an hour of income-generating activity in SL?

below 100
100-500
500-750
750-1,250
1,250-2K
2K-4K
4K-8K
above 8K

Question 3 of 28
Roughly how many hours per week do you spend playing SL?

below 3
3-5
6-10
11-15
16-20
21-30
31-40
41-60
above 60

Question 4 of 28
What is your average hourly wage in first-life (Real World) paid work (in USD)?

below 5.85
5.85-8
8-12

12-15
15-20
20-40
above 40
No answer

Question 5 of 28
What is your average monthly wage in first-life (Real World) paid work (in USD)?

below 500
500-1K
1K-1.5K
1.5K-2K
2K-3K
3K-4K
4K-6K
6K-8K
8K-12K
above 12K

Question 6 of 28
How many hours a week do you work for a wage or to generate income (Real World)?

below 10
11-20
21-30
31-40
41-50
51-60
61-70
above 70

Question 7 of 28
How do you characterize your level of educational attainment:
1) No high school completed
2) Some high school completed
3) High school graduate
4) Some college
5) College graduate
6) Completed graduate study

1
2
3
4

5
6

Question 8 of 28
Employment/Education Status:
1) Work Full Time
2) Full-Time Student
3) Part-Time Education/Employment

1
2
3

Question 9 of 28
How would you characterize your current profession?
1) non profit
2) business executive
3) business employee
4) education/academic
5) government
6) military

1
2
3
4
5
6

Question 10 of 28
Gender:

male
female

Question 11 of 28
Race:
1) Caucasian
2) African American
3) Asian/Pacific Islander
4) Hispanic/Latino
5) American Indian
6) Multiracial
7) No answer

1

2
3
4
5
6
7

Question 12 of 28
Where were you born?
1) North America
2) Latin America and Caribbean
3) West Europe
4) East Europe
5) South Asia
6) East or South East Asia
7) Australia, New Zealand
8) Africa
9) North Africa, Middle East

1
2
3
4
5
6
7
8
9

Question 13 of 28
Where do you currently live?
1) North America
2) Latin America and Caribbean
3) West Europe
4) East Europe
5) South Asia
6) East or South East Asia
7) Australia, New Zealand
8) Africa
9) North Africa, Middle East

1
2
3
4
5

6
7
8
9

Question 14 of 28
Number of Siblings:

0
1
2
3
4
5-6
7+

Question 15 of 28
Number of Younger Siblings:

0
1
2
3
4
5-6
7+

Question 16 of 28
Marital Status:
1) Single
2) Engaged
3) Married/Domestic Partnership
4) Divorced/Separated
5) Widowed

1
2
3
4
5

Question 17 of 28
Number of Children:

0
1

2
3
4
5-6
7+

Question 18 of 28
Currently under the influence of mind-altering drugs or alcohol?

yes
no
I'd like to

Question 19 of 28
Which of the following best describes the frequency with which you attend a church or other religious service?
1) Never
2) Once a year
3) A few times a year
4) Monthly
5) Weekly
6) Multiple times a week

1
2
3
4
5
6

Question 20 of 28
Have you contributed money or time to a charitable organization or cause during the past year?

yes
no

Question 21 of 28
Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?
1) Most people can be trusted.
2) You can't be too careful.

1
2

Question 22 of 28
In your opinion, are fellow gamers more trustworthy, less trustworthy, or about the same as people in the general population?
1) More trustworthy
2) About the same
3) Less trustworthy

1
2
3

Question 23 of 28
In your opinion, are fellow gamers more deserving of reciprocal treatment, i.e. deserving of having a kindness returned, than people in the general population?
1) More deserving of reciprocal treatment
2) About the same
3) Less deserving of reciprocal treatment

1
2
3

Question 24 of 28
Where did you find out about this experiment?
1) Just wandering by
2) Saw event posted
3) Word of mouth

1
2
3

********** starting from here: different questions for A and B **********

Question 25 of 28
How much $L do you think B would have sent to you if you sent L$100 to them, out of a total of L$300?

0
1-50
51-100
101-150
151-200
201-250
251-299

300

Question 25 of 28
How much of L$300 would you have sent if A sent you L$100?

0
1-50
51-100
101-150
151-200
201-250
251-299
300

Question 26 of 28
In your opinion, what would B send back if you sent L$50 to them, out of a total of $150?

0
1-25
26-50
51-75
76-100
101-125
126-149
150

Question 26 of 28
How much of L$150 would you have sent if A sent you L$50?

0
1-25
26-50
51-75
76-100
101-125
126-149
150

Question 27 of 28
How much $L do you think the average person from the general (Real World) population would send to B if they were making the same decision as you?

0
1-10
11-20

21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-99
100

Question 27 of 28
In your opinion, what percentage of L$ would the average person from the general (Real World) population have sent to A, i.e. if they were presented with the B decision you were presented with before?

0%
1-10%
11-20%
21-30%
31-40%
41-50%
51-60%
61-70%
71-80%
81-90%
91-99%
100%

Question 28 of 28
In your opinion, if someone sent that average amount you just identified, what percentage of its tripled amount in $L would B send back to A?

0%
1-10%
11-20%
21-30%
31-40%
41-50%
51-60%
61-70%
71-80%
81-90%
91-99%
100%

Question 28 of 28

In your opinion, how much $L would the average person from the general (Real World) population have sent if they were asked Player A's question?

0
1-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-99
100

References

Abrams, David. "A Primer on the Economics of Virtual Worlds," Working paper, University of Chicago Law School. 2007.

Ben-Ner, Avner, Fanmin Kong and Louis Putterman. "Share and Share Alike? Gender-Pairing, Personality, and Cognitive Ability as Determinants of Giving." Journal of Economic Psychology 25.5 (2004): 581-589.

Ben-Ner, Avner, Louis Putterman and Ting Ren. "Lavish Returns on Cheap Talk: Non-Binding Communication in a Trust Experiment." Working paper. Department of Economics, Brown University.

Bloomfield, Robert. "Blast from the Past: Metanomics 101, Revisited." Metanomics 2008 – Business and Policy in the Metaverse. Jan 29 2008. < http://www.metanomics.net/29-jan-2008/blast-past-metanomics-101-revisited >

Bloomfield, Robert. "Worlds for Study: Invitation; Virtual Worlds for Studying Real-World Business (and Law, and Politics, and Socioogy, and…)." Working paper, Johnson Graduate School of Management, Cornell University. 2008.

Bochet, Oliver, Talbot Page and Louis Putterman. "Communication and Punishment in Voluntary Contribution Experiments." *Journal of Economic Behavior & Organization. 60.1 (2006): 11-26.*

Berg, Joyce, John Dickhaut, and Kevin A. McCabe. "Trust, Reciprocity, and Social History." Games and Economic Behavior 10 (1995): 122-142.

Castronova, Edward. "Virtual Worlds: A First-Hand Account of Market and Society on the Cyberian Frontier." Working paper, Department of Telecommunications, Indiana University Bloomington.

Castronova, Edward et. al. "Synthetic Worlds as Experimental Instruments." *Video Game Theory Reader II*. Routledge. Spring 2008.

Charness, Gary, Ernan Haruvy and Doron Sonsino. "Social Distance and Reciprocity: An Internet Experiment." *Journal of Economic Behavior & Organization.* 63.1 (May 2007): 88-103.

Chesney, Thomas, Swee-Hoon Chuah and Robert Hoffmann. “Virtual World Experimentation: An Exploratory Study.” Working paper, Nottingham University Business School. 2007.

Duffy, John. “Trust in Second Life.” Working paper, Department of Economics, University of Pittsburgh. 2008.

Fehr, Ernst and Simon Gächter. “Reciprocity and Economics: The Economic Implications of *Homo Reciprocans*.” European Economic Review 42.3 (1998): 845-859.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr. “Are People Conditionally Cooperative? Evidence From a Public Goods Experiment.” Economics Letters 71 (2001): 397-404.

Gordon, Mary. “Follow Up Comments.” *Griefers in Virtual Research and the Truth Serum.* Metanomics. March 2008. < http://metanomics.net/25-mar-2008/griefers-virtual-research-and-truth-serum#comments >.

Hoffman, Elizabeth, Kevin A. McCabe, and Vernon L. Smith. “Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology.” Economic Inquiry 36.3 (1998): 335-352.

Linden Labs. “Land: islands.” *Second Life I Land: Islands*.” Linden Labs. Feb 2008. < http://secondlife.com/community/land-islands.php >.

Linden Labs. “The World.” *Second Life | The World. 2008.* Linden Labs. Feb 2008. < http://secondlife.com/whatis/world.php >.