SSOAR

Open Access Repository
www.ssoar.info

Surveying the virtual world: a large scale survey in second life using the Virtual Data Collection Interface (VDCI)

Bell, Mark W.; Castronova, Edward; Wagner, Gert G.

Veröffentlichungsversion / Published Version
Forschungsbericht / research report

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
SSG Sozialwissenschaften, USB Köln

Empfohlene Zitierung / Suggested Citation:
Bell, M. W., Castronova, E., & Wagner, G. G. (2009). *Surveying the virtual world: a large scale survey in second life using the Virtual Data Collection Interface (VDCI).* (RatSWD Research Notes, 40). Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD). https://nbn-resolving.org/urn:nbn:de:0168-ssoar-416287

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

German Council for Social
and Economic Data (RatSWD)

www.ratswd.de

RatSWD
Research Notes

Research Note No. 40

Surveying The Virtual World

A Large Scale Survey in Second Life
Using the Virtual Data Collection Interface (VDCI)

Mark W. Bell, Edward Castronova, Gert G. Wagner

June 2009

Federal Ministry
of Education
and Research

Research Notes of the Council for Social and Economic Data (RatSWD)

The *RatSWD Research Notes* series publishes empirical research findings based on data accessible through the data infrastructure recommended by the RatSWD. The pre-print series was launched at the end of 2007 under the title *RatSWD Working Papers*.

The series publishes studies from all disciplines of the social and economic sciences. The *RatSWD Research Notes* provide insights into the diverse scientific applications of empirical data and statistics, and are thus aimed at interested empirical researchers as well as representatives of official data collection agencies and research infrastructure organizations.

The *RatSWD Research Notes* provide a central, internationally visible platform for publishing findings based on empirical data as well as conceptual ideas for survey design. The *RatSWD Research Notes* are non-exclusive, which means that there is nothing to prevent you from publishing your work in another venue as well: all papers can and should also appear in professionally, institutionally, and locally specialized journals. The *RatSWD Research Notes* are not available in bookstores but can be ordered online through the RatSWD.

In order to make the series more accessible to readers not fluent in German, the English section of the *RatSWD Research Notes* website presents only those papers published in English, while the the German section lists the complete contents of all issues in the series in chronological order.

The views expressed in the *RatSWD Research Notes* are exclusively the opinions of their authors and not those of the RatSWD.

The *RatSWD Research Notes* are edited by:

Chair of the RatSWD (2007/ 2008 Heike Solga; 2009 Gert G. Wagner)

Managing Director of the RatSWD (Denis Huschka)

Surveying The Virtual World

A Large Scale Survey in *Second Life* Using the Virtual Data Collection Interface (VDCI)[i]

Mark W. Bell

Indiana University: bellmw[at]indiana.edu

Edward Castronova

Indiana University and DIW Berlin: castro[at]indiana.edu

Gert G. Wagner

Berlin University of Technology, DIW Berlin, and Max Planck Institute for Human Development, Berlin: gwagner[at]diw.de

Abstract

Technology has always introduced changes in the way researchers administer surveys. A new technology known as *virtual worlds* has now emerged that promises to change data collection once again. Virtual worlds are persistent, online, computer-rendered spaces populated by hundreds, thousands, or even millions of people at a time. Previously, this population has only been surveyed in ways that required respondents to exit the virtual world before giving their answers. No survey method has existed whereby they could be surveyed while remaining present in the virtual space. Needless to say, this is less than ideal for any survey about the respondent's attitudes, perceptions, and behavior within the virtual world itself. This study introduces a method for solving this problem and a tool that allows surveys entirely within a virtual environment. The method is introduced as Virtual Assisted Self Interview (VASI), and the tool for implementing it, the Virtual Data Collection Interface (VDCI). The tool was created and deployed in the virtual world *Second Life* (SL), where users were asked questions about demographics and quality of

i This study was conducted in a virtual environment created with funding from the German Federal Ministry of Education and Research (BMBF), Grant No. 01UW0706 - PT-DLR.

life. The valid response numbers for the survey (N=2094) make it the largest in-virtual-world data collection seen so far. This paper discusses the VDCI and describes several different sampling methods, as well as results that provide unique, new insights into virtual world populations. It is found, for example, that the demographic make-up of SL is unlike that of other virtual worlds. Moreover, the SL population is unlike that of other worlds in its approach to gender-switching. The limitations and new hazards of virtual world survey research are also discussed, especially survey "hacking" by individuals hoping to exploit the survey for financial gain. Despite the challenges, the results generally suggest that the VDCI is a valuable new research tool for obtaining representative data on virtual world population.

Keywords: VASI, VDCI, Second Life, Survey Plan, Fieldwork
JEL Classification: C81, C82, Z19

1. Introduction

In the past few years, a new form of computer media environment has emerged from the Internet. This new environment is known as a “virtual world," which can be defined as a "synchronous, persistent, network of people, represented by avatars, facilitated by computers." (Bell & Robbins-Bell, 2008). Examples of virtual worlds include Linden Lab's *Second Life* (SL), Blizzard's *World of Warcraft* (WoW), and Sony Entertainment's *Free Realms*. All told, virtual worlds today have millions of users interacting 24 hours a day 365 days a year. Yet despite having populations larger than many cities and even countries, there is today no systematic data collection among the virtual world population. One explanation might be that these places are “mere games” and not worthy of attention. However, virtual worlds are now used by businesses, governments, and foundations to conduct serious activities. Another explanation relates to the lack of government jurisdiction – no virtual world is a “part” of any Earth country, any more than the internet is. The populations of virtual worlds are mixed samples of the populations of all the countries on Earth, and no one country’s central statistical office is likely to decide that this population is part of its target population. Finally, and most critically, there is not even a tool for conducting survey research inside virtual spaces. At best, current practice is to contact members of a virtual world community and invite them to go to a website to take a survey. There is no known method for surveying them inside the virtual world itself. Thus while virtual worlds have emerged as an important site of human activity, there is literally no general information available about the humans who spend their time there.

The lack of a tool that preserves immersion is a significant problem. When a user is in a virtual world they have a sense of immersion that involves feelings and perceptions that are particular to that environment (Bystrom et al., 1999). Surveys have been done of virtual world in the past but have involved the respondent leaving the virtual world experience and completing a web-based survey (Griffiths, Davies, & Chappell (2003), Seay, Jerome, Sang Lee, & Kraut (2004), Yee (2006a), Yee (2006b), Yellowlees & Cook (2006), Williams (2006c), and Kemp & Livingstone (2006)). Moving out of the virtual world to answer a survey creates a break in immersion and thus potentially prevents accurate recall of the virtual environment.

To move forward on the absent-tool problem, this paper reports the deployment of a new survey method and research tool for collecting data entirely within a virtual world, that is, without requiring the respondent to break out of the virtual environment to a web survey. To combat the break in immersion, a new data collection method has been proposed, the Virtual-Assisted Self Interviewing (VASI), and a specific instrument, the Virtual Data Collection Interface (VDCI), has been developed (Bell, Castronova and Wagner, 2008). These new methods and tools were given their first test in the virtual world *Second Life*, and the results are reported here. They produced one of the largest surveys ever done in a virtual world, and perhaps the first to have a claim to genuine representativeness and full immersion in the space.

After this introduction of basic issues, Section 2 of the paper considers some of the unique features of any survey in virtual worlds (as opposed to other, more mundane environments). Section 3 discusses our survey method, and Section 4 the specific sampling procedures. Section 5 discusses the process by which these methods were implemented within Second Life, and the specific challenges that came up. Section 6 provides some basic results. Section 7 narrows focus to one example – "gender-bending" behavior – to illustrate the research capabilities that the survey provides. Section 8 concludes.

2. Surveying the Virtual World

At first glance, obviously, there are genuine differences between the real world and the virtual world. The real world is not mediated by a screen. The real world requires no computer interaction. The virtual world allows for fluid identity, fantastic traits and adventures that defy the biology and physics of the real world. However, the human interactions in virtual spaces are absolutely real. When someone sends the message of "I hate you," as opposed to "I love you," the human feelings being delivered are genuine even though the symbols used to send the message may vary and are in a sense epiphenomenal. Sometimes the symbols will be letters on paper; sometimes bytes rendered on a screen (as now), and sometimes the gestures of a computer avatar. The virtual world is not "real" but the people and their passion, and hence their society, certainly is (Castronova & Falk, 2008).

This new form of computer mediated environment has raised research questions in several areas such as, social norms (Yee et al., 2007), online learning (Childress & Braswell 2006) and embodied

communication (Cassell, 2000). These studies, and many others, have begun to contribute to the understanding of the mechanics and complexity of virtual worlds. As with any new research space, old techniques need not be thrown out and replaced with new ones simply because of a change in technical possibilities. Lessons learned in other environments may be translatable to new technologies. Survey researchers have integrated new technologies as they expand their researcher's toolkit.

Though surveying the virtual world follows many existing methods of real world survey research there are some different aspects that need to be understood and discussed. Moschini (2008) states that virtual world research presents both obstacles and opportunities. Based on Moschini these include venue choice, expertise, ethical considerations, and data. The Virtual Data Collection Interface (described in detail in Bell, Castronova and Wagner, 2008) was designed to implement a Virtual Assisted Self-Interview (VASI) method in a way that accommodates the issues particular to virtual environments. VASI is a method that, in essence, allows a user who is in a virtual space to click on a survey tool that is in there as well and take the survey by himself. The VDCI is a specific tool for VASI, developed within the virtual environment of *Second Life.* The VDCI is a box that a user finds when wandering about in the virtual world. Click the box, and a display appears on screen that only the user sees. The display presents questions and has buttons for the user to click. As each question is answered, the display sends another. When the survey is done, the display goes away. Users can stop and start, skip around, and so on, just as with web-based surveys. The box can send a payment as a reward for taking the survey. The box also can implement a prior consent protocol and ensure anonymity.

The goal in designing VASI and the VDCI was to produce a data set as close as possible to the standards of survey research in the offline world. Thus, it will now be considered whether the VDCI handles the issues raised by Moschini (2008) in a way that preserves the real-world validity of the results.

To begin with venues: Survey research in the real world can take place in many different venues. People can be contacted in a public space, in their homes (door-to-door, telephone or web surveys) or in a more controlled environment (a research lab). To handle the venue issues in a virtual world, one needs to create an instrument that allows any user to take the survey anywhere in the virtual space. The VDCI is a box that can be placed anywhere in the virtual world for which it was designed, *Second Life*.

As for expertise, in the real world a survey protocol and method must be simple enough that anyone can be a respondent. In a virtual world, this means the survey must be accessible to any user, regardless of whether they are expert users of the space. The VDCI was placed on an ordinary island in SL (the IU-DIW Island) and also at 13 other locations that included all sorts of content. Anyone minimally conversant with getting around in Second Life could encounter the survey box. Once they did, a set of clear instructions led them into the survey, and answering was a simple matter of clicking on boxes. This protocol created a virtual (occurring in the virtual space), assisted (because of the VDCI), self-interview, that anyone could take.

The virtual world also adds new ethical considerations to survey research. To begin, existing research ethics should be maintained and followed. The participants in virtual worlds are humans (represented by their "avatars") and all previous human subjects protocols must be followed (consent, IRB review, and transparency). However, new ethical considerations arise in the virtual world. First, SL has a Terms of Service (ToS) agreement that all users must abide by. Any research in SL must follow these terms of service. This includes no clandestine monitoring or recording of text or voice and using it in research while identifying the avatar. Second, as with the real world consent must be obtained. Finally, identity in SL is fluid so ensuring respondent integrity is difficult.

The VDCI addresses all of these issues. When a user sees the box, she knows it is an academic survey right away. The boxes were clearly labeled with institutional logos and contact information. When the user clicks on the box, she sees a notecard that requests consent. No one else can see that she is taking a survey. Once she agrees to take part in the survey the VDCI captures the avatars name (till the end of the survey period) to ensure that the same avatar does not take the survey multiple times. This does not prohibit a user with multiple avatars from taking the survey multiple times. Thus, the VDCI ensures openness, privacy, and consent throughout.

Moschini's (2008) final issue with virtual world surveys is the data. The intent of the VDCI was to collect data in "real-world" format. Data was written to a database and handled, from that point forward, just like any other survey data set. Thus, the VDCI handles the data issue as well.

The VASI method implemented with the VDCI thus seems to address the main issues raised a priori by authors like Moschini (2008). More details about the VDCI and the Virtual Assisted Self Interview method are available in the earlier working paper (Bell, Castronova and Wagner, 2008).

There are other issues involving virtual surveying that have not, however, been addressed by prior writers. One is technical problems. Every survey method can have technical issues – someone

might spill coffee on a paper survey, for example. However, in virtual environments technical issues are typically more global. If someone spills coffee on a computer, it ruins the entire database, not just one survey response. A second overlooked issue is interference. It is not usual, in real-world survey collection, for other people to actively attempt to disrupt a survey in progress. Possible examples of this, however, might be when people not affiliated with an organization make fake survey calls in its name, to dupe the respondent into sending a "donation" to the caller. In virtual environments, there is a general absence of law and some degree of social indifference, so that this kind of interference is more common. In Section 5.2 below, as the specific data collection experience is being described, some responses to actual technical and interference events will also be discussed. It would be useful for researchers to develop conceptual protocols, as Moschini (2008) has done, for dealing with similar issues in the future.

3. Survey Method

Though the survey uses a unique new form of data collection, the initial stages of survey creation were exactly like those for a real-world survey (see Bell, Castronova, & Wagner, 2008). The unit of analysis was chosen to be an individual avatar. Following this, the variables to be collected were defined and a sampling plan was created.

3.1 VASI and VDCI

As described above, a virtual-assisted self-interview protocol was implemented through the creation and placement of virtual data collection interfaces, the VDCI boxes. This appears to be the first ever such implementation. The VDCI was created specifically for research in the virtual world of *Second Life*.

Second Life (SL) is a virtual world created by Linden Lab, Inc., a media company in San Francisco USA. Users connect to SL by downloading and installing a software client that connects to the SL servers allowing the user access to the virtual world. SL also allows creation of objects and scripting of those objects. These two features of SL allow the VDCI to be built and distributed completely within the virtual world of SL. The VDCI in SL, when clicked, generates a Heads-Up Display (HUD) on the user's screen. The avatar would "wear" the HUD, therefore displaying the survey instrument on the user interface allowing the user to answer the questions through the SL

client interface. The VDCI uses LSL (Linden Scripting Language) which formatted HTTP calls that use PHP to write the respondent's answers to a MySQL database.

3.2 Specific Instrument Creation

The main aim of our first survey is to obtain some basic demographic figures of our respondents and to get their permission for a longitudinal survey (see for an example Schimmack et al. 2009). Beyond the basic demographic questions we asked some questions about well-being, since possible differences in real-life second-life satisfaction might explain why some people 'go virtual' and others do not. In order to make our future research comparable with research about real life (RL) we use the basic demographic concepts which are common for research in RL. And especially we use some of the questions of the German Socio-Economic Panel Study (SOEP) which is a widely used longitudinal survey (Siedler et al., 2009). For the basic questions see Appendix E.

We conducted our first survey just in English. No translations were done. The answers were written in the data based whenever a respondent clicked a survey answer so even incomplete surveys had some data collected. Once a survey was complete (there was an answer clicked for the final question) the survey was marked in the database as complete. When an avatar had completed a survey, the respondent was unable to answer the questions again. By means of the very last question of the survey it is possible to ask the respondents whether they give permission for a re-contact which allows us to establish a longitudinal panel study.

The VDCI survey was placed within a "kiosk" (see Figure 1 -- image of a kiosk) that identified the researchers (logos of institutions) and displayed text that asked the user to click on the box to receive a survey. This kiosk automatically dispenses the survey HUD along with a note card containing instructions and consent information (see Appendix A). Both of these items were marked in a manner that would not allow the user to modify the survey or note card in any way.

The user was also not allowed (by using object permissions) to give the survey tool to another user or create copies. These software measures were taken to ensure one avatar could only respond to the survey HUD they were given.

Once the user has been given the survey HUD and the note card (which opens automatically once the HUD is attached), the first question of the survey asks if the user has read the note card text which covers consent information for the survey. Answering "Yes" to this question meant the user gave informed consent.

The final part of the survey is the payment which is administered by means of a "payment box". Once the user had answered the first question (either Yes or No) they were then allowed to click on the payment box on the kiosk. This then paid the avatar 250 linden dollars. During the survey period 250 linden dollars was roughly equivalent to 0.97 USD. This payment comes from the research avatar which we used for administering the survey (see Figure 2). The money is paid directly to the avatar who has taken the survey. Once one avatar has been paid it was recorded in the database and the avatar was not permitted to be paid again. [1]

4. Sampling

Sampling the virtual world is not similar to sampling in the real world. The population our research was interested in was the universe of individuals who use SL. Three sampling methods were devised to collect different samples of the SL population.

The basic problem is that there is no register of SL users (inhabitants) which can be used for drawing a random sample. Linden Lab does not provide this information for commercial or for research purposes. In SL, there are also no permanent addresses of avatars. So it is not possible either to apply a kind of "random walk procedure" in order to create a random sample with equal probability weights. Other methods which all rely on some sort of "self selection" are necessary. A look in the literature and other information about surveys which took place in SL reveal that up to now no random sample was administered. In fact, most surveys are heavily selective[2].

For our survey, we used three survey methods so we can compare results and give an estimate of the impact of different methods. First, three large email lists were contacted. Second, a classified ad was placed in SL. Third, a quasi-random method for placing survey kiosks was used. Each of these methods had a unique survey HUD so that the sampling method could be recorded in the database.

(1) The first method used the sampling frame of three SL related mailing lists. These lists were the SL Educators List (5400 individuals), SL Researchers List (1900 individuals) and the Association of Internet Researchers List (about 4500 individuals). These lists were contacted via e-mail (Appendix B) and asked to go to a specific location in SL and answer the survey. This resulted in principle in a random sample of the members of those lists. Compared with the universe of all SL users it is most likely a heavily biased sample because SL users who are interested in education and research most likely did not represent the "average" user community. But, because the

characteristics of all SL users are not known yet it is not easy to estimate the degree of the bias. Another reason for a bias is because only users who are able to read English can participate and give valid answers.

(2) Secondly, using SL's in-world classified ad system a classified ad was placed (Appendix C). This ad creates a convenience sample because respondents select themselves into the survey. And, first of all, our second sampling method requires potential respondents to know about the SL classified system, be searching for a way to make money inside SL or to participate in surveys (the latter was a reason for very few respondents, we believe). Again, potential respondents must be able to read and understand English. Before they received to the survey the user must be able to follow the instructions (in English) in the classified ad . This ad included a teleport to a unique location for the classified ad survey. All in all this resulted in a convenience sample of SL users who are able to read English. However, because every user can read the ads there is no bias produced by the advertisement (whereas the advertisement via the lists we used for our first sampling method result in a bias).

(3) Finally, a random location protocol was used to create a quasi-random sample of the population. In order to try to obtain a random sample of SL users we used a randomizing script (by SL Name: SignpostMarv Martin) that randomly selected a location from the searchable locations in SL (random location protocol). A location can be searchable or not. By default all locations are searchable. To be removed from the list, the land owner needs to manually switch the setting for the land to "not searchable". So, most of the locations are searchable. However, there might be some bias created by restricting a random placement to searchable locations.

The script returns a random location from the list of searchable locations. We recorded the name of the land and its SL address or "SLURL"[3]. Then our research avatar (figure 2) teleported to the location and gathered the land owner's name. If the land was group owner (a common practice in SL) the owner of the group was used. This owner was then sent an instant message explaining who the researchers were, the nature of the survey and asked if a survey kiosk could be placed on the land. The date of this contact was recorded. If the owner agreed the researcher placed a kiosk as close to the random location as the owner would allow. A screen shot was taken of the kiosk at the random location (see Appendix D with the SLURL of all randomly dropped kiosks). This method tried to maximize the amount of randomness in users taking the survey.

However, even this method do not result in a conventional random sample because the kiosks do not “ask” (in an active manner) all avatars who come by. So no participation rate can be estimated and no analysis of the structure of response (which most likely is not completely at random) can be applied. The random kiosks only attract respondents who are curious and click on the kiosk. According to the rules of Terms of Service the kiosks do not “scan” their neighborhoods and they do not count the avatars (and do not scan their characteristics either) who come by without taking the survey. Thus - comparable to the ad-sample - the random-kiosks-sample resulted in an “unbiased” convenience sample of SL users who are able to read English.

To summarize: we developed and applied three different channels for doing the sampling of our survey:

1. e-mail lists
2. classified ad
3. quasi-random survey protocol (random location protocol)

All channels – and especially the classified ad sample – were also susceptible to "snowballing". Many of the respondents probably told their friends where to find a kiosk with the survey tool as a way to make money. The number of respondents may have been increased through word-of-mouth. Therefore, all the subsamples cannot be seen as all coming from the sampling frame which we recorded.

We expect that the respondents of channel 1 (e-mail-lists) will differ from the respondents from channels 2 and 3 (ad and random kiosks). It could be possible that the respondents of channels 2 and 3 may not be significantly different from each others. If this would be the case than we can draw the conclusion that the class ad method comes close to a quasi-random method – but because class ad is much cheaper it would be the more efficient method.

5. Data Collection

5.1 Overview

The survey ran for 30 days in early 2009 (02/03 to 03/05). The VDCI allowed data to be collected 24 hours a day with no permanent researcher involvement.

At the beginning of the study the e-mails were sent out (02/03/09). On the same day, the classified ad was placed in the Second Life Classified Ad system. And then the random kiosks were

placed. Using the random protocol one hundred and twenty five locations were contacted and kiosks were placed at 13 locations. Thus the quasi-random channel is hard to apply. We were only able to place kiosks at 10 % of the randomly chosen islands. However, a look at the map of SL shows that the dropped kiosks were spread around all the populated territories of SL (figure 3). There is a concentration of contacted places and dropped kiosks on the eastern site but this is not a bias because the eastern part of SL is much more "populated" than the western site.

In conclusion, the survey resulted in the largest survey ever done of SL (N=2127). Also, the random protocol resulted in the largest (however in fact not huge) quasi-random sample of SL residence ever collected (N=79). For more details see section 6 below.

5.2 Technical and Interference Events

As mentioned above, though virtual world research should be built on a solid grounding of existing research, there are challenges that are unique to the environment. Each research environment has its own unique challenges that need to be understood as work in those environments continues. In virtual worlds, research instruments may be exposed to hackers or technological problems that might not be predicted at first. This section covers how these events were seen in the survey research with the VDCI.

Hacking is manipulating a digital environment to your advantage or entertainment and usually at odds with the intentions of the environments creators. Hacking can be destructive or creative. For example, using a program to find the user names and passwords of bank customers is an example of hacking. This phenomenon has begun to come under the interest of researchers (Consolvo, 2007)

When paying in SL, the money given to another avatar must come from a source avatar. In the case of the VDCI survey, there was a "research avatar" (Figure 2) whose account needed to have sufficient funds available to pay surveys. This avatar did not need to be logged in to pay but had to have sufficient funds associated with its account. Early in the second week of the VDCI survey (02/17/09) it was discovered the research avatar had no funds in their account. The account funds were replenished. Less than 12 hours later, the funds were once again gone. So we consulted the SL transaction record to discover a small number of individuals had been paid multiple times to do the survey. A small group of individuals had figured out away to hack the system and drain the research account of funds.

We closed the research island and picked up all the randomly placed kiosks. Over the next few hours it was determined the hackers had taken advantage of a programming flaw [4] to gain payment multiple times. This was not a simple flaw. To get multiple payments, the hackers would have had to spent hours figuring out the exploit and then doing the actions took a substantial period of time. In the end, the hackers had worked hours to receive a few meager dollars. The best guess to the motive of the hackers is the utility more in gaming the system, than the money received. The programming flaw was corrected and the survey resumed. Any surveying in a programmatic environment, like SL, are susceptible to hacking and given a incentive, such as with this study, the likelihood of attempted hacking is increased.

Investigating the individuals who hacked the survey, it was discovered through looking at the profiles of the avatars that had been paid multiple times identified themselves as users who are living in a certain country of residence in the real world or were part of the respective country of residence based user group in SL. Later conversations with these and other individuals we learned they people who had exploited the survey were part of an SL organized crime group. It is doubtful these individuals were actually part of an offline organized crime group but more a SL created group of mischievous users who "grief" in SL. Griefing is malicious misuse of the systems in SL for the benefit or pleasure of the griefer. When the problem was fixed the individuals returned and again tried to hack the system but were unable to and contacted the researchers saying the survey was not working.

Also, when working in a programmatic environment that is not fully controlled by the researcher problems out of their control can arise. Commonly, as in the VDCI survey, the environment is not in complete control of the environment. The servers that run the SL code are susceptible to slow downs, errors and being inaccessible. With the VDCI survey this resulted in a race condition in the payment process. A race condition is when multiple programming events conflict to change a variable and results in the variable being set incorrectly at the incorrect time. So for instance, with the VDCI if the user clicked the payment box multiple times, too quickly, it resulted in the variable for payment being set to "paid" before they were actually paid. This, of course, was not preferable to the respondent but all payment discrepancies were solved whenever a non-satisfied respondent contacted the research avatar.

In some cases the channel by which the respondent was contacted could not be recorded. This was again due to the race conditions described above; unavoidable lags in network speeds or

processor speeds meant that, on occasion, the database did not record whether this was a person responding due to a classified ad, an email, or encounter with a random kiosk. This will be referred to as “no record” in the channel variable.

6. Basic Results

6.1 Cleaning Data

Once all the data was collected, the names of each avatar were removed and all the analyses were then done on a file with no personal identifiers. From the total number of surveys started (N=2181), the surveys that did not give consent (N=54) were removed. The non-consent surveys broke down by channel as follows: Classified ad N=19, Email N=2, Quasi-random N=11, and No record N=22). The number of surveys where consent was given was thus N=2127. The non-consent among Quasi-random surveys is higher because the consent form notecard was the first information this group received about the survey (or even what the kiosk was for). Respondents from the other two channels would have read an email or classified ad, and then voluntarily gone to a kiosk, before encountering the notecard.

We deleted not only the data from the people who gave no consent, but we removed also the names of the avatars who did gave consent to become re-contacted (by saying yes to question 36 which asked for permission to store the name for re-contact). Thus we deleted the names of avatars of 387 ("No" N=318, No entry N=69) respondents and we keep 1707. Even for the names we keep, the names are stored in another database than the responses.

For all our analyses we exclude 33 observations who marked for all questions the first option in the questionnaire (Appendix E) We believe that such a consistent and unlikely pattern is in fact an indicator for non-serious response behavior. It is suspected that respondents who displayed this strange behavior belong most likely to the group of "SL criminals" we discussed above. Because the number of those non-serious respondents is so small (33 out of 2127) the impact on the overall results is not significant, in either the descriptive figures or in regression results. So we display results only based on the valid responses (N=2094). They are:

- 151 respondents with unknown channel (no record in the channel variable),
- 325 respondents based on the e-mail-survey,
- 1543 respondents based on classified ad, and

- 75 respondents based on the quasi-random protocol.

However, due to missing values for the different questions in most of our analyses the number of cases is smaller than 2094. Missing values are common in all surveys. In fact, the share of missing values is small for most of our questions. For example, five of the valid respondents did not give us their country of residence. So, in table 3 the total number of cases only is 2089.

6.2 Basic Results

For this survey the three sampling methods resulted in the majority of respondents coming from the classified ad method (N=1543 or 73.7%) and email respondents second largest (N=325 or 15.5%) with quasi-random respondents being small (N=75 or 3.6%) (See Table 1). The respondents who did not have a sample variable set were small (N=157 or 7.2%).

For the e-mail-list method we can calculate a response rate. It is 325 / (5400+1990+4500) = 2.73 % This is a low response rate. However such low rates are not unknown for “mail surveys” in the real world which provide no direct interaction between the “survey provider” and the “survey taker”. This kind of survey never has response rates above 20 percent.

We can calculate to a quasi-response rate for all players within SL who were active in SL during the time when the survey took place. According to Linden Lab, from 02/03/09 to 03/05/09 about 1.1 million avatars were active in SL. So, for the overall sample the quasi-response rate is 2094 / 1100000 = 0.2 %. For the random location sample we can calculate an almost clean response rate, which is 75 / 1100000 = 0.006 %. This is a small participation rate indeed, but it is to be noted that random samples of the entire US adult population – more than 200 million people - often have no more than 1,000 respondents. Results from such surveys are of course valid, because the absolute number of cases is what determines the statistical power of a survey. In most circumstances involving human population, having 2,000 respondents offers sufficient power.

6.3 Participation over Time

Table 2 displays the number of responses collected each day of the elapsed time. Technical and interference events described in Section 5 above appear in the time pattern of response. The race condition issue (different speeds of the network and the computers at either end) would occasionally make the survey type appear as a blank in the data. These “blank types” appeared predominantly at the beginning (50% or the blanks (N=76) appeared in the first 9 days of the survey). Fourteen days

from the start of the survey it is clearly observable when the survey was down to the criminal attack (described in section 5.2 above) (half of 02/17/09, full day of 02/18/09 and some hours of (02/19/09).

The criminal attack gave us – by accident – an opportunity to run a "natural experiment" about the fieldwork strategy in SL. Due to the act-of-crime we must close the kiosks for two half and one full day. During this time we removed the class ad too. So when we opened the kiosk again we placed a new class ad in the Employment section of SL. The result was an increased number of respondents after the new placement of the class ad (more than 100 respondents in the days of re-opening compared to about 30 ones in the days before the closure of the kiosks; see table 2) So we might speculate that, a survey researcher might be able to stimulate the respondent rates in SL by a "rotating" scheme of new respectively quasi-new class ads during the fieldwork duration of a survey in SL.

6.4 Impact Of The Three Sampling Modes On The Structure Of Subsamples

In this section we analyze the possible effects of the three modes (respectively four subgroups, including respondents with unknown modes) on the demographics of the resulting subsamples. We expect that the respondents with no record in the channel variable are not significantly different from the respondents with known channels of recruiting. However, because the blanks occurred more at the beginning of the survey it is likely that the group of "blanks" is more similar to the e-mail group than to the others (especially to the quasi-random group which was started later than the e-mail and class ad groups).

The central question is if the respondents who came via e-mail and classified ad are different from the ones who came via the random location protocol. This is an important question because the quasi-random location protocol is the best one in terms of survey theory. However it is expensive (and furthermore only a quasi-random protocol). So the question arises, if it is worthwhile to run the expensive quasi-random mode, or if the classified-ad method delivers almost the same structures of respondents. Due to the high self-selection in favor of special interests it is most likely that the e-mail method delivers different demographics.

SL Land Ownership. Avatars can own land in Second Life. An avatar can own a small parcel of land or an entire island. The survey results show (Table 3) that 32.6% of the avatars (N=2089) who took the survey were land owners. This was almost constant across the classified ad and quasi-

random survey types. The numbers were slightly different with surveys collected from the email respondents who may have slightly higher ownership due to need for land for educational or research installations. However, respondents from the class ad and the quasi-random groups show almost no difference.

SL Avatar Gender. An avatar has a gender also. This gender is set in two ways. First, known only to the owner of the avatar, the avatar has a male or female setting in its Appearance settings. This setting is not visible to other avatars or users and is not definitive for gender because a female setting can be changed to look like a male person. Also, the individual may choose one gender and may choose another's appearance or an un-gendered appearance. The results show that slightly more avatars are female than male (40.0% Male vs. 52.2% Female N=1932) (See Table 4). With Transgendered, Other and No Answer all less than 3%. There was no significant difference among the survey types.

Conversely, World of Warcraft (WoW) the most popular Massively-multiplayer Online Role Playing Game (MMORPG) has a very different avatar gender make up (65% male vs. 35% female (Yee, 2005).

RL Gender. The owner of each avatar also has a gender. This can be the same or different than the avatar's gender. Our analysis shows in SL, the majority of users are female (51.4 % female vs. 43% male, N=1967). The survey type had no significant association with RL gender (p=.216).

Multiple Avatars. A Second Life user can have more than one avatar. All a user needs is another email address to use for the second or more avatar. This second avatar is called an "alternative" or simply "alt". Users with multiple avatars is reported to inflate SL population statistics (Terdiman, 2007; Boelstorff, 2008). The survey shows 52% of respondents report not having another avatar (See Table 5) This level was almost constant across the classified ad and quasi-random survey types. The numbers were slightly different with surveys collected from the email respondents who may have slightly lower due to need retain a stable identity for educational or research activities. However, respondents from the class ad and the quasi-random groups show almost no difference.

Country of Residence. Each respondent was asked for their country of residence based on a list of the top 10 countries based on a list provided by Reuters (2007) (See Table 6). The results show country of residence along these 10 countries with the USA, UK and Canada with slightly higher percentages. This is most likely due to the survey being in English and so English speaking countries being slightly higher in the representation (See Table 7). Also, there is a higher percentage

of German residents in the random sample group. None of the random locations had any particular German connection but the DIW Berlin logo displayed on the kiosk may have attracted the attention of more German respondents, because “DIW Berlin” is a well known institute in Germany.

RL Age. Each respondent was asked for their real life age in 5 discrete categories. The results show in total 67% of respondents were 35 or younger (See Table 8). When looking at the survey type, the quasi-random sample showed a more spread out age representation (50% of respondents 35 and under N=37). With more random respondents this distribution may have began to look more like the other survey types.

RL Income. Each respondent was asked for their real life income in 8 discrete categories. The results show in total 55% of respondents had an annual income of 20,000 USD or less (See Table 9). When looking at the survey type, the quasi-random sample showed a more spread out income representation (37% of respondents 20,000 and under N=27). With more random respondents this distribution may have began to look more like the other survey types.

RL Education Level. It is difficult in the international context of SL to capture the level of education of respondents. As a proxy, we asked respondents to give the age of their final year of education. The results show over 80% of all respondents (N=1691) (See Table 10) had finished education at the age 35 or less. The email sample had less (75% who represents N=238) respondents finishing education under the age of 35 but since they were predominantly academics and researchers a commitment to lifelong education may be expected.

Permission for a re-contact. At the very end of the questionnaire the respondents were asked if they would agree to be re-contacted (for the exact wording of the question see Appendix E, question 36). The good message is: about 80 % of the respondents are willing to get re-contacted. And this is a share which is in the range of such a willingness in real world surveys. However, a closer look on the structure of the willingness reveals remarkable (and plausible) differences (see table 11).

The willingness to go into a panel is about twice as high in the class ad-group of respondents than in the quasi-random group (90 % vs. 45 %). The e-mail-group is very similar to the quasi-random group. And the blank group is – as we know – a mixture of the different groups. Thus the blanks display a willingness which is close to the average willingness. The large differences between the three survey modes are highly plausible because the randomly picked respondents and the ones, who were asked via e-mail-lists, were by definition less selective than the ones who respondent to the class ad. A lot of the latter probably participated in the survey only because they

wanted to earn some money. In this case, it is likely that the want to repeat this "job".

A logistic regression analysis (not displayed) shows that in fact the different survey modes – and thus the different kinds of self selection of respondents – are the most important factors which explain differences in the willingness to participate in a panel survey. There is no difference between female and male respondents. However, those who do not display their gender are less willing which is no surprise to get re-contacted. Land owners (within SL) have a slightly higher willingness to participate in a panel study. This is no surprise either because house owners in real life (who have more stable lives than other people) are participating in panels more frequently than non-owners. The same might by true for land owners in SL, who most likely are more stable visitors in SL than other players. However, the difference in willingness to participate in an "SL panel" is not statistical significant between land owners and others.

There are two significant effects for the countries of residence. Again it is no surprise that those respondents who were not willing to display their country of residence are less willing to participate in a panel study. Much harder to explain is the lower willingness for German respondents (by controlling for thee survey mode). In fact we do not have an explanation.

Thus the differences of giving permission for a longitudinal survey by personal characteristics of the respondents are rather small. Whereas the differences are huge with respect to survey mode. But this will be no problem for a panel study because we can calculate longitudinal weighting factors based on the first-wave-characteristics (Spiess, 2009).

Conclusion: Random versus Classified Ads. As for the main question, of whether the quasi-random sample and the much cheaper classified ad sample are the same, the results are tentatively in favor. For most demographic categories, the two samples seem the same. There are some areas of difference as well, but these seem small. Thus, the evidence seems to support an argument that classified-ad sampling obtains a representative sample of the SL population.

7. One Example for Substantial Analysis: Gender Bending

The data collected in this test run of the VDCI allow a number of substantial analysis. Due to space considerations, only one issue is considered: the sex of the avatar that a user creates.

Almost all virtual worlds allow the user to create an avatar. This avatar represents the user in the virtual world. The avatar created by the user is in no way related to the physical characteristics of

the user within the limits of the system. For example, a 24 year old woman in Indiana could create an avatar in SL that was a 74 year old male Asian body builder. This is especially true in the area of gender and gender-bending. In the real world, gender-bending happens when a person of one sex wears clothes, male-up or even has surgical procedures to appear to be or become a member of the opposite sex. In the virtual world, when a user of one gender chooses an avatar of another gender, they can be said to be "gender-bending".

Existing work in this area suggests a hypothesis that male players (= males in RL) will more often choose a female avatar than females (= females in RL) will choose male avatars. One reason is because male players have a stronger incentive to switch over, because attractive female avatars appear to have a greater likelihood to make friends or receive help than a male avatar (Hussain & Griffiths, 2008). Gender-bending in game virtual worlds such as *World of Warcraft* (WoW) and *EverQuest* (Yee, 2003; From et al., 2007) seems to support this hypothesis.

Our survey collected some initial statistics around gender-bending including asking if the user had created an avatar of the opposite gender (72% or respondents (N=1511) reported not adopting the opposite gender for their avatar). Using the same model as Yee (2005), 11% of RL males (N=98) used an opposite gender SL avatar compared with 5% of RL females (N=58). Therefore, in SL, male users are twice as likely to gender-bend than female user, in contrast to 7-8 times more in WoW (Yee, 2005). Continuing Yee's model, given a hypothetical pool of 1000 users 514 would be female and 430 would be male. Of the 514 female users, 25.7 would be male avatars and of the 430 male users, 47.3 would be female avatars. This is vastly different than the WoW statistics Yee found. More comparisons of the gender make-up of virtual worlds need to be gathered and compared.

The gender of an avatar is not static. It can be changed back and forth at any time in Second Life. The results show that only 23.2% of avatars do this (N=486) (See Table 4) . The percentage is slightly higher amongst email respondents (31%) perhaps because educators and researchers promote and demonstrate experimentation. Also, SL gender is positively associated ($p<0.0001$) with change of avatar gender. This means a male avatar survey respondent is more likely to have changed avatar gender, at some point in the past, than a female avatar respondent (50% male vs. 37% female) regardless of survey mode.

8. Conclusions

Surveying the virtual world is a relatively new research method and the state of the art is to send virtual world respondents outside the virtual world, to a web-based survey. The VDCI (Virtual Data Collection Interface) introduced here allows survey research to become completely immersed in the virtual world. This tool also allows automated data collection and respondent payment that allowed the survey to run 24 hours a day with little researcher involvement. This resulted in three different sample protocols and one of the largest group of valid respondents (N=2094) surveyed in the virtual world.

When doing such a survey the expertise of both researcher and respondents must be taken into account. The VDCI creation required extensive SL, LSL, PHP and MySQL knowledge that may not be available to all researchers. If other researchers are entering the field of virtual world research some knowledge of these skills would be invaluable. Second, when creating a tool like the VDCI it is important to assess the skill level of the respondents also. It cannot be assumed all virtual world residents have the same level of skills. With the survey the instruction note card included detailed instructions of how to put the HUD on and how to answer the questions. Having a number of methods of contacting the researcher (IM or email) the respondent could have technical questions answered quickly.

Our main methodological result may be the fact that while there are differences in the demographic structure of respondents who responded to the three different survey modes we applied (e-mail lists, class ad, and a quasi-random protocol), *the differences between the superior quasi-random protocol and the classified ad mode are rather small*. So we believe that we have evidence that the easy-to-handle (and thus economically efficient) classified ad-mode delivers sufficiently valid results about the population of SL players/users.

Because we wanted to get large numbers of respondents we paid a small honorarium (250 Linden Dollars = about 1 $US) to all respondents. So we were not able to test the impact of different incentive schemes (for example reaching from “nothing” to a far higher amount than 250 Lindens). An experimental study of different incentive schemes would be a good avenue for further research.

By accident, we ran a “natural experiment” about the fieldwork strategy in SL. Due to an act-of-crime (see section 5.2) we were forced to close the kiosks for two half days and one full day. During this time we removed the classified ad too. When we opened the kiosk again we placed a new

classified in the Employment section of SL. The result was an increased number of respondents after the new placement of the class ad. So we might speculate that a survey researcher might be able to stimulate the response rates in SL by a “rotating” scheme of new or quasi-new class ads during the fieldwork duration of a survey in SL.

Overall, the successful fielding of a fully-immersive survey within a virtual world confirms that the virtual world population can be treated as a “real” population for survey research. Their “country” has been identified, and tools and methods have been created for obtaining representative random samples from them. As the virtual world population grows, the future will likely see larger-scale surveys and more general conclusions about who these people are, how they feel, and what they do.

References:

Bell, Mark W., Castronova, Edward and Wagner, Gert G. (2008). Virtual Assisted Self Interviewing (VASI): An Expansion of Survey Data Collection Methods to the Virtual Worlds by Means of VDCI. Berlin DIW Berlin Data Documentation No. 37. Available at http://www.diw.de/documents/publikationen/73/88367/diw_datadoc_2008-037.pdf

Bell, Mark. W. and Robbins-Bell, Sarah (2008). Toward an Expanded Definition of “Virtual Worlds". New Digital Media: Audiovisual, Games and Music. Rio de Janeiro: e-papers.

Boellstorff, T. (2008). Coming of age in second life: An anthropologist explores the virtual human. Princeton, NJ: Princeton University Press.

Bystrom, K.-E., Barfield, W., & Hendrix, C. (1999). A Conceputal Model of the Sense of Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments, 8(2), 241-244.

Cassell, J. (Apr. 2000). More than just another pretty face: Embodied conversational interface agents. Communication. ACM 43, 4 70–78.

Castronova, E. (2004) Synthetic worlds. Chicago: The University of Chicago Press.

Castronova, E.; & Falk, M. (2008). Virtual Worlds as Petri Dishes for the Social and Behavioural Sciences. Berlin: RatSWD Working Paper No. 47. Available at http://www.ratswd.de/download/workingpapers2008/47_08.pdf.

Childress, M., & Braswell, R. (2006). Using massively multiplayer online role-playing games for online learning. Distance Education, 27(2), 187-196.

Consalvo, M. (2007) Cheating: Gaining Advantage in Videogames. Cambridge: MIT Press.

Friedman, D., Steed, A., and Slater, M.. 2007. Spatial social behavior in second life. Intelligent Virtual Agents, 4722(2007), 252-263.

Fron, J., Fullerton, T., Ford Morie, J. and Pearce, C. (2007).Playing dress-up: costumes, roleplay and imagination. Proceedings of Women in Games. University of Wales, Newport, U.K,

Griffiths, M. D., Davies, M. N., & Chappell, D. (2003). Breaking the Stereotype: The Case of Online Gaming. CyberPsychology & Behavior , 6 (1), 81-91.

Hussain, Z. & Griffiths, M.D. (2008). Gender swapping and socializing in cyberspace: An exploratory study. CyberPsychology & Behavior, 2008, 11, 47-53.

Jennings, N. & Collins, C. (2008). Virtual or Virtually U: Educational Institutions in Second Life, International Journal of Social Sciences, 2(3).

Kemp, J., & Livingstone, D. (2006). Putting a Second Life “Metaverse” skin on learning management systems. Proceedings of the Second Life Education Workshop at the Second Life Community Convention (pp. 13-18). San Francisco: The University Of Paisley.

Kuntze, M. F., Stoermer, R., Mueller-Spahn F., & Bullinger, A. H. (2002). CyberPsychology & Behavior, 5(3): 203-206.

Schimmack, U., Krause, P., Wagner, G. G., & Schupp, J., (2009). Stability and Change of Well Being: An Experimentally Enhanced Latent State-Trait-Error Analysis. Social Indicator Research (forthcoming).

Seay, A. F., Jerome, W. J., Sang Lee, K., & Kraut, R. E. (2004). Project Massive: A study of online gaming communities. CHI '04 extended abstracts on Human factors in computing systems (pp. 1421 - 1424). Vienna, Austria : ACM.

Siedler, T., Schupp, J., Spiess, C. K., & Wagner, G. G. (2009). The German Socio-Economic Panel (SOEP) as Reference Data Set. Schmollers Jahrbuch, 129(2) (forthcoming).

Spiess, Martin (2009), Compensating for missing data in the SOEP. DIW Data Documentation No. 41: DIW Berlin.

Terdiman , D.(2007, January 3). Counting the real 'Second Life' population. cnet news. Retrieved May 26, 2009, from http://news.cnet.com/Counting-the-real-Second-Life-population/2100-1043_3-6146943.html

Ting-Jui Chou, Chih-Chen Ting (2003). The Role of Flow Experience in Cyber-Game Addiction. CyberPsychology & Behavior., 6(6): 663-675.

Williams, D. (2006). On and Off the 'Net: Scales for Social Capital in an Online Era. Journal of Computer-Mediated Communication, 11.

Yee, N. (2006).The demographics, motivations, and derived experiences of users of massively-multiuser online graphical environments. PRESENCE: Teleoperators and Virtual Environments, 15:309–329.

Yee, N. (2006a). The Demographics, Motivations and Derived Experiences of Users of Massively-Multiuser Online Graphical Environments. PRESENCE: Teleoperators and Virtual Environments, 15: 309-329.

Yee, N. (2006b). Motivations for Play in Online Games. CyberPsychology & Behavior, 9 (6), 772-775.

Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & Merget, D. (2007). The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. CyberPsychology & Behavior, 10, 115-121.

Yee, N. (2003, September 3) The Demographics of Gender-Bending. the daedalus project. Retreived May 26, 2009, from http://www.nickyee.com/daedalus/archives/000551.php?page=1

Yee, N. (2005, July 28). WoW Gender-Bending. the daedalus project. Retreived May 26, 2009, from http://www.nickyee.com/daedalus/archives/001369.php

Yellowlees, P. M., & Cook, J. N. (2006). Education About Hallucinations Using an Internet Virtual Reality System: A Qualitative Survey. Academic Psychiatry,30, 534-539.

Appendix A

Note Card Containing Instructions And Consent Information

IRB Study #08-13679
INDIANA UNIVERSITY BLOOMINGTON
INFORMED CONSENT STATEMENT

You are invited to participate in a research study of demographics and quality of life of virtual world residents. . You were selected as a possible respondent because you have clicked on a survey kiosk. We ask that you read this form and ask any questions you may have before agreeing to be in the study.

The study is being conducted by Edward Castronova (Indiana University), Mark W. Bell (Indiana University) Gert G Wagner (Berlin University of Technology). It is funded by the German Socio-economic Panel Study (SOEP)

STUDY PURPOSE

The purpose of this study is to gain demographic and quality of life information from the residents of virtual worlds to compare them to real world residents.

PAYMENT

You will receive payment for taking part in this study. When accept to take part in the survey you will be eligible to be paid $250 linden dollars. Click on the payment box ONCE at the kiosk to receive your payment.

NUMBER OF PEOPLE TAKING PART IN THE STUDY:

If you agree to participate, you will be one of 50-10,000 respondents who will be participating in this research.

PROCEDURES FOR THE STUDY:

If you agree to be in the study, you will do the following:

Attach VDCI HUD

1. Search your inventory for the "HUDSVY01-2.1CLASS".
2. Right click on the object and choose “Wear”.
3. Confirm you have read this material.
4. Answer the survey questions. Depending on lag it may take some time to rez the survey questions.

RISKS OF TAKING PART IN THE STUDY:

While on the study, the risks are:

Possible verbal harassment from other residents beyond the control of the researchers.

BENEFITS OF TAKING PART IN THE STUDY:

There is no direct benefit to the respondent.

ALTERNATIVES TO TAKING PART IN THE STUDY:

Instead of being in the study, you have the option to choose not to participate. Simply, close this notecard and do not attach the Survey HUD.

CONFIDENTIALITY

Efforts will be made to keep your personal information confidential. In any case, we will only record your avatar name, not your personal data. We cannot guarantee absolute confidentiality, because information about your avatar might be disclosed if required by law. However, your avatar identity will be held in confidence in reports in which the study may be published and databases in which results may be stored.

Organizations that may inspect and/or copy the survey data (which do not contain your avatars name) for quality assurance and data analysis include groups such as the study investigator and his/her research associates, the IUB Institutional Review Board or its designees, the study sponsor, and (as allowed by law) state or federal agencies.

COSTS

There is no cost to the you.

CONTACTS FOR QUESTIONS OR PROBLEMS

For questions about the study or a research-related injury, contact the researcher Mark W. Bell at 317.XXX.XXXX.

For questions about your rights as a research respondent or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IUB Human Subjects office, 530 E Kirkwood Ave, Carmichael Center, L03, Bloomington IN 47408, 812-855-3067 or by email at iub_hsc@indiana.edu

VOLUNTARY NATURE OF STUDY

Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty or loss of benefits to which you are entitled. Your decision whether or not to participate in this study will not affect your current or future relations with the investigator(s).

Thanks a lot for you cooperation.

Appendix B

Email Contact Letter

Subject: Respondents required for survey research - 250 L$

Hey folks! The Synthetic Worlds Initiative (Edward Castronova's group) at Indiana University is rolling out an in-world survey tool. Our first survey is on demographics and the quality of life for virtual world residents.

Respondents will be paid L$ 250.

All you need to do is take a survey from a kiosk located on our island: http://slurl.com/secondlife/Synthetic%20World/127/130/32

Thanks for your willingness to help out a fellow educator/researcher.
Mark
--
Mark Bell
PhD student in Indiana University's Telecommunications program
SL: Typewriter Tackleberry
http://www.indiana.edu/~telecom/
http://swi.indiana.edu/
http://blog.markwbell.com
"Communications tools don't get socially interesting until they get technologically boring." - Clay Shirky

Appendix C

Classified Ad

Subject: 250 L$ for taking a survey

The Synthetic Worlds Initiative at Indiana University is rolling out an in-world survey tool. Our first survey is on demographics and the quality of life for virtual world residents.
Respondents will be paid L$ 250.
All you need to do is take a survey from a kiosk located on our island: http://slurl.com/secondlife/Synthetic%20World/127/130/32

Appendix D

Quasi-randomly Dropped Kiosks

Island name	slurl	Date Contacted	Kiosk placement
Hinode Shima	http://slurl.com/secondlife/Hinode%20Shima/209/59/26	2/3/2009	http://slurl.com/secondlife/Hinode%20Shima/105/125/22
Meola	http://slurl.com/secondlife/Meola/128/128/2	2/5/2009	http://slurl.com/secondlife/Dauphin/21/138/101
Spinolds Flat	http://slurl.com/secondlife/Spinolds%20Flat/220/169/58	2/6/2009	http://slurl.com/secondlife/Spinolds%20Flat/177/192/58
Tamaqua	http://slurl.com/secondlife/Tamaqua/85/128/24	2/11/2009	http://slurl.com/secondlife/Tamaqua/147/115/23
Ebusiness	http://slurl.com/secondlife/Ebusiness/56/124/33	2/12/2009	http://slurl.com/secondlife/Ebusiness/199/120/27
Mowry	http://slurl.com/secondlife/Mowry/56/124/32	2/12/2009	http://slurl.com/secondlife/Helvellyn/45/218/31
Horisme	http://slurl.com/secondlife/Horisme/90/251/101	2/12/2009	http://slurl.com/secondlife/Horisme/225/166/102
Rosewood	http://slurl.com/secondlife/Rosewood/56/128/32	2/21/2009	http://slurl.com/secondlife/Cybertopia%20North/189/173/26
Kandor	http://slurl.com/secondlife/Kandor/206/41/30	2/21/2009	http://slurl.com/secondlife/Kandor/206/41/30
The Sardar Passage	http://slurl.com/secondlife/The%20Sardar%20Passage/248/116/23	2/21/2009	http://slurl.com/secondlife/Big%20Serenity/190/251/25
Nijole	http://slurl.com/secondlife/Nijole/144/124/107	2/21/2009	http://slurl.com/secondlife/Nijole/144/124/107
Kaolin	http://slurl.com/secondlife/Kaolin/32/49/26	2/12/2009	http://slurl.com/secondlife/Kaolin/32/49/26
Sardar Forest	http://slurl.com/secondlife/Sardar%20Forest/69/72/436	2/21/2009	http://slurl.com/secondlife/Sardar%20Forest/69/72/436

Appendix E

Survey Questions

		Answer										
#	Question	1	2	3	4	5	6	7	8	9	10	11
1.	Have you read the pretest notecard and wish to take part in the survey?	Yes	No									
2.	Do You own land in Second Life?	Yes	No									
4.	What is the gender of this avatar?	Male	Female	Trans-gendered	Other	No Answer						
5.	Have you ever change the gender of this avatar?	Yes	No	No Answer								
6.	Do you have a second avatar?	Yes	No	No Answer								
7.	What is the gender of the second avatar?	Male	Female	Trans-gendered	Other	No Answer						
8.	Is the present avatar your main one?	Yes	No	No Answer								

9.	What is your country of Residence	United States	Germany	France	United King-dom	Nether-lands	Spain	Brazil	Canada	Belgium	Italy	Other
10.	What is your gender?	Male	Female	Transgen-dered	No Answer							
11.	What is your current age?	18-25	26-35	36-45	46-55	56+						
12.	Approxi-mately how many hours per week do you use the computer for any purpose?	0-10	11-20	21 - 30	31-40	41+						
13.	Approxi-mately how many hours per week do you spend in "Second Life"?	0-10	11-20	22 - 30	31-40	41+						
14.	Approxi-mately how many hours per week do you play video games?	0-10	11-20	23 - 30	31-40	41+						
15.	Approxi-mately how many hours per week do you play massively multi-player online games (For example World of Warcraft)?	0-10	11-20	24 - 30	31-40	41+						

16.	Approxi-mately how many hours per week do you use online social spaces (i.e., MMOGs, MUDs, MySpace, etc.)?	0-10	11-20	25 - 30	31-40	41+						
17.	At what age did you stop college/ university education ?	18-25	26-35	36-45	46-55	56+						
18.	What is your current employment status?	student	Employer	Employee	Unemployed	Non-Employed	Other					
19.	What is your current yearly income?	$0-$10,000	$10,001-$20,000	$20,001-$50,000	$50,001-$75,000	$75,001-$100,000	$100,001-$150,000	$150,001-$200,000	$200,001 or more			
35.	How would you rate this survey experi-ence?	Completely satisfied / 10	9	8	7	6	5	4	3	2	1	Completely dissatisfied / 0
36.	Are you interested in being contacted for further surveys?	Yes	No	No Answer								

Table 1

Response By Survey Type

Survey Type	N	Percentage
No record	151	7.2%
Class	1543	73.7%
Email	325	15.5%
Quasi-random	75	3.6%
Total	2094	100%

Table 2

Surveys Collected Each Day By Sample Method

	Survey Type												
Date	No record	%	Cumulative %	Class	%	Cumulative %	Email	%	Cumulative %	Quasi-random	%	Cumulative %	Total
2/3/2009	11	7%		14	1%		33	10%		0	0%		58
2/4/2009	15	10%	17%	22	1%	2%	24	7%	18%	0	0%	0%	61
2/5/2009	13	9%	26%	49	3%	6%	13	4%	22%	1	1%	1%	76
2/6/2009	7	5%	30%	26	2%	7%	3	1%	22%	3	4%	5%	39
2/7/2009	7	5%	35%	48	3%	10%	4	1%	24%	3	4%	9%	62
2/8/2009	3	2%	37%	18	1%	11%	35	11%	34%	7	9%	19%	63
2/9/2009	6	4%	41%	16	1%	13%	2	1%	35%	5	7%	25%	29
2/10/2009	9	6%	47%	26	2%	14%	7	2%	37%	1	1%	27%	43
2/11/2009	5	3%	50%	16	1%	15%	6	2%	39%	8	11%	37%	35
2/12/2009	12	8%	58%	8	1%	16%	8	2%	42%	11	15%	52%	39
2/13/2009	9	6%	64%	27	2%	17%	25	8%	49%	2	3%	55%	63
2/14/2009	9	6%	70%	35	2%	20%	8	2%	52%	5	7%	61%	57
2/15/2009	6	4%	74%	33	2%	22%	21	6%	58%	4	5%	67%	64
2/16/2009	3	2%	76%	27	2%	24%	12	4%	62%	1	1%	68%	43
2/17/2009	1	1%	77%	9	1%	24%	1	0%	62%	0	0%	68%	11
2/18/2009	0	0%	77%	0	0%	24%	0	0%	62%	0	0%	68%	0

2/19/2009	2	1%	78%	131	8%	33%	14	4%	66%	2	3%	71%	149
2/20/2009	3	2%	80%	106	7%	40%	13	4%	70%	0	0%	71%	122
2/21/2009	4	3%	83%	123	8%	48%	20	6%	77%	0	0%	71%	147
2/22/2009	5	3%	86%	73	5%	52%	14	4%	81%	1	1%	72%	93
2/23/2009	2	1%	87%	72	5%	57%	9	3%	84%	1	1%	73%	84
2/24/2009	1	1%	88%	51	3%	60%	5	2%	85%	0	0%	73%	57
2/25/2009	2	1%	89%	50	3%	64%	3	1%	86%	1	1%	75%	56
2/26/2009	2	1%	91%	64	4%	68%	6	2%	88%	2	3%	77%	74

Survey Type

Date	No record	%	Cumulative %	Class	%	Cumulative %	Email	%	Cumulative %	Quasi-random	%	Cumulative %	Total
2/27/2009	0	0%	91%	63	4%	72%	6	2%	90%	0	0%	77%	69
2/28/2009	1	1%	91%	83	5%	77%	9	3%	93%	0	0%	77%	93
3/1/2009	2	1%	93%	69	4%	82%	9	3%	95%	0	0%	77%	80
3/2/2009	2	1%	94%	80	5%	87%	6	2%	97%	7	9%	87%	95
3/3/2009	3	2%	96%	88	6%	92%	4	1%	98%	6	8%	95%	101
3/4/2009	3	2%	98%	95	6%	99%	5	2%	100%	4	5%	100%	107
3/5/2009 (closing time: 9am EST)	3	2%	100%	21	1%	100%	0	0%	100%	0	0%	100%	24
Total	151			1543			325			75			2094

Table 3

SL Land Ownership By Sample Method

	Survey Type									
Land Ownership	No record	%	Class	%	Email	%	Quasi-random	%	Total	%
Yes	44	29%	469	30%	141	44%	28	37%	682	37%
No	106	71%	1071	70%	183	56%	47	63%	1407	63%
Total	150		1540		324		75		2089	

Table 4

SL Gender By Sample Method

	Survey Type									
SL Gender	No record	%	Class	%	Email	%	Quasi-random	%	Total	%
Male	67	44%	619	40%	124	38%	28	38%	838	40.0%
Female	73	48%	812	53%	171	53%	38	51%	1094	52.2%
Transgendered	3	2%	38	2%	12	4%	0	0%	53	2.5%
Other	8	5%	32	2%	13	4%	2	3%	55	2.6%
No record	0	0%	40	3%	5	2%	6	8%	51	2.4%

Table 5

Second Avatar By Sample Method

	Survey Type									
Second avatar	No record	%	Class	%	Email	%	Quasi-random	%	Total	%
Yes	82	54%	631	41%	169	52%	42	57%	904	43%
No	84	56%	847	55%	136	42%	26	35%	1093	52%
NA	5	3%	62	4%	20	6%	6	8%	93	4%
Total	151		1540		325		74		2090	

Table 6

Country of Residence of Respondents Compared to Reuters

Country of Residence	N	%	Reuters	Difference
United States	839	40%	31.19%	9%
Germany	143	7%	12.73%	-6%
France	114	5%	10.46%	-5%
United Kingdom	185	9%	8.09%	1%
Netherlands	54	3%	6.55%	-4%
Spain	62	3%	3.83%	-1%
Brazil	45	2%	3.77%	-2%
Canada	117	6%	3.30%	2%
Belgium	8	0%	2.63%	-2%
Italy	50	2%	1.93%	0%
Other	459	22%		
Total	2076			

Table 7

Country of Residence By Sample Method

	Survey Type									
Country of Residence	No record	%	Class	%	Email	%	Quasi-random	%	Total	%
United States	58	39%	616	40%	132	41%	33	44%	839	40%
Germany	9	6%	91	6%	30	9%	13	17%	143	7%
France	7	5%	89	6%	16	5%	2	3%	114	5%
United Kingdom	17	11%	125	8%	35	11%	8	11%	185	9%
Netherlands	4	3%	42	3%	6	2%	2	3%	54	3%
Spain	3	2%	46	3%	13	4%	0	0%	62	3%
Brazil	2	1%	41	3%	2	1%	0	0%	45	2%
Canada	7	5%	90	6%	15	5%	5	7%	117	6%
Belgium	0	0%	5	0%	3	1%	0	0%	8	0%
Italy	1	1%	37	2%	11	3%	1	1%	50	2%
Other	40	27%	351	23%	57	18%	11	15%	459	22%
Total	148	1	1533	1	320	1	75	1	2076	

Table 8

RL Age By Sample Method

	Survey Type									
RL Age	No record	%	Class	%	Email	%	Quasi-random	%	Total	%
18-25	54	36%	554	36%	106	33%	17	23%	731	35%
26-35	46	31%	490	32%	104	33%	20	27%	660	32%
36-45	23	15%	281	18%	51	16%	21	28%	376	18%
46-55	17	11%	150	10%	35	11%	12	16%	214	10%
56+	9	6%	58	4%	22	7%	5	7%	94	5%
Total	149		1533		318		75		2075	

Table 9

RL Annual Income By Sample Method

	Survey Type									
Annual Income	No record		Class		Email		Quasi-random		Total	
$0-$10,000	51	35%	581	38%	101	32%	16	22%	749	36%
$10,001-$20,000	28	19%	280	18%	63	20%	11	15%	382	19%
$20,001-$50,000	29	20%	314	21%	63	20%	21	29%	427	21%
$50,001-$75,000	18	12%	163	11%	33	11%	12	16%	226	11%
$75,001-$100,000	9	6%	98	6%	24	8%	6	8%	137	7%
$100,001-$150,000	4	3%	42	3%	15	5%	5	7%	66	3%
$150,001-$200,000	1	1%	19	1%	2	1%	1	1%	23	1%
$200,001 or more	6	4%	30	2%	10	3%	1	1%	47	2%
Total	146		1527		311		73		2057	

Table 10

Age Education Stopped By Sample Method

	Survey Type									
Age Education Stopped	No record	%	Class	%	Email	%	Quasi-random	%	Total	%
18-25	77	52%	907	59%	149	47%	42	57%	1175	57%
26-35	39	27%	370	24%	89	28%	18	24%	516	25%
36-45	11	7%	148	10%	46	15%	10	14%	215	10%
46-55	11	7%	60	4%	16	5%	2	3%	89	4%
56+	9	6%	47	3%	15	5%	2	3%	73	4%
Total	147		1532		315		74		2068	

Table 11

Willingness to Be Re-contacted By Sample Method

	Survey Type			
Willingness to Be Re-contacted Stopped	No record	Class	Email	Quasi-random
Yes	31.8%	47.7%	9.3%	54.7%
No	68.2%	52.3%	90.7%	45.3%
Total	100	100	100	100

Figure Caption

Figure 1. Survey Kiosk

Figure Caption

Figure 2. Research Avatar

–

Figure Caption

Figure 3. Map of randomly chosen islands and the dropped kiosks. Successful placements are represented by a circle (O).

End Notes:

---

[1] Due to payment problems created by race conditions, roughly 1.5 % (N= 33) people were not paid immediately. Once they contact the research avatar they were promptly paid.

[2] See Yellowlees and Cook, 2006; Yee, 2006; Jennings and Collins, 2008; Yee and Bailenson, 2007; Friedman, Steed and Slater, 2007.

[3] SLURLS are addresses in SL that allow the user to teleport their avatar to that location. It is like an URL for a web page but has exact co-ordinates in SL.

[4] This programming flaw was caused by the survey complete variable being set back to the default state. This was due to a series of putting on and taking off the kiosk and picking up a new version of the HUD from the kiosk.