



HAL
open science

Virtual world experimentation: An exploratory study

Thomas Chesney, Swee-Hoon Chuah, Robert Hoffmann

► **To cite this version:**

Thomas Chesney, Swee-Hoon Chuah, Robert Hoffmann. Virtual world experimentation: An exploratory study. *Journal of Economic Behavior and Organization*, 2009, 72 (1), pp.618. <10.1016/j.jebo.2009.05.026>. <hal-00707106>

HAL Id: hal-00707106

<https://hal.science/hal-00707106v1>

Submitted on 12 Jun 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Accepted Manuscript

Title: Virtual world experimentation: An exploratory study

Authors: Thomas Chesney, Swee-Hoon Chuah, Robert Hoffmann

PII: S0167-2681(09)00168-1
DOI: doi:10.1016/j.jebo.2009.05.026
Reference: JEBO 2413

To appear in: *Journal of Economic Behavior & Organization*

Received date: 14-4-2008
Revised date: 29-5-2009
Accepted date: 29-5-2009

Please cite this article as: Chesney, T., Virtual world experimentation: An exploratory study, *Journal of Economic Behavior and Organization* (2008), doi:10.1016/j.jebo.2009.05.026

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Virtual world experimentation: An exploratory study

Thomas Chesney, Swee-Hoon Chuah, Robert Hoffmann^{*},

*International Centre for Behavioural Business Research
Nottingham University Business School
Jubilee Campus, Wollaton Road
Nottingham NG8 1BB
United Kingdom.
Tel: +44(0)115 846 66 02
Fax: +44(0)115 846 66 67*

Abstract

We explore the scientific potential of virtual worlds for experimental economics in terms of the subject pools and experimental platforms they present. Our results offer tentative, qualified support for virtual world experimentation. Overall, the behaviour of virtual subjects recruited, incentivised and observed within *Second Life* across a range of five standard experimental games was not found to differ significantly from established standard results. In addition, we identify certain methodological opportunities and challenges which confront virtual world experimenters.

Key words: virtual worlds, experiments

JEL classification: C72; C88; C99; Z13

- 1 Acknowledgements: We are grateful to Simon Gächter, Sourafel Girma, Martin
- 2 Sefton and Chris Starmer for a number of helpful suggestions.

^{*} Corresponding author.

Email addresses: thomas.chesney@nottingham.ac.uk (Thomas Chesney),
swee-hoon.chuah@nottingham.ac.uk (Swee-Hoon Chuah),
robert.hoffmann@nottingham.ac.uk (Robert Hoffmann).

Virtual world experimentation: An exploratory study

Abstract

We explore the scientific potential of virtual worlds for experimental economics in terms of the subject pools and experimental platforms they present. Our results offer tentative, qualified support for virtual world experimentation. Overall, the behaviour of virtual subjects recruited, incentivised and observed within *Second Life* across a range of five standard experimental games was not found to differ significantly from established standard results. In addition, we identify certain methodological opportunities and challenges which confront virtual world experimenters.

Key words: virtual worlds, experiments

JEL classification: C72; C88; C99; Z13

1 Introduction

2 Social scientists are becoming increasingly interested in virtual worlds, three-
3 dimensional environments in which communities of networked individuals in-
4 teract (Castranova, 2005; Bainbridge, 2007; Bloomfield, 2007). There are two
5 reasons. First, the growing number of users and the scope and nature of socio-
6 economic activity between them are seen as interesting phenomena that merit
7 investigation in their own right (Castranova, 2005). Virtual worlds present
8 evolving cultures with independent social institutions that are becoming more
9 significant to society at large (Noveck, 2004). In economic terms, their evo-
10 lution from specialised video game networks to general social platforms has
11 generated a global industry of firms that leverage installed user bases for sub-
12 scription fees, advertising opportunities or virtual support services (Cagnina
13 and Poian, 2007). Many virtual worlds have evolving economies with fully con-
14 vertible currencies as well as functioning financial, labour and product markets
15 that are capable of producing a host of micro and macroeconomic phenomena
16 (Guest, 2007).

17 Second, the computer technology underlying virtual worlds provides novel
18 methods of conducting social science research (Bainbridge, 2007). To begin
19 with, it facilitates the economical and large-scale recruitment of diverse sub-
20 jects from different cultural-geographical and socio-economic groups for par-
21 ticipation in interviews, focus groups, surveys or experiments. In addition,
22 it affords control of the environment in which they decide and interact that
23 can be used to manipulate decision conditions, observe behaviour and collect
24 data. Conversely, however, both these features also present potential method-
25 ological problems. As subjects, virtual world users may not reflect standard
26 populations in terms of demographic or cultural characteristics and there-
27 fore may display different behaviours. The electronic interface that moderates
28 communication and interaction between them precludes physiological signals
29 and proximity that moderate economic behaviour in physical settings. Vir-
30 tual world culture, social institutions and conventions that evolve as a result
31 may shape economic interactions in ways that differ from traditional social
32 settings. The anonymity of the interface may hamper quality control in the
33 data collection process.

34 The current study is intended as a first, exploratory step towards the method-
35 ological issue. While virtual worlds may provide useful research tools for a
36 number of social science disciplines, we concentrate here on their potential as
37 platforms for designing and conducting economic experiments, an area which
38 may be especially conducive to benefit from the new methods virtual worlds of-
39 fer (Bainbridge, 2007; Castranova, 2006). Traditional experimental economics
40 involves testing economic theories by observing the incentivised decisions of
41 representative subjects under choice conditions systematically manipulated in

42 laboratory settings. Virtual worlds may provide opportunities for methodolog-
43 ical innovation here. The discipline has recently begun to broaden its scope
44 by exploring new methods and applications outside the standard controlled
45 laboratory environment commonly populated by Western student subjects.
46 There are two related ways in which experimentalists are trying to improve
47 the realism of the behaviour they observe. First, field studies in naturalistic
48 settings are being proposed as a way of avoiding the distorting effects artificial
49 laboratory settings may have on subject behaviour (Harrison and List, 2004).
50 Second, new recruitment techniques and sampling locations are being used
51 to overcome the reliance of experimentalists on Western university students
52 to generate results (Anderhub et al., 2001; Henrich et al., 2004, e.g.). Virtual
53 worlds may give an opportune impetus to both of these concerns. First, due
54 to their computerised interfaces, they may provide relatively controlled en-
55 vironments for conducting experiments while remaining within a naturalistic
56 setting familiar to subjects. Second, virtual worlds may be inhabited by a
57 wider cross section of people such that sampling from different cultures and
58 more heterogeneous backgrounds may be possible in a single location acces-
59 sible to experimentalists. In this sense, virtual worlds may bridge the gap
60 between laboratory experiments and field studies, allowing researchers to use
61 representative subjects in more natural environments to study the relationship
62 between the conditions of interaction and the evolution of social institutions
63 in a controlled manner.

64 We assess to what extent virtual worlds can be used in this context. We ap-
65 proach the issue in two ways, by replication and by observation. First, virtual
66 world experimentation can be a useful, alternative experimental tool to the
67 extent that the results it generates for particular tasks and conditions are the
68 same as those generated by traditional experimental methods. We assess this
69 aspect by conducting virtual experiments with a range of standard tasks in
70 standard conditions and comparing virtual subject behaviour with that of tra-
71 ditional pools reported in existing work. The suitability of virtual experimenta-
72 tion as an alternative would be supported to the extent that no differences are
73 found. As the observed subject behaviour may be related to their underlying
74 culture, demographics and values, we also used a survey instrument to collect
75 data on these which can be compared to standard populations. The difference
76 or similarity of virtual users to these provides additional insight into their
77 suitability as experimental subjects representative of economic agents gener-
78 ally. This first part of our approach tests the scope virtual worlds hold for
79 traditional economic experimentation, rather than for new avenues of experi-
80 mental research they may promise. We conceive of it as measuring the ‘output’
81 of the virtual experimentation method. The second part of our approach is
82 more qualitative and focuses on its ‘input’ side. This involves gathering in-
83 formal insights about the practical feasibility of economic experimentation in
84 virtual environments from the process of conducting experiments. We hoped
85 to learn by observation to what extent virtual worlds can provide a suitable

86 platform for experimental research generally, what the advantages and disad-
87 vantages are, and what modifications may be made to render virtual worlds
88 more amenable to experimentation. This second part may also provide in-
89 sights into what opportunities for new research approaches or methods virtual
90 worlds hold.

91 The rest of the paper proceeds as follows. In the next section, we discuss
92 the features of virtual worlds, their significance for experimentalists and our
93 procedure of methodologically assessing them. The results we obtained are
94 reported in section 3. Section 4 discusses our general observations from the
95 experiment in terms of the methodological issues we consider. The final section
96 contains concluding remarks.

97 **2 Virtual Experimentation**

98 *2.1 Virtual Worlds*

99 While there is considerable variation between the many alternative virtual
100 worlds that exist, they typically reproduce features of the physical world such
101 as a three-dimensional topography containing virtual objects obeying simu-
102 lated physical laws as well as the possibility of communication, social interac-
103 tion and economic exchange between users virtually represented by *avatars*.
104 We chose Second Life (SL, see Linden-Labs 2008) as the virtual platform for
105 our study. At the time of writing (November 2008), SL has over 15.7 million
106 registered avatars.¹ Accounting for multiple and dormant registrations, there
107 are an estimated one million regular users who spend over twenty million hours
108 logged in per month. Between twenty and thirty thousand users are online at
109 any one time. In terms of demographics, the majority of these are from popu-
110 lous and industrialised countries including the USA, the UK, Germany, Brazil,
111 France and Japan, with a median age of 36 and 57% being male.

112 SL is divided into individual sectors with topographical features in which
113 avatars can operate, including oceans, rivers, mountains and beaches as well
114 as flora. A typical location is displayed in figure 1. Avatars are capable of loco-
115 motion, including walking, running and flying and are immune to destruction.
116 They communicate using instant text messaging (IM) and can signal voice
117 intonation such as whispering and shouting as well as use gestures and body
118 language. Public IM can be received by all avatars in the vicinity, while private
119 IM is transmitted only between two avatars irrespective of location. Internet

¹ Economic and general statistics concerning SL are available at:
<http://secondlife.com/whatis/economy.php> and <http://blog.secondlife.com/>.



Fig. 1. Typical SL-screenshot showing the user's avatar (male foreground figure), the surrounding SL-environment and interface controls along the bottom.

120 telephony has recently been introduced to SL. Users can edit the appearance
 121 of avatars in terms of physical features, clothing and accessories. As a result,
 122 avatars can assume the form of humans, animals, fantasy creatures or objects.
 123 Avatars are associated with user accounts that include money balances in Lin-
 124 den dollars (L\$) which can be bought from or sold to Linden Lab, the creators
 125 and owners of SL, at a relatively stable exchange rate of about 270 L\$ per
 126 1 U.S. dollar. A total of 5.3 billion L\$ (U.S. \$19.7 million) are currently in
 127 circulation. SL provides an interface feature that allows immediate and direct
 128 account-to-account transfers. These balances can be used to purchase a port-
 129 folio of tradable virtual objects including land, buildings, vehicles, clothing,
 130 accessories and tools.

131 2.2 *Experimental Economics*

132 Virtual worlds such as SL may have potential as powerful new platforms for
 133 designing and conducting experimental research. Bainbridge (2007) makes the
 134 following case:

135 Virtual worlds such as SL provide environments and tools that facilitate
 136 creating online laboratories that can automatically recruit potentially thou-
 137 sands of research subjects, over a period of months, at low cost. SL offers
 138 scripting and graphics tools that allow anyone to build a virtual laboratory
 139 building, functioning equipment to run the experiment, and incentives to

140 motivate participation. (p. 473)

141 Conversely, however, the very technology that generates these advantages may
142 give rise to a number of *a priori* concerns about virtual experimentation. Prin-
143 cipally, experimenters know little about the identity or state of the subjects
144 who control the participating avatars. This may make it difficult to recruit
145 appropriate subjects, to ensure discipline in the virtual laboratory, to prevent
146 repeat participation and subject collusion and to engender subjects' trust and
147 confidence in the experiment. There is a possibility of demographic or cul-
148 tural idiosyncrasies of virtual subjects generally. This may generate a sample
149 bias that renders virtual experimentation inappropriate to test general eco-
150 nomic theories. They may have more hedonistic or short-term tendencies or
151 show less conformity than the average person. In addition, virtual behaviour
152 is not moderated by physical presence and may therefore not be comparable
153 to traditionally-generated results.

154 2.3 *Experimental Design*

155 The purpose of our study is to conduct experiments within SL to assess the
156 overall feasibility of virtual experimentation. Our approach is to gauge to what
157 extent the behaviour and values of virtual subjects conform to those of stan-
158 dard subjects. In the following, we outline the general working procedure that
159 we developed and deployed over the course of our experiments in terms of five
160 stages of which individual experimental sessions consisted. All our experimen-
161 tal sessions were conducted during standard GMT working hours between July
162 and November 2007. Experimental instructions are available upon request.

163 In the recruitment stage, we solicited participation by approaching online users
164 *in situ* immediately prior to a particular experimental session in the following
165 manner. Half an hour before a scheduled experimental session, we used a search
166 feature in the SL-inteface to identify the currently busiest locations in terms of
167 number of avatars present (excluding locations with an adult thematic focus).
168 Next, each of the three experimenters used their avatar to access one of these
169 locations and to address groups of avatars gathered there using public IM with
170 a standardised recruitment message. This message was in English and stated
171 our institutional affiliation and general information about the nature of the
172 task, its duration and incentivisation. Whenever interested users responded,
173 we answered any additional questions and informed volunteers of the time
174 and venue of the session. This process was repeated for a number of locations
175 and avatar groups in each until the recruitment of the desired number of
176 participants was complete. The thirty minute period was in almost all cases
177 sufficient to recruit between four and seven subjects.



Fig. 2. A typical experimental session in progress. The experimenters' avatars are standing.

178 Participants were transported to our virtual experimental laboratory in a ded-
 179 icated virtual building with controllable access rights and purpose-built labo-
 180 ratory furniture. In the briefing stage, subjects who have arrived (typically in
 181 groups between two and seven depending on the task) were given virtual doc-
 182 uments containing general information on experimental etiquette, anonymity,
 183 confidentiality and incentivisation. The two to three experimenters present at
 184 all times communicated with subjects using either public or private (i.e. one-
 185 to-one) IM. Once they have finished reading the briefing documents, subjects
 186 were asked to occupy cubicles that were purpose-built to restrict their vision
 187 and communication in order to prevent collusion between them. They were
 188 then given virtual documents containing the experimental instructions and
 189 a comprehension quiz. The decision task stage commenced after all subjects
 190 completed the quiz successfully. Experimenters instructed individually when
 191 subjects were initially unable to do so. Subjects communicated their decisions
 192 to the lead experimenter and received feedback via private IM. Next, in the
 193 survey stage, subjects were sent the URL of a webform containing a values
 194 survey as well as some demographic questions which they had to fill out. In the
 195 final, payment stage of the experimental session, subjects were paid earnings in
 196 \$L on the spot using the SL payment transfer feature. A typical experimental
 197 session in progress is shown in figure 2.

198 Table 1 provides some general information about the decision tasks of our ex-
 199 periments. Our choice of tasks was guided by our objective to assess whether a
 200 virtual subject pool may be appropriate in testing economic theories. In par-
 201 ticular, we wanted to examine whether virtual behaviour conforms to estab-
 202 lished results generated in conventional experimentation. As a result, we chose

203 the ultimatum (UG), dictator (DG), public good (PGG), guessing (GG) and
 204 minimum effort (MEG) games. Previous experimental results for all of these
 205 standard games abound for a variety of conditions as well as demographic and
 206 cultural groups and provide ready benchmarks for our own results. They also
 207 permit eliciting a broad spectrum of different types of strategic choice. In the
 208 following, we do not explain or analyse these standard games in detail, but
 209 report data from our and those previous studies most appropriate for compar-
 210 ison. We also report results from tests of differences in means, medians and
 211 overall distributions between them using t -tests, Mann-Whitney U (MW) and
 212 Kolmogorov-Smirnov Z (KS) tests respectively. While means tests can indi-
 213 cate differences between the overall behavioural propensities in two pools of
 214 subjects, distribution tests can also reveal differences in the incidence of a vari-
 215 ety of behaviours when average behaviour does not differ. For experimental
 216 tasks with multiple decision rounds, we also used regression analysis to test
 217 for differences with previous results. In particular, we pooled available data
 218 from our own and the previous study used as a comparator and estimated the
 219 following regression equation:

$$220 \quad Y_i^t = \alpha + \beta Y_i^{t-1} + \gamma X_i + \delta n_i \quad (1)$$

221 where Greek letters represent constant and parameters, Y is observed be-
 222 haviour, t the task round, n experimental group size and X a dummy variable
 223 for the comparator study. No differences between SL and comparator study
 224 behaviour exist to the extent that the coefficient for the latter variable is in-
 225 significant. The inclusion of the lagged variable on the right-hand side was
 226 intended to reduce omitted-variable bias in our model. In particular, it is well
 227 established that simple learning processes may explain some changes in be-
 228 haviour over time in specific game and choice contexts (see, e.g., Camerer
 229 1987, Erev and Roth 1998). As a result, we opted for a specification simi-
 230 lar to a partial adjustment model, where the behaviour in the current period
 231 is adjusted to that in the previous one. These kinds of dynamic model have
 232 been previously applied to the three games for which we seek to estimate be-
 233 haviour, i.e. the PGG (Healy, 2006), the GG (Kurz, 2008) as well as the MEG
 234 (Crawford, 1995).

235 It should be noted that our design makes no provision for establishing a control
 236 treatment by replicating our virtual experiments in a standard physical set-
 237 ting with otherwise identical experimental parameters. While this alternative
 238 has certain advantages, our approach was to rely instead on the replicability
 239 of existing studies and to design virtual experiments that mirror their task
 240 conditions such as to permit using their results as a comparator.

241 An additional avenue for testing subject pool suitability is to survey and com-
 242 pare our subjects' values and demographics to those of standard experimental
 243 subjects and general populations. Values provide a measurement of a respon-

Task	UG	DG	GG	PGG	MEG	ESS
Subjects (N)	64	60	31	32	31	113
Subjects per session (n)	4-5	4-5	3-7	4	5-6	n/a
Average pay (U.S. \$)	5.25	1.95	2.30	20.15	8.25	3.85
Duration (minutes approx.)	25	10	25	35	20	10
Rounds (r) or questions	1	1	10	10	10	21

Table 1

Summary statistics for experimental games and survey.

244 dent’s cultural orientation and are known to affect behaviour (Rokeach, 1973;
 245 Chuah et al., 2006). We used the human values survey designed by Shalom
 246 Schwartz for the European Social Survey (ESS) project (Schwartz, 2002). Like-
 247 wise, a number of demographics such as gender, age, and nationality are known
 248 to affect behaviour (see Camerer 2003 for an overview). In the following sec-
 249 tions, we report the results we obtained from the game tasks and survey.

250 3 Experimental Results

251 3.1 Subject Demographics

252 Subjects’ basic demographical data are summarised in figure 3. The average
 253 age of respondents was 32, with the youngest at 18 and the oldest at 64. Com-
 254 pared with the general population of the European Union (EU), the age range
 255 20-40 years was over represented, an expected result given the technological
 256 and cultural status of virtual worlds. In line with SL generally, most subjects
 257 were from populous Western nations, although UK and European countries
 258 were somewhat over-represented in our sample. The reason may lie in using
 259 the English language and our institutional affiliation in recruitment. Recruit-
 260 ing during GMT daytime hours further bias sample selection in terms of time
 261 zone. In terms of gender, exactly half of our respondents were male.

262 3.2 Ultimatum Game

263 Separate sessions with UG-proposers and responders were conducted on 6,
 264 25 and 26 July 2007. In the proposer sessions, subjects were given the task to
 265 decide how to share L\$3000 (U.S. \$11.50) with a randomly-chosen co-player
 266 from a responder session who had the choice to accept or reject the split,

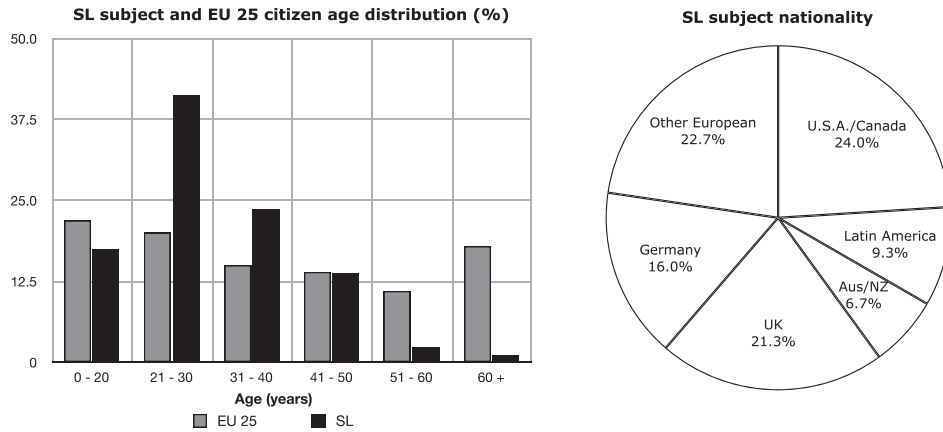


Fig. 3. Age and nationality distribution of SL-subjects.

267 resulting in the proposed shared being paid out or neither player receiving
 268 anything.

269 Although there is little evidence for stake size effects in the UG (see Camerer
 270 2003), we aimed for comparability by using a stake in the U.S. \$10-15 interval
 271 used in many previous studies, as well as for easy mental divisibility. Theory
 272 predicts that, because instrumentally-rational responders should accept any
 273 share of the stake, rational proposers should offer the minimum. However,
 274 proposers in previous studies offer in the region of 42-48% (see table 2.2.
 275 in Camerer 2003), reflecting a mixture of altruistic and strategic thinking on
 276 their part (Forsythe et al., 1994). In standard task conditions and subject pools
 277 recruited in industrialised nations, UG-results are relatively robust. Roth et al.
 278 (1991) (RPOZ) found little difference between offers made by urban subjects
 279 recruited in the U.S. (RPOZ 1), Tokyo (RPOZ 2), Yugoslavia and Israel.
 280 However, alternative cultural and demographic characteristics can generate
 281 differences (Camerer, 2003; Oosterbeek et al., 2004). Buchan et al. (1997) and
 282 Chuah et al. (2007) (CHJW) identified slightly but significantly higher offers
 283 of South-East Asian subjects potentially linked to their collectivist values.
 284 Henrich et al. (2004) found a much wider range of offers (between 25-57%)
 285 in a series of experiments with traditional, small-scale societies across the
 286 developing world.

287 Table 2 reports summary statistics of UG bargaining by SL-subjects compared
 288 with behaviour reported by RPOZ (1 and 2), by Hoffman et al. (1994) for
 289 U.S. subjects (HMSS) and by CHJW for UK subjects. The SL mean offer
 290 is 45.73% of the stake with a modal offer of half. These central tendencies
 291 in the proposals are very similar to those reported for comparable samples.
 292 Figure 4 shows the distributions of offers in all these experiments. With the
 293 exception of a small number of hyper-fair outliers among SL-subjects, the
 294 distribution we found is also very similar to those in the previous studies.
 295 Statistical tests bear these observations out. As the UK formed the largest

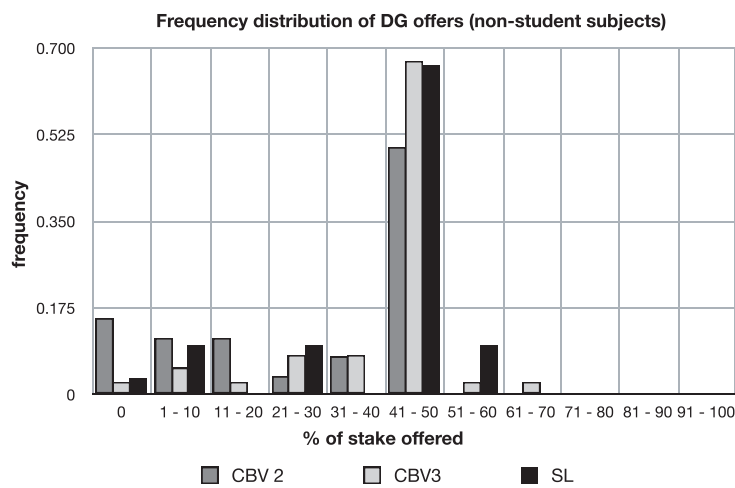
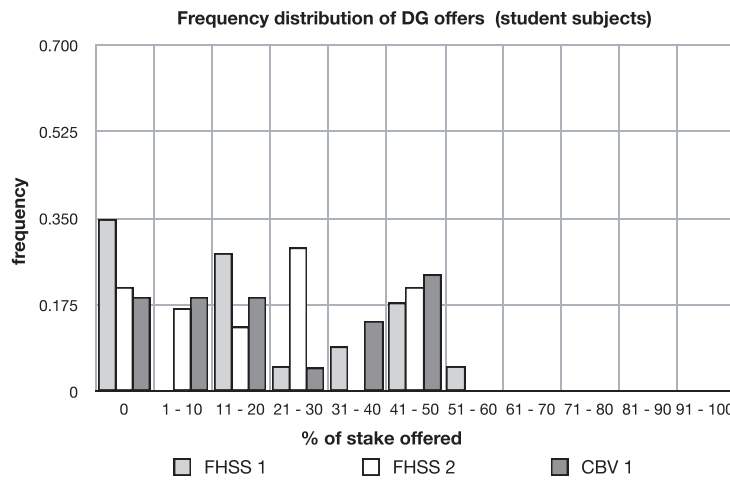
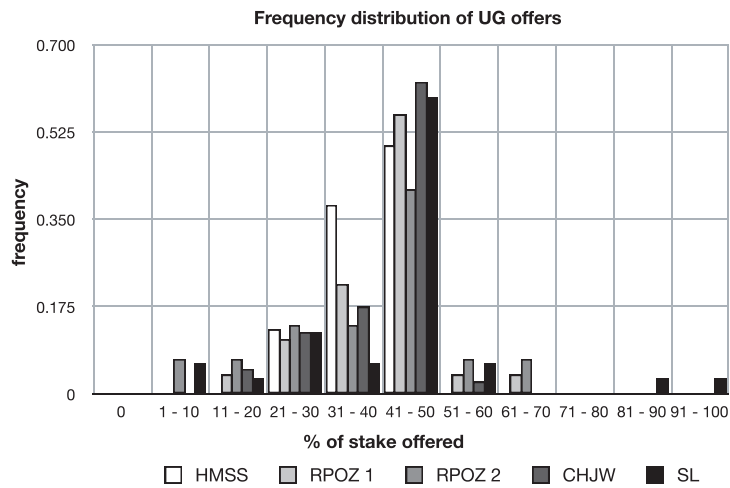


Fig. 4. Distribution of UG and DG offers in SL as well as in selected previous studies.

	SL	HMSS	RPOZ 1	RPOZ 2	CHJW
$N/2$	32	24	27	29	40
Stake	11.50	10	10	10	16
Offers					
Mean	45.73	44	45	45	44
Mode	50	50	50	50	50
St. Dev.	18.6	7.2	9.6	21.0	9.5
Rejections					
% of offers < 20%	33.33	-	-	50	-
% of all offers	6.25	8.3	22	24	15

Table 2

Summary statistics of ultimatum game offers (in % of the U.S. \$ stake) and rejections for $N/2$ subject pairs in SL as well as in selected previous studies.

296 national group among our subjects (see section 3.7), we used UK subject data
 297 from CHJW as a comparator for our findings. No differences in the mean
 298 ($t=0.216$, $p=0.829$), median ($U=2706.5$, $p=0.422$) or distribution ($Z=0.595$,
 299 $p=0.870$) of offers were found between their and SL subjects.

300 3.3 Dictator Game

301 DGs were also conducted in separate sessions for proposers and responders,
 302 except that responders were not given the opportunity to accept or reject
 303 offers. The sessions were conducted on 27 and 31 July 2007. As a number of
 304 previous studies employed stakes divisible by 10, and since stake size effects
 305 are not noticeable between studies with significantly different stakes (see table
 306 3), we opted for a stake size of 1000 \$L (U.S. \$3.90). The DG was originally
 307 conceived as a way of separating altruistic and strategic motives in UG-offer
 308 behaviour (Forsythe et al., 1994). While instrumentally rational players should
 309 keep all of the stake, experimental subjects offer in the region of 20-35% to
 310 responders, reflecting altruistic preferences. DG-behaviour is sensitive to a host
 311 of experimental conditions such as anonymity, source and destination of the
 312 stake (see Camerer 2003 for an overview). In addition, subject demographics
 313 influence offers.

314 Table 3 reports summary statistics of SL-dictator behaviour compared to sub-
 315 jects in comparable studies by Forsythe et al. (1994) (FHSS) and Carpenter
 316 et al. (2005) (CBV). Figure 4 displays the distributions of offers in the experi-
 317 ments reported there. The first two of these studies (centre panel of the figure)

	FHSS 1	FHSS 2	CBV 1	CBV 3	SL	CBV 2
$N/2$	24	45	21	26	30	37
Stake	10	5	100	100	3.90	100
Offers						
Mean	24	24	25	33	43	45
Mode	30	0	50	50	50	50
Median	25	20	20	45	50	50
St. Dev.	17.68	20.44	19	20	16.17	12

Table 3

Summary statistics of dictator game offers (in % of the U.S. \$ stake) for $N/2$ subject pairs in SL as well as reported in selected previous studies.

318 report offers made by standard college student subjects which tend to be in
 319 the region of 23-24% of the stake (see also Hoffman et al. 1996, Cason and
 320 Mui 1998), although some studies, such as Schotter et al. (1996), have found
 321 offers close to 40%. Of particular interest to us is the study by CBV, who
 322 identified marked differences in DG offer levels based on age and experimental
 323 location (bottom panel of figure 4). In their study, they compare offers made
 324 by students (average age: 19 years) in standard college settings (CBV 1), by
 325 older community college students (27, CBV 2) and by workers in a warehouse
 326 setting (37, CBV 3).

327 The data show the DG offers made by SL-subjects to be higher than those
 328 reported in standard college settings, but similar to those made by older sub-
 329 jects in CBV. These results reflect the greater average age of our subjects (see
 330 section 3.7) and the fact that DG-offers are sensitive to age (Harbaugh et al.,
 331 2003). Previous and current DG-results pertaining to older subjects are shown
 332 in the bottom panel in figure 4. It is also noteworthy that in our experiment,
 333 proposers communicated their offers to the experimenter directly using pri-
 334 vate IM rather than using forms collected and delivered in stacks by monitors
 335 as tends to be practiced in physical locations. Our treatment provides more
 336 scope for social influence and demand effects that would be expected to raise
 337 offers.

338 The age similarity between warehouse workers in CBV 3 to our own SL-
 339 subjects provides us with an appropriate benchmark for the comparison of
 340 DG-behaviour. No statistically significant differences were found between the
 341 means ($t=-0.700$, $p=0.485$) medians ($U=981.5$, $p=0.823$) and distributions
 342 ($Z=0.383$, $p=0.999$) of DG offer data in these two pools.

343 3.4 *Public Good Game*

344 The PGG sessions were conducted on 25 October and 2 November 2007. In
345 them, subjects in groups of $n = 4$ were asked to divide a stake of L\$400 (U.S.
346 \$1.50) between a private and a group fund and explained that their total earn-
347 ings would be their private allocation plus $a = 0.4$ times the total of all group
348 allocations. This was repeated $r = 10$ times. The parameter values for n , r
349 and a were chosen with comparability with other studies in mind (see table 4).
350 The PGG is a n -person version of the prisoner's dilemma and pits subjects'
351 self-serving motives against their desire to further the benefit of the group.
352 Instrumentally-rational play involves complete free-riding and allocating the
353 whole endowment to the private fund. In repeated PGGs, players decisions
354 may be guided both by strategic considerations of reciprocation and purely
355 altruistic motives. A large literature exists that identifies the experimental
356 conditions that elicit cooperative behaviour. In general, subjects contribute
357 positive amounts to the public good that steadily decline as the game is re-
358 peated. The studies reporting PGG games under standard conditions serve as
359 benchmarks for the behaviour of our SL-subjects. We compare the behaviour
360 of SL-subjects with those in experiments with comparable conditions reported
361 by Andreoni (1988, 1995) (A (88) and A (95)) as well as Fehr and Gächter
362 (2000) (FG), who used values for parameter a of 0.5, 0.5 and 0.4 respectively.
363 Table 4 reports summary statistics of SL-PGG behaviour compared to sub-
364 jects in these three.

365 The top panel in figure 5 shows the average contribution to the group fund
366 subjects made in SL and in the three previous studies over ten rounds. SL-
367 subjects contribute marginally more than subjects in the other pools in all
368 rounds. The average contribution decays over rounds in similar ways in all
369 studies. The higher average we find is not unusual within the context of find-
370 ings made using variegated subject pools. For instance, Henrich et al. (2004)
371 report on PGGs played with traditional society subjects in many continents
372 and find mean contribution rates to vary between 22 and 65%. The SL sub-
373 jects differ from standard college students in a number of ways, age being one.
374 Our result may also be due to the apparent greater altruism of SL-subjects
375 compared with students we observed in the DG.

376 For our statistical tests of PGG behaviour, we chose A (95)'s Western student
377 subject data as a benchmark. It should be borne in mind that this experiment
378 differs from our study in two ways; the differences in experimental platform
379 we are assessing, and the differences in subject demographics. We performed
380 mean, median and distribution tests between the offers for each of the ten
381 rounds played by A (95) and SL subjects (see table 5). Only one of the resulting
382 thirty test statistics was significant ($Z_{n=10} = 1.370, p=0.047$). As the repeated
383 testing procedure amplifies the probability of Type I errors, we also estimated

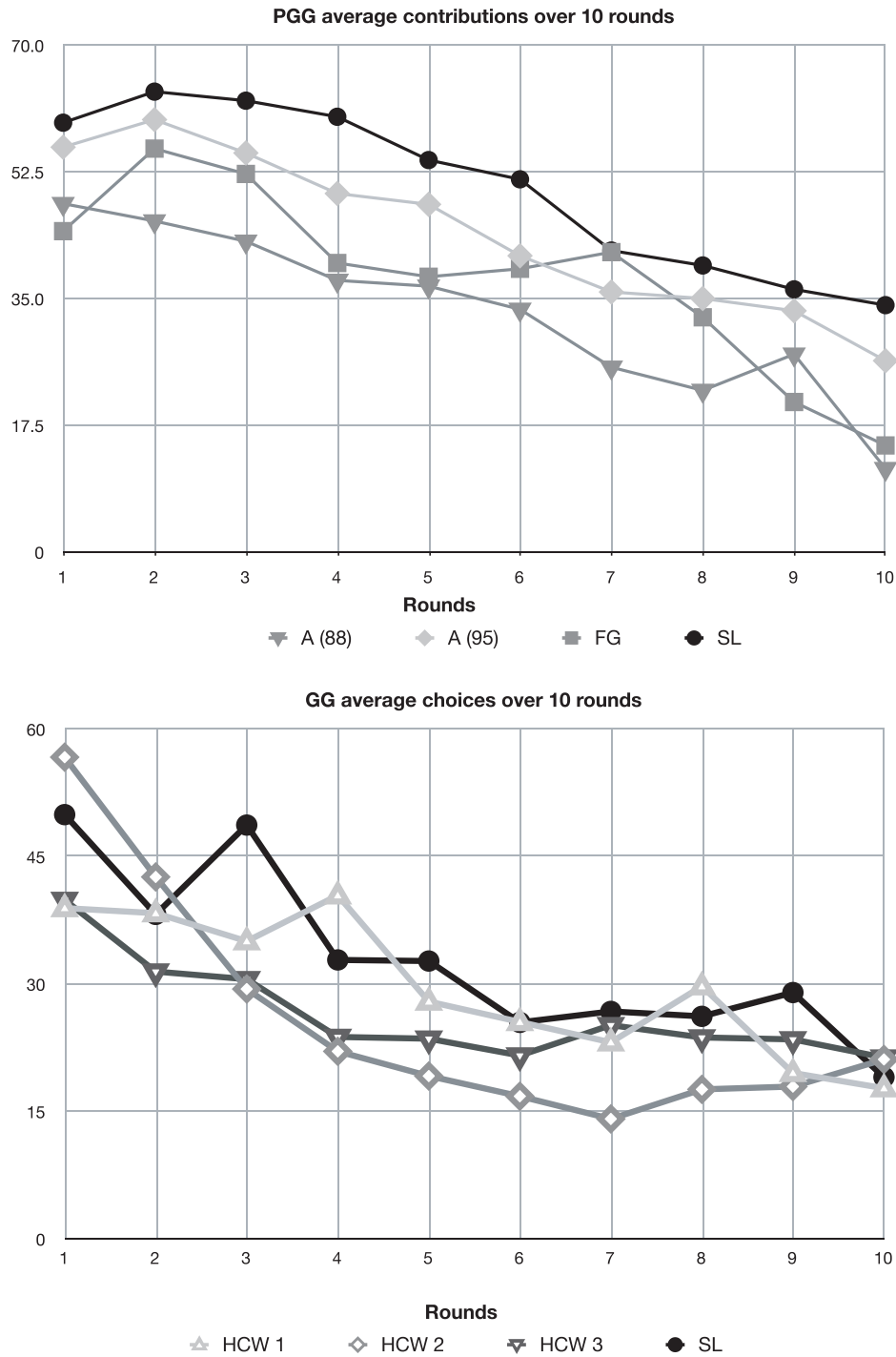


Fig. 5. Average subject decisions in GG and PGG over $r=10$ rounds in SL and selected previous studies.

384 equation 1 to compare the two data sets. The factor n could not be entered
 385 due its insufficient variation in the data set. The regression results are given
 386 in table 6 and show an insignificant coefficient for X , leading us to conclude
 387 that no behavioural differences are in evidence.

	A (88)	A (95)	FG	SL
N	30	40	24	32
n	5	5	4	4
r	10	10	10	10
Stake	0.50	0.60	0.86	1.50
α	0.5	0.5	0.4	0.4
Contributions				
Mean	33.20	44.09	37.94	50.34
Median	32.00	42.50	40.25	45.63
St. Dev.	21.65	27.47	16.89	22.54

Table 4

Summary statistics of public good game contributions (in % of the U.S. \$ stake averaged over r rounds) for N subjects playing in groups of n in SL as well as reported in selected previous studies. Stakes are given as U.S.\$-values of tokens subjects were asked to allocate per round.

388 3.5 Minimum Effort Game

389 The MEG sessions were conducted between 16 and 21 November 2007. In
 390 them, groups of $n = 5$ to 6 subjects were asked to choose an integer in the
 391 interval $[1, 7]$ and informed that payoffs would be determined by the smallest
 392 number chosen within the group according to the payoff matrix adapted from
 393 Van Huyck et al. (1990) (VBB) and shown in table 7. Each group played ten
 394 rounds of this game. Again, these parameter values are standard to the extent
 395 that they have been adopted by the majority of previous studies. The game
 396 has multiple equilibria in which all players make the same choice, which payoff
 397 dominate each other in turn with a unique Pareto-efficient equilibrium in every
 398 player choosing 7. The game represents situations where a group's ability to
 399 coordinate on the individually as well as collectively best outcome may be
 400 undermined by individuals' pessimistic expectations of others' reasoning. A
 401 typical example is punctuality (Camerer, 2003). While everyone arriving on
 402 time for a meeting is mutually the best outcome, an individual may arrive late
 403 to avoid a wait expecting others to also be late. After a number of meetings,
 404 such expectations may become increasingly self fulfilling as general punctuality
 405 disintegrates. Previous experimental evidence shows this kind of convergence
 406 on payoff-dominated outcomes to be dependent on the size of the group, the
 407 size of payoffs and information players receive about the choices of others.

408 Figure 6 shows the round-to-round changes in the choices and minimum

Task	r	t	MW U	KS Z
PGG	1	0.431 (0.667)	619.0 (0.808)	1.054 (0.216)
	2	0.499 (0.619)	611.5 (0.743)	0.949 (0.329)
	3	0.864 (0.391)	572.0 (0.436)	0.764 (0.603)
	4	1.231 (0.223)	536.5 (0.235)	1.001 (0.269)
	5	0.697 (0.488)	567.0 (0.403)	0.817 (0.517)
	6	1.231 (0.222)	534.5 (0.227)	0.870 (0.436)
	7	0.673 (0.503)	573.5 (0.446)	0.738 (0.648)
	8	0.568 (0.572)	544.5 (0.274)	1.370 (0.047**)
	9	0.372 (0.711)	567.5 (0.405)	0.817 (0.517)
	10	0.926 (0.358)	539.5 (0.240)	1.133 (0.153)
MEG	1	1.482 (0.141)	982.5 (0.139)	0.895 (0.452)
	2	1.218 (0.226)	1023.0 (0.236)	0.833 (0.491)
	3	1.927 (0.057)	931.0 (0.070*)	1.109 (0.171)
	4	2.660 (0.009***)	822.5 (0.011**)	1.353 (0.051*)
	5	1.449 (0.150)	986.0 (0.153)	0.713 (0.690)
	6	1.382 (0.170)	990.0 (0.162)	0.983 (0.289)
	7	1.571 (0.119)	955.5 (0.102)	0.888 (0.409)
	8	0.785 (0.435)	1059.0 (0.351)	0.951 (0.326)
	9	0.518 (0.606)	1073.5 (0.406)	1.042 (0.228)
	10	2.364 (0.020**)	841.0 (0.014**)	1.347 (0.053*)
GG	1	1.798 (0.078*)	219.0 (0.079*)	0.928 (0.355)
	2	0.091 (0.928)	305.0 (0.923)	0.478 (0.976)
	3	2.195 (0.033**)	212.5 (0.060*)	1.226 (0.099*)
	4	-1.090 (0.281)	268.5 (0.423)	0.821 (0.510)
	5	1.003 (0.321)	289.5 (0.692)	0.664 (0.771)
	6	0.032 (0.974)	280.5 (0.569)	0.703 (0.706)
	7	0.538 (0.593)	283.0 (0.602)	0.664 (0.771)
	8	-0.552 (0.583)	278.5 (0.543)	0.652 (0.788)
	9	2.107 (0.041**)	250.5 (0.250)	1.277 (0.077*)
	10	0.279 (0.781)	292.5 (0.735)	1.063 (0.209)

Table 5

Test statistics for differences in mean (t), median (U) and distribution (Z) of behaviour between SL subjects and those in selected previous studies for $r=10$ rounds. Corresponding p -values are given in parentheses. The symbols *, ** and *** denote significance at the 10, 5 and 1% levels respectively.

PGG	Estimate	<i>t</i> -value	<i>p</i> -value
Constant	14.24	6.39	0.000***
Y_{t-1}	0.67	22.63	0.000***
X	-2.47	-1.15	0.252
	R^2 (adj.)= 0.45	$F= 260.69$	$p=0.000***$
MEG	Estimate	<i>t</i> -value	<i>p</i> -value
Constant	1.22	1.36	0.17
Y_{t-1}	0.61	24.03	0.000***
n	0.06	0.43	0.664
X	-0.34	-2.12	0.034**
	R^2 (adj.)=0.39	$F=204.55$	$p=0.000***$
GG	Estimate	<i>t</i> -value	<i>p</i> -value
Constant	26.84	11.38	0.000***
Y_{t-1}	0.25	8.09	0.000***
n	-1.11	-2.82	0.005**
X	-1.36	-0.81	0.416
	R^2 (adj.)=0.09	$F=30.42$	$p=0.000***$

Table 6

Regression results for experimental behaviour across three tasks in SL and one comparator study respectively. The symbols *, ** and *** denote significance at the 10, 5 and 1% levels respectively.

409 choices averaged over experimental groups in SL and comparable previous
 410 studies of Knez and Camerer (1994) (KC), Bornstein et al. (2002) (BGN),
 411 Devetag (2005) (DT) and VBB. Table 8 reports summary statistics of SL-
 412 PGG behaviour compared to subjects in these studies. All these studies used
 413 VBB's payoff matrix and had groups between 5-7 subjects except VBB, which
 414 had groups of 14-16. The figure shows similar declines in choices in all these
 415 studies. On the other hand, there appears to be greater variability in the over-
 416 all level of average choices, with SL-averages appearing higher than those in
 417 other studies.

418 We used the data reported by DT for the comparison with SL-observations.
 419 In terms of means, medians and distributions for $r=10$ rounds, round four
 420 and ten behaviours were different in terms of all three at the 10%-level of
 421 significance (see table 5). With one exception ($U_{n=3} = 931.0, p = 0.070$), the
 422 other twenty-four tests were negative, suggesting no differences exist in the

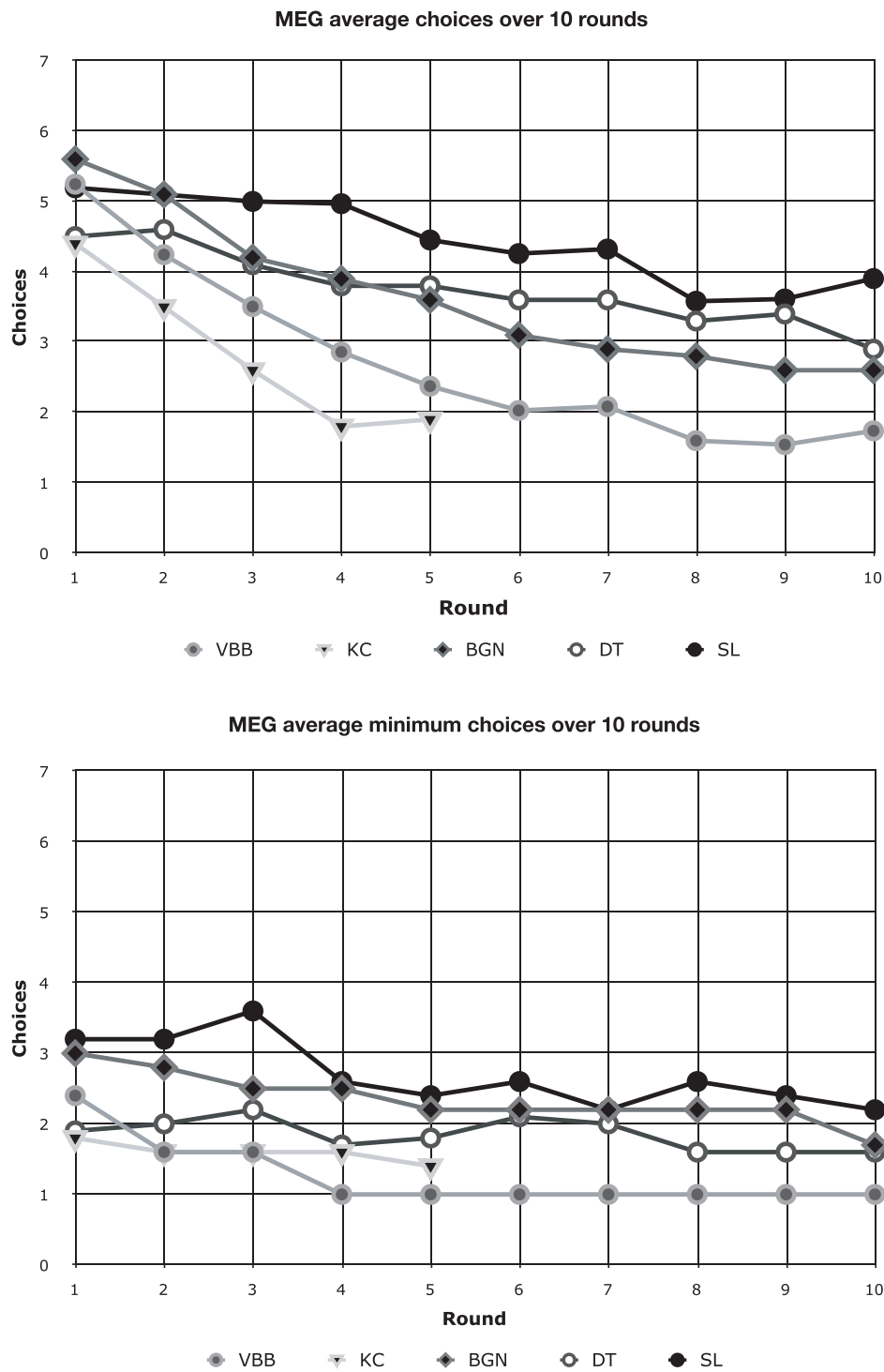


Fig. 6. Average and average minimum MEG choices over $r=10$ rounds in SL and selected previous studies.

	Smallest choice in group						
	7	6	5	4	3	2	1
7	390	330	270	210	150	90	30
6	-	360	300	240	180	120	60
5	-	-	330	270	210	150	90
4	-	-	-	300	240	180	120
3	-	-	-	-	270	210	150
2	-	-	-	-	-	240	180
1	-	-	-	-	-	-	210

Table 7

MEG payoff matrix (in L\$). The first column represents player choices which, combined with the smallest choice in the group, determines payoffs. Dashes denote logically impossible outcomes.

	VBB	KC	BGN	DT	SL
<i>N</i>	107	30	42	77	31
<i>n</i>	14-16	6	7	7	5-6
<i>r</i>	10	5	10	14	10
Stake	1.30	1.30	1.30	1.82	1.46
Choices					
Mean	2.72	2.87	3.65	3.75	4.44
Median	2.50	2.80	2.40	3.60	4.60
St. Dev.	1.30	1.07	1.34	1.57	1.51

Table 8

Summary statistics of minimum effort game choices over r rounds for N subjects playing in groups of n in SL as well as reported in selected previous studies. Stakes are given as U.S.\$-value of payoff associated with unique Pareto-efficient outcome.

423 rounds concerned. Again, we regressed equation 1 for the combined data set
 424 (table 6). The results show that at the 95% significance level, our data are
 425 different to those of DT as the coefficient for X is significant ($p = 0.034$). It
 426 should be noted that the same model also yields differences between the data
 427 of DT and BGN ($p = 0.084$) as well as between SL and BGN ($p = 0.002$). As a
 428 result, for the MEG, these findings do not provide firm conclusions about the
 429 ability of virtual world experimentation to replicate laboratory results. The
 430 two comparator experiments differ from ours in an additional, demographical
 431 dimension and also differ from each other in terms of results. The reason may

432 lie in greater general variability in MEG-behaviour due to the presence of
 433 multiple equilibria.

434 3.6 *Guessing Game*

435 The GG sessions were conducted on 8 and 15 November 2007. In them, $n=3$
 436 to 7 subjects were asked to choose integers in the interval $[0,100]$ and informed
 437 that the subject with a response closest to $g = 0.7$ times the average of all
 438 choices would receive L\$200 (U.S. \$0.75). Ties were resolved by dividing this
 439 sum among the winners. Each group of subjects played $r = 10$ rounds of this
 440 game.

441 The GG (sometimes known as the beauty contest game) is used as a tool to
 442 identify what levels of reasoning subjects employ in strategic thinking (Nagel,
 443 1995; Duffy and Nagel, 1997; Camerer, 1997). A zero-order (i.e. unstrategic)
 444 player may choose randomly or use a focal point such as the median of the
 445 interval (50 in our case). First-order choosers may take others into consider-
 446 ation but assume these to be of order 0. An optimal first-order choice would
 447 be in the interval $[0,70]$ accounting for the impossibility of the group average
 448 to exceed 70. In particular, a choice of 35 (0.7×50) may reflect a belief that
 449 zero-order guessers choose 50 on average. Second-order players who assume
 450 others to use order 1 will not choose above 49 (0.7×70), and may opt for 25
 451 (0.7×35) believing order 1 choices to average 35 and so forth. The iterative
 452 application of increasingly higher levels of reasoning will eventually yield an
 453 equilibrium choice of 0.

454 The average and distribution of GG-choices therefore provides insights not
 455 only to what levels of reasoning subjects use, but also what levels they at-
 456 tribute to others. Equilibrium choices may reflect higher orders of reasoning
 457 but be ineffective when other players operate at lower levels. In addition, re-
 458 peated GGs show to what extent subjects learn to adjust their choices on the
 459 basis of previous rounds' results. Table 9 shows statistics concerning subjects'
 460 choices in single or first rounds of repeated games played in groups of differ-
 461 ent sizes with a parameter $g = 0.7$. The Singaporean student data are from
 462 10-round GG-experiments reported in Ho et al. (1998) (HCW). The HCW 1
 463 pool consisted of 3-player groups playing the game for the first time. Subjects
 464 in HCW 2 also played in 3-player groups but had experience of one previous
 465 game with a different g -value. Finally, HCW 3 was composed of inexperienced
 466 7-subject group players. In all HCW-treatments, the winning subject received
 467 50 Singapore cents (ca. U.S.\$ 0.34). The U.S. study of Kovalchik et al. (2005)
 468 (KCGPA) compares one-round choices by college students (KCGPA 1) with
 469 those of mentally healthy senior citizens with an average age of 82 (KCGPA
 470 2). Our experimental settings of group size, g -value and repetition are the

Subjects	Mean	Median	St. Dev.	% 0	<i>N</i>
Caltech students	21.88	23.00	10.35	0.07	27
Portfolio managers	24.31	24.35	16.15	0.08	26
Economics PhDs	27.44	30.00	18.69	0.13	16
U.S. high school students	32.45	28.00	18.61	0.04	52
College students (KCGPA 1)	35.00	35.00	12.86	0.00	51
Singaporean students (HCW 1)	36.45	35.00	24.28	0.00	21
German students	36.73	33.00	20.21	0.03	67
Senior citizens (KCGPA 2)	37.00	33.00	17.46	0.00	50
University CEOs	37.81	36.50	18.92	0.03	73
Wharton students	37.92	35.00	18.84	0.00	35
Singaporean students (HCW 3)	39.78	35.00	25.46	0.02	49
SL	50.00	56.00	27.10	0.00	31
Singaporean students (HCW 2)	58.27	50.00	26.98	0.05	21

Table 9

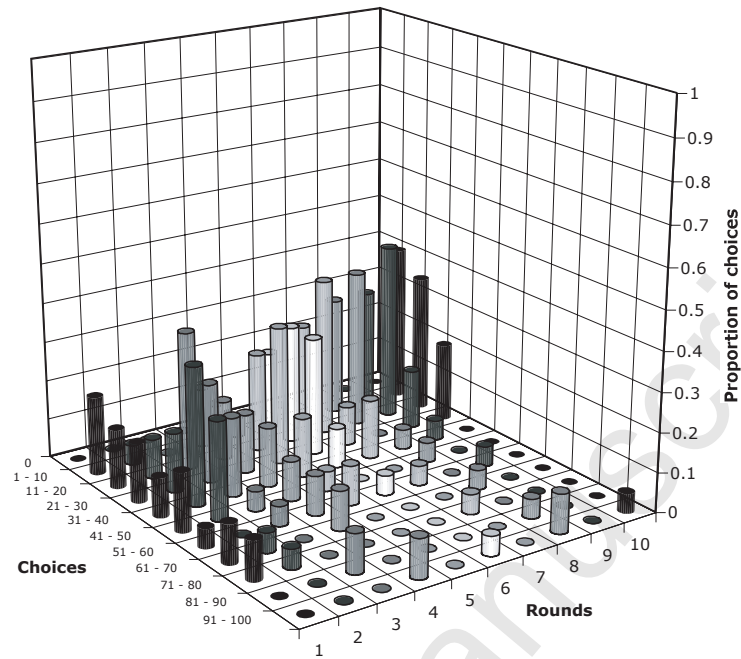
Summary statistics for round 1 GG choices in n -subject pools in SL as well as reported by Camerer (2003), Camerer (1997) and Kovalchik et al. (2005). The percentage of subjects choosing 0 is given by %0.

471 same as in HCW 1, which is most useful for a direct comparison.

472 SL first round choices are relatively high (especially compared to our bench-
473 mark HCW 1) but by no means outside the range of previous results. The bot-
474 tom panel in figure 5 shows mean choices over ten rounds among SL-subjects
475 and Singaporean students (HCW). Table 9 reports summary statistics of SL-
476 GG behaviour compared to subjects in this study. Our subjects did appear to
477 converge towards the equilibrium at similar rates to the latter. The frequency
478 distribution of individual SL-choices over all ten rounds is displayed in figure 7,
479 along with the corresponding data for HCW 1 reported in Ho et al. (1998)(p.
480 955, figure 2E). Both distributions are similar in that a greater proportion
481 of choices are low in later rounds. The SL-data appear different mainly in
482 the more equal distribution in early rounds. However, towards the end of the
483 game, the distributions are more similar, reflected in the convergence of curves
484 in figure 5.

485 GG data generally show divergence in first-round average choices. Part of the
486 reason may be the role that players' common knowledge of rationality has in
487 equilibrium reasoning. Lower choices are not merely associated with greater
488 strategic sophistication among players, but also with greater expectations con-

Frequency distribution of GG choices (HCW 1)



Frequency distribution of GG choices (SL)

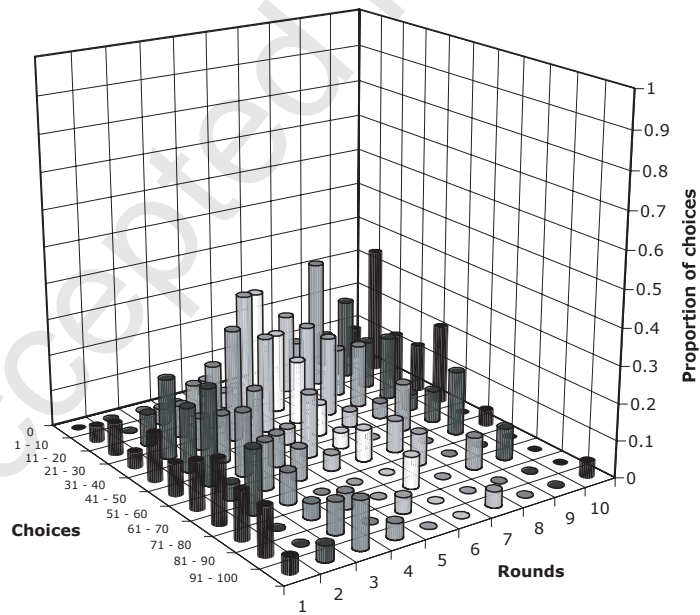


Fig. 7. Subject choice frequency distributions over $r=10$ rounds (group size 3, $p=0.7$) in HCW 1 and SL.

cerning the sophistication of others. Groups that are more sophisticated as well as more uniformly so, such as Caltech students, may therefore be expected to exhibit lower choices than comparatively heterogeneous groups such as SL where little is known about others who take part. Our first-round results may have not been much different had our pool consisted of anonymous and mutually unaware game theorists disguised by avatars. The fact that SL-subjects' learning resulted in similar final-round choices supports this possibility. The anonymity of SL, potentially subverting the common knowledge of rationality, may therefore partly explain any differences in round one choices in SL.

We compared the means, medians and distributions of SL choices with HCW 1 over $r=10$ rounds (see table 5). Rounds 1, 3 and 9 show differences in all three dimensions. In total, seven of the thirty tests were positive, most only at the 10%-significance level. Table 6 shows the regression results for equation 1 pooling SL data with HCW 1 and 3. The latter study was not used for the tests as its larger subject group size rendered it inappropriate for a direct comparison; however, we were able to control for that difference using variable n in the regression. The results show an insignificant coefficient for X ($p = 0.416$). We conclude differences are not in evidence between the data sets.

3.7 *Universal Human Values*

In order to assess whether an idiosyncratic cultural environment exists within SL, we administered the ESS human values survey. This survey is based on Schwartz's portrait values questionnaire, a well-tested instrument for identifying ten universal value dimensions (listed in figure 8). An individual's scores are calculated on the basis of responses on a 6-point Likert scale indicating own similarity with 21 hypothetical value portraits. Subjects completed the survey on a webform immediately after the decision task stage of the session. Upon completion, each subject was paid L\$1000 (ca. U.S. \$3.85) for the survey in addition to the pay-outs from the decision task.

Again, a host of existing data for this survey generates scope for comparing SL-subjects with standard populations. Cultural and demographic factors may have an influence on economic behaviour as they shape an individual's social interaction and socialisation into particular values. Values are therefore an important indicator of how representative particular subject pools are of the underlying population to which economic theory relates. We conducted the human values survey in order to ascertain to what extent SL-residents resemble standard experimental subjects culturally. Figure 8 shows the average value orientations of our subjects compared with those of respondents of the 2002-2003 ESS, as well as a standard sample of thirty-six UK university stu-

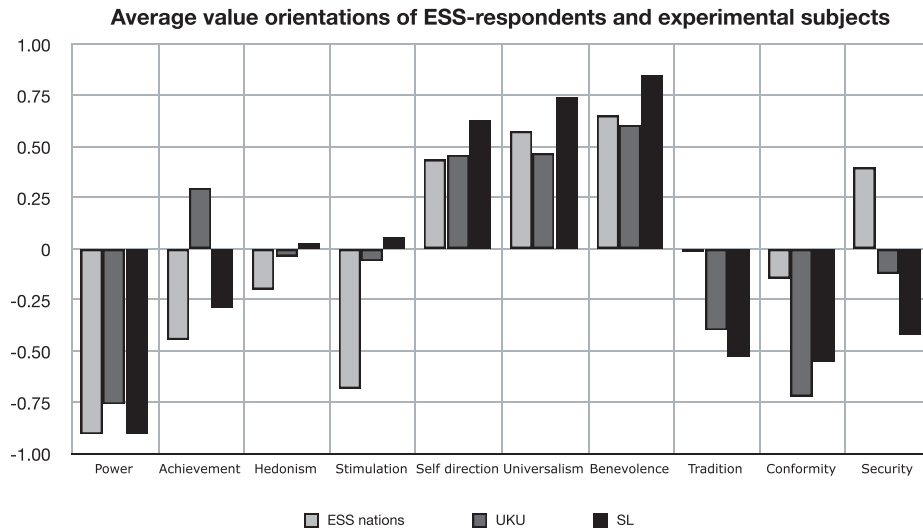


Fig. 8. Average orientations of ESS-respondents (ESS), SL and UK student subjects (UKU) according to Schwartz' ten value dimensions.

528 dents (UKU) we also administered the questionnaire to. The ESS randomly
 529 samples more than 1500 adults from each participating nation's resident popu-
 530 lation. The students were UK nationals invited randomly by automated email
 531 from the experimental subject database maintained by the *Centre for Decision*
 532 *Research and Experimental Economics*. For comparative purposes, we follow
 533 the ESS practice of presenting averages of ipsative scores, i.e. an individual's
 534 Likert-scale responses standardised in terms of his or her overall response av-
 535 erage and variance. Ipsatised scores for different value dimensions have the
 536 advantage of being comparable in terms of relative strength.

537 Schwartz' ten human values are shown along the horizontal axis of figure 8.
 538 They have established empirical interrelationships that are commonly used
 539 to reduce them to two basic dimensions shown along the two respective axes
 540 in figure 9. The first dimension, *self-transcendence v. self-enhancement*, en-
 541 compasses six values: hedonism, stimulation and self direction relative to tra-
 542 dition, conformity and security. The former three values express underlying
 543 motivations such as pleasure, sensuous gratification, excitement, novelty and
 544 independence, while the latter express respect and acceptance of norms, self-
 545 restraint and harmony. The remaining four universal values are contained in
 546 the the second dimension, *openness to change v. conservatism*. It weighs the
 547 values of universalism and benevolence against those of power and achieve-
 548 ment. The former two values express motivations including tolerance and care
 549 for the welfare of others, while the latter two encompass social status, personal
 550 success and dominance over others. Figure 9 plots nations and subject pools
 551 according to the two overall dimensions.

552 Our survey data indicate that while SL-users' value orientations differ from
 553 those of ESS-respondents, they do so to a lesser extent than those of the UK

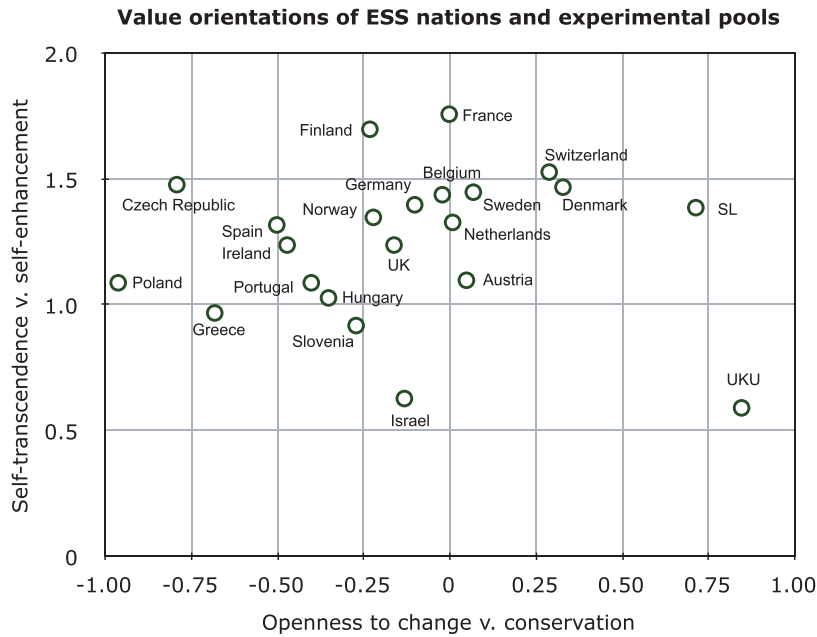


Fig. 9. Average orientations of ESS-respondents by nationality, SL and UK student subjects (UKU) according to Schwartz' two composite value dimensions.

554 student subjects. The SL and student average value orientations correlate at
 555 90% with each other, and respectively at about 70 and 64% with the averaged
 556 overall ESS-orientation of EU respondents. By comparison, individual national
 557 samples within ESS correlate with the average EU-values profile at about
 558 94%. The graph shows a relatively small distance between randomly-sampled
 559 individuals from European nations to SL-users and UK students. The students
 560 place a greater importance on the factors underlying self-enhancement, as can
 561 be verified in figure 8. This is consistent with age effects found in previous value
 562 surveys comparing students and teachers (Schwartz, 2001). Another reason for
 563 the difference may lie in a slightly higher relative socio-economic background
 564 and educational potential of students. However, caution has to be exercised
 565 due to our small sample size.

566 4 Methodological Discussion

567 Our experience of conducting experiments in SL suggests a number of ad-
 568 vantages and disadvantages of virtual experimentation generally as well as
 569 practical steps to adapt the platform for experimental purposes.

570 It was possible, with little organisation and preparation, to recruit subjects
 571 *in situ* in the numbers we could manage within the SL-interface. SL's fea-
 572 tures make it simple to create and maintain a database of subjects for future
 573 use. On the other hand, this procedure is prone to biased sample selection

574 on the basis of choosing busy recruitment locations, of solicitation, in the
575 recruitment language, time and institutional affiliation we used. In addition,
576 the relative anonymity that avatars confer on subjects makes it difficult in
577 practice to prevent financially-motivated repeat participation or the recruit-
578 ment of unfit (tired or intoxicated) or non-eligible or non-targeted subjects.
579 While these issues may not be completely resolvable, we attempted to miti-
580 gate both repeat and unsuitable participation by disqualifying avatars using
581 the following criteria. First, to avoid repeats, we excluded avatars who partici-
582 pated previously, who were created after the first experimental session or who
583 made unsolicited approaches to us. To avoid unsuitable participants, we also
584 excluded avatars less than a month old and potentially insufficiently familiar
585 with the SL-environment, avatars referred by previous subjects who may have
586 prior knowledge of the task, and avatars representing users who appeared to
587 be in an unfit state. An additional identity issue both in our and in other
588 virtual world studies concerns the potential for a disparity between user and
589 avatar characteristics. For many users, the attraction of SL consists of the
590 potential for using an avatar to assume a new and different identity. While
591 our study was designed to elicit the behaviour and values of users and not
592 avatars, we cannot be certain to what extent this was practised by subjects
593 responding through their avatars.

594 Our demographical and values survey shows that virtual worlds provide oppor-
595 tunities for recruiting subjects who are demographically more representative
596 than university students. In addition, targeting particular types of individuals
597 is possible within those groups represented in virtual worlds, such as partic-
598 ular nations. Clearly, some groups are currently not sufficiently represented
599 in virtual worlds, including individuals from smaller and traditional societies.
600 However, the bias of SL towards industrialised nations is likely to change as
601 economic development provides greater access to the Internet to more people
602 worldwide.

603 The relatively sophisticated SL-economy provides some scope for appropriate
604 incentive mechanisms. In particular, SL has developed informal labour and
605 product markets which generate incentivisable subjects as well as money or
606 in-kind rewards that can be delivered easily. Many users regularly participate
607 in paid online activities for returns which are modest compared with those of
608 standard economic experiments. In addition, the developed markets for virtual
609 objects provide alternative in-kind incentives.

610 While the computerised interface of SL provides an economical experimental
611 environment that is well suited for data generation, collection and storage, it
612 also has certain disadvantages. Communicating with subjects using IM makes
613 it difficult to deal with more than a handful per session. In addition, private IM
614 makes it hard to detect collusive behaviour or conferring amongst subjects.
615 While is it not possible to override the communication mechanisms of SL,

616 we developed virtual laboratory furniture that alerts the experimenter to the
617 potential for clandestine communication between subjects (visible in figure
618 2). In particular, upon entering the virtual laboratory, subjects were asked
619 to sit in cubicles and to enter *mouselook*, a SL-mode under which avatars
620 are restricted to frontal vision and where private IM is suspended, in line
621 with standard experimental conditions. Once activated, the furniture indicates
622 whenever a subject suspends the mouselook mode and is therefore able to use
623 private IM. While this furniture assured discipline in practice, it is in theory
624 possible for experts to circumvent such mechanisms. On the other hand, this
625 requires not only significant expertise on the part of a subject, but matching
626 skills of and prior collusion with another subject present in order to establish a
627 clandestine communication channel. Another problematic issue is establishing
628 subject trust in the experimenters. Because of the nature of virtual worlds,
629 it is difficult to convince subjects of the genuine nature of the experiment
630 and incentivisation. A further problem involves the potential for disruption
631 of experimental sessions by other users. This, however, may be controlled by
632 restricting access to the virtual laboratory.

633 The absence of physical signals and presence in virtual worlds creates clear
634 differences between virtual and physical experimental conditions. Virtual ex-
635 periments preclude physical presence that may influence behaviour through
636 involuntary non-verbal communication that reveals emotional states. In ad-
637 dition, the potential for anonymity means that the social consequences of
638 virtual behaviour are different to those in physical laboratories. These factors
639 may limit the comparability of virtual and physical experimental results in
640 many cases. Virtual experimentation is clearly not appropriate when physical
641 phenomena are part of the experimental treatment, such as when the effect of
642 face-to-face interactions is tested.

643 5 Conclusion

644 Despite the non-standard nature of the SL-subject pool and certain imper-
645 fections of the experimental environment that it provides, we were unable to
646 detect significant and systematic overall differences between their behaviours
647 and those observed in traditional settings. In particular, given SL-users' demo-
648 graphics in terms of age and cultural background, behaviour closely matched
649 expectations based on a host of existing experimental evidence for a range
650 of five important games. These results suggest tentatively that virtual world
651 economic phenomena are based on similar behavioural regularities observed
652 in standard economic settings and can be tested experimentally within the
653 virtual environment.

654 In addition, there is a slightly lesser cultural and age bias within SL than

655 at the average university campus. Users' values are more in line with those
656 of general populations of economic agents. There was little evidence of users'
657 niche interests or motivations generating an unsuitable subject pool. Our work
658 therefore supports Yee (2006), whose study of virtual world demographics
659 dispels the popular notion that they are predominantly the domain of a male,
660 adolescent sub-culture with niche interests. His data indicate that usage and
661 appeal are equally strong over gender and age groups as well as based more
662 on general social motivations (such as relationship building) than escapism.

663 It should be noted that our study was not designed to provide support for
664 or against virtual world experimentation as a method in absolute terms. In-
665 stead, we adopted a less ambitious research question regarding its ability to
666 reproduce the results of traditional experimentation in physical laboratories
667 with standard subjects. As a result, the absence of observed behavioural dif-
668 ferences between the two environments does not necessarily make a case for
669 virtual experiments *per se*, but rather suggests they may be a valid alterna-
670 tive to traditional method, subject to similar methodological advantages and
671 limitations. Conversely, the presence of such differences would not necessarily
672 invalidate virtual experimentation to the extent that the standard physical
673 laboratory method is not without imperfections. As a result, these method-
674 ological issues remain and may benefit from renewed debate in the context of
675 virtual experimentation.

676 While the above suggests that virtual world experimentation has potential as
677 an economical and practical alternative to standard laboratory experiments,
678 there are certain disadvantages associated with virtual worlds as experimental
679 platforms which suggest that their suitability depends on the type of experi-
680 ment planned. For instance, studies that consider the effects of physical sig-
681 nals or depend on recruiting specific types of subjects will find little value in
682 virtual experimentation. On the other hand, suitably adapting experimental
683 procedures to the virtual world environment makes it possible to effectively
684 and cheaply recreate many standard decision tasks. In addition, virtual world
685 users appear to constitute suitable subject pools to the extent that they dis-
686 play many of the economic behaviours associated with standard subjects. The
687 future development of this technology will further increase the sophistication
688 of the virtual experimental platform.

689 References

- 690 Anderhub, V., Müller, R., Schmidt, C., 2001. Design and evaluation of an
691 economic experiment via the internet. *Journal of Economic Behavior and*
692 *Organization* 46 (2), 227–247.
- 693 Andreoni, J., 1988. Why free ride? Strategies and learning in public goods
694 experiments. *Journal of Public Economics* 37, 291–304.

- 695 Andreoni, J., 1995. Cooperation in public-goods experiments: Kindness or
696 confusion? *American Economic Review* 85 (4), 891–904.
- 697 Bainbridge, W. S., 2007. The scientific research potential of virtual worlds.
698 *Science* 317, 472–476.
- 699 Bloomfield, R. J., 2007. Worlds for study: Invitation - virtual worlds for study-
700 ing real-world business (and law, and politics, and sociology, and....). SSRN
701 eLibrary.
- 702 Bornstein, G., Gneezy, U., Nagel, R., 2002. The effect of intergroup competi-
703 tion on group coordination: An experimental study. *Games and Economic*
704 *Behavior* 41, 1–25.
- 705 Buchan, N. R., E., J., Croson, R. T. A., 1997. Culture, power and legitimacy:
706 Contrasting influences on fairness beliefs and negotiation behavior in Japan
707 and the United States, University of Wisconsin Department of Marketing
708 Working Paper.
- 709 Cagnina, M. R., Poian, M., 2007. How to compete in the metaverse: The
710 business models in second life. Tech. rep.
711 URL <http://ssrn.com/abstract=1088779>
- 712 Camerer, C. F., 1987. Do biases in probability judgment matter in markets?
713 *American Economic Review* 77 (5), 981–997.
- 714 Camerer, C. F., 1997. Progress in behavioral game theory. *Journal of Economic*
715 *Perspectives* 11 (4), 167–188.
- 716 Camerer, C. F., 2003. *Behavioral Game Theory*. Princeton University Press,
717 Princeton, NJ.
- 718 Carpenter, J. P., Burks, S., Verhoogen, E., 2005. Comparing students to work-
719 ers: The effect of social framing on behavior in distribution games. In: Car-
720 penter, J. P. (Ed.), *Field Experiments in Economics*. Vol. 10. Elsevier, Ch. 7,
721 pp. 261–290.
- 722 Cason, T. N., Mui, V. L., 1998. Social influence and the strategy method in the
723 sequential dictator game. *Journal of Mathematical Psychology* 42, 248–265.
- 724 Castranova, E., 2005. *Synthetic Worlds: The Business and Culture of Online*
725 *Games*. University of Chicago Press, Chicago, IL.
- 726 Castranova, E., 2006. On the research value of large games: Natural experiments
727 in *Norrath and Camelot*. *Games and Culture* 1, 163–186.
- 728 Chuah, S. H., Hoffmann, R., Jones, M. K., Williams, G. A., 2006. An economic
729 anatomy of culture: Attitudes and behaviour in inter- and intra-national
730 ultimatum game experiments, *Occasional Papers* 13, Industrial Economics
731 Division, Nottingham University Business School.
- 732 Chuah, S. H., Hoffmann, R., Jones, M. K., Williams, G. A., 2007. Do cultures
733 clash? Evidence from cross-national ultimatum game experiments. *Journal*
734 *of Economic Behavior and Organization* 64, 35–48.
- 735 Crawford, V. P., 1995. Adaptive dynamics in coordination games. *Economet-*
736 *rica* 63 (1), 103–143.
- 737 Devetag, G., 2005. Precedent transfer in coordination games: An experiment.
738 *Economics Letters* 89, 227–232.
- 739 Duffy, J., Nagel, R., 1997. On the robustness of behaviour in experimental

- 740 ‘beauty contest’ games. *Economic Journal* 107, 1684–1700.
- 741 Erev, I., Roth, A. E., 1998. Predicting how people play games: Reinforce-
742 ment learning in experimental games with unique, mixed strategy equilibria.
743 *American Economic Review* 88 (4), 848–881.
- 744 Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods ex-
745 periments. *American Economic Review* 90 (4), 980–994.
- 746 Forsythe, R., Horowitz, J. L., Savin, N. E., Sefton, M., 1994. Fairness in simple
747 bargaining experiments. *Games and Economic Behavior* 6, 347–369.
- 748 Guest, T., 2007. *Second Lives*. Random House, London, UK.
- 749 Harbaugh, W. T., Krause, K., Liday, S. G., 2003. Bargaining by children,
750 working paper.
- 751 Harrison, G. W., List, J. A., 2004. Field experiments. *Journal of Economic*
752 *Literature* 42 (4), 1009–1055.
- 753 Healy, P. J., 2006. Learning dynamics for mechanism design: An experimen-
754 tal comparison of public goods mechanisms. *Journal of Economic Theory*
755 129 (1), 114–149.
- 756 Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. (Eds.),
757 2004. *Foundations of Human Sociality: Economic Experiments and Ethno-*
758 *graphic Evidence from fifteen Small-Scale Societies*. Oxford University
759 Press, Oxford, UK.
- 760 Ho, T., Camerer, C., Weigelt, K., 1998. Iterated dominance and iterated best
761 response in experimental ‘p-beauty contests’. *American Economic Review*
762 88 (4), 947–969.
- 763 Hoffman, E., McCabe, K., Shachat, K., Smith, V. L., 1994. Preferences, prop-
764 erty rights and anonymity in bargaining games. *Games and Economic Be-*
765 *havior* 7, 346–380.
- 766 Hoffman, E., McCabe, K., Smith, V. L., 1996. Social distance and other-
767 regarding behavior in dictator games. *American Economic Review* 86 (3),
768 653–660.
- 769 Knez, M., Camerer, C., 1994. Creating expectational assets in the laboratory:
770 Coordination in ‘weakest-link’ games. *Strategic Management Journal* 15,
771 101–119.
- 772 Kovalchik, S., Camerer, C. F., Grether, D. M., Plott, C. R., Allman, J. M.,
773 2005. Age and decision making: A comparison between neurologically
774 healthy elderly and young individuals. *Journal of Economic Behavior and*
775 *Organization* 58, 79–94.
- 776 Kurz, M., 2008. Beauty contests under private information and diverse beliefs:
777 How different? *Journal of Mathematical Economics* 44 (7-8), 762–784.
- 778 Linden-Labs, 2008. Second life website.
779 URL <http://secondlife.com/>
- 780 Nagel, R., 1995. Unraveling in guessing games: an experimental study. *Amer-*
781 *ican Economic Review* 85 (5), 1313–1326.
- 782 Noveck, B. S., 2004. Introduction: The state of play. *New York Law School*
783 *Law Review* 49 (1), 1–18.
- 784 Oosterbeek, H., Sloof, R., van de Kuilen, G., 2004. Cultural differences in

- 785 ultimatum game experiments: evidence from a meta-analysis. *Experimental*
786 *Economics* 7, 171–188.
- 787 Rokeach, M., 1973. *The Nature of Human Values*. Free Press, New York, NY.
- 788 Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., Zamir, S., 1991. Bargaining
789 and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An
790 experimental study. *American Economic Review* 81 (5), 1068–1095.
- 791 Schotter, A., Weiss, A., Zapater, I., 1996. Fairness and survival in ultimatum
792 and dictatorship games. *Journal of Economic Behavior and Organization*
793 31, 37–56.
- 794 Schwartz, S. H., 2001. Value hierarchies across cultures: Taking a similarities
795 perspective. *Journal of Cross-Cultural Psychology* 32 (3), 268–290.
- 796 Schwartz, S. H., 2002. A proposal for measuring value orientations across
797 nations. In: *European Social Survey Questionnaire Development Report*.
798 European Social Survey, London, UK.
- 799 Van Huyck, J., Battalio, R., Beil, R., 1990. Tacit coordination games, strategic
800 uncertainty, and coordination failure. *American Economic Review* 80, 234–
801 248.
- 802 Yee, N., 2006. The demographics, motivations and derived experiences of users
803 of massively-multiuser online graphical environments. *Presence: Teleopera-*
804 *tors and Virtual Environments* 15, 309–329.

Virtual world experimentation: An exploratory study

Abstract

We explore the scientific potential of virtual worlds for experimental economics in terms of the subject pools and experimental platforms they present. Our results offer tentative, qualified support for virtual world experimentation. Overall, the behaviour of virtual subjects recruited, incentivised and observed within *Second Life* across a range of five standard experimental games was not found to differ significantly from established standard results. In addition, we identify certain methodological opportunities and challenges which confront virtual world experimenters.

Key words: virtual worlds, experiments

JEL classification: C72; C88; C99; Z13

1 Introduction

2 Social scientists are becoming increasingly interested in virtual worlds, three-
3 dimensional environments in which communities of networked individuals in-
4 teract (Castranova, 2005; Bainbridge, 2007; Bloomfield, 2007). There are two
5 reasons. First, the growing number of users and the scope and nature of socio-
6 economic activity between them are seen as interesting phenomena that merit
7 investigation in their own right (Castranova, 2005). Virtual worlds present
8 evolving cultures with independent social institutions that are becoming more
9 significant to society at large (Noveck, 2004). In economic terms, their evo-
10 lution from specialised video game networks to general social platforms has
11 generated a global industry of firms that leverage installed user bases for sub-
12 scription fees, advertising opportunities or virtual support services (Cagnina
13 and Poian, 2007). Many virtual worlds have evolving economies with fully con-
14 vertible currencies as well as functioning financial, labour and product markets
15 that are capable of producing a host of micro and macroeconomic phenomena
16 (Guest, 2007).

17 Second, the computer technology underlying virtual worlds provides novel
18 methods of conducting social science research (Bainbridge, 2007). To begin
19 with, it facilitates the economical and large-scale recruitment of diverse sub-
20 jects from different cultural-geographical and socio-economic groups for par-
21 ticipation in interviews, focus groups, surveys or experiments. In addition,
22 it affords control of the environment in which they decide and interact that
23 can be used to manipulate decision conditions, observe behaviour and collect
24 data. Conversely, however, both these features also present potential method-
25 ological problems. As subjects, virtual world users may not reflect standard
26 populations in terms of demographic or cultural characteristics and there-
27 fore may display different behaviours. The electronic interface that moderates
28 communication and interaction between them precludes physiological signals
29 and proximity that moderate economic behaviour in physical settings. Vir-
30 tual world culture, social institutions and conventions that evolve as a result
31 may shape economic interactions in ways that differ from traditional social
32 settings. The anonymity of the interface may hamper quality control in the
33 data collection process.

34 The current study is intended as a first, exploratory step towards the method-
35 ological issue. While virtual worlds may provide useful research tools for a
36 number of social science disciplines, we concentrate here on their potential as
37 platforms for designing and conducting economic experiments, an area which
38 may be especially conducive to benefit from the new methods virtual worlds of-
39 fer (Bainbridge, 2007; Castranova, 2006). Traditional experimental economics
40 involves testing economic theories by observing the incentivised decisions of
41 representative subjects under choice conditions systematically manipulated in

42 laboratory settings. Virtual worlds may provide opportunities for methodolog-
43 ical innovation here. The discipline has recently begun to broaden its scope
44 by exploring new methods and applications outside the standard controlled
45 laboratory environment commonly populated by Western student subjects.
46 There are two related ways in which experimentalists are trying to improve
47 the realism of the behaviour they observe. First, field studies in naturalistic
48 settings are being proposed as a way of avoiding the distorting effects artificial
49 laboratory settings may have on subject behaviour (Harrison and List, 2004).
50 Second, new recruitment techniques and sampling locations are being used
51 to overcome the reliance of experimentalists on Western university students
52 to generate results (Anderhub et al., 2001; Henrich et al., 2004, e.g.). Virtual
53 worlds may give an opportune impetus to both of these concerns. First, due
54 to their computerised interfaces, they may provide relatively controlled en-
55 vironments for conducting experiments while remaining within a naturalistic
56 setting familiar to subjects. Second, virtual worlds may be inhabited by a
57 wider cross section of people such that sampling from different cultures and
58 more heterogeneous backgrounds may be possible in a single location acces-
59 sible to experimentalists. In this sense, virtual worlds may bridge the gap
60 between laboratory experiments and field studies, allowing researchers to use
61 representative subjects in more natural environments to study the relationship
62 between the conditions of interaction and the evolution of social institutions
63 in a controlled manner.

64 We assess to what extent virtual worlds can be used in this context. We ap-
65 proach the issue in two ways, by replication and by observation. First, virtual
66 world experimentation can be a useful, alternative experimental tool to the
67 extent that the results it generates for particular tasks and conditions are the
68 same as those generated by traditional experimental methods. We assess this
69 aspect by conducting virtual experiments with a range of standard tasks in
70 standard conditions and comparing virtual subject behaviour with that of tra-
71 ditional pools reported in existing work. The suitability of virtual experimenta-
72 tion as an alternative would be supported to the extent that no differences are
73 found. As the observed subject behaviour may be related to their underlying
74 culture, demographics and values, we also used a survey instrument to collect
75 data on these which can be compared to standard populations. The difference
76 or similarity of virtual users to these provides additional insight into their
77 suitability as experimental subjects representative of economic agents gener-
78 ally. This first part of our approach tests the scope virtual worlds hold for
79 traditional economic experimentation, rather than for new avenues of experi-
80 mental research they may promise. We conceive of it as measuring the ‘output’
81 of the virtual experimentation method. The second part of our approach is
82 more qualitative and focuses on its ‘input’ side. This involves gathering in-
83 formal insights about the practical feasibility of economic experimentation in
84 virtual environments from the process of conducting experiments. We hoped
85 to learn by observation to what extent virtual worlds can provide a suitable

86 platform for experimental research generally, what the advantages and disad-
87 vantages are, and what modifications may be made to render virtual worlds
88 more amenable to experimentation. This second part may also provide in-
89 sights into what opportunities for new research approaches or methods virtual
90 worlds hold.

91 The rest of the paper proceeds as follows. In the next section, we discuss
92 the features of virtual worlds, their significance for experimentalists and our
93 procedure of methodologically assessing them. The results we obtained are
94 reported in section 3. Section 4 discusses our general observations from the
95 experiment in terms of the methodological issues we consider. The final section
96 contains concluding remarks.

97 **2 Virtual Experimentation**

98 *2.1 Virtual Worlds*

99 While there is considerable variation between the many alternative virtual
100 worlds that exist, they typically reproduce features of the physical world such
101 as a three-dimensional topography containing virtual objects obeying simu-
102 lated physical laws as well as the possibility of communication, social interac-
103 tion and economic exchange between users virtually represented by *avatars*.
104 We chose Second Life (SL, see Linden-Labs 2008) as the virtual platform for
105 our study. At the time of writing (November 2008), SL has over 15.7 million
106 registered avatars.¹ Accounting for multiple and dormant registrations, there
107 are an estimated one million regular users who spend over twenty million hours
108 logged in per month. Between twenty and thirty thousand users are online at
109 any one time. In terms of demographics, the majority of these are from popu-
110 lous and industrialised countries including the USA, the UK, Germany, Brazil,
111 France and Japan, with a median age of 36 and 57% being male.

112 SL is divided into individual sectors with topographical features in which
113 avatars can operate, including oceans, rivers, mountains and beaches as well
114 as flora. A typical location is displayed in figure 1. Avatars are capable of loco-
115 motion, including walking, running and flying and are immune to destruction.
116 They communicate using instant text messaging (IM) and can signal voice
117 intonation such as whispering and shouting as well as use gestures and body
118 language. Public IM can be received by all avatars in the vicinity, while private
119 IM is transmitted only between two avatars irrespective of location. Internet

¹ Economic and general statistics concerning SL are available at:
<http://secondlife.com/whatis/economy.php> and <http://blog.secondlife.com/>.



Fig. 1. Typical SL-screenshot showing the user's avatar (male foreground figure), the surrounding SL-environment and interface controls along the bottom.

120 telephony has recently been introduced to SL. Users can edit the appearance
 121 of avatars in terms of physical features, clothing and accessories. As a result,
 122 avatars can assume the form of humans, animals, fantasy creatures or objects.
 123 Avatars are associated with user accounts that include money balances in Lin-
 124 den dollars (L\$) which can be bought from or sold to Linden Lab, the creators
 125 and owners of SL, at a relatively stable exchange rate of about 270 L\$ per
 126 1 U.S. dollar. A total of 5.3 billion L\$ (U.S. \$19.7 million) are currently in
 127 circulation. SL provides an interface feature that allows immediate and direct
 128 account-to-account transfers. These balances can be used to purchase a port-
 129 folio of tradable virtual objects including land, buildings, vehicles, clothing,
 130 accessories and tools.

131 2.2 *Experimental Economics*

132 Virtual worlds such as SL may have potential as powerful new platforms for
 133 designing and conducting experimental research. Bainbridge (2007) makes the
 134 following case:

135 Virtual worlds such as SL provide environments and tools that facilitate
 136 creating online laboratories that can automatically recruit potentially thou-
 137 sands of research subjects, over a period of months, at low cost. SL offers
 138 scripting and graphics tools that allow anyone to build a virtual laboratory
 139 building, functioning equipment to run the experiment, and incentives to

140 motivate participation. (p. 473)

141 Conversely, however, the very technology that generates these advantages may
142 give rise to a number of *a priori* concerns about virtual experimentation. Prin-
143 cipally, experimenters know little about the identity or state of the subjects
144 who control the participating avatars. This may make it difficult to recruit
145 appropriate subjects, to ensure discipline in the virtual laboratory, to prevent
146 repeat participation and subject collusion and to engender subjects' trust and
147 confidence in the experiment. There is a possibility of demographic or cul-
148 tural idiosyncrasies of virtual subjects generally. This may generate a sample
149 bias that renders virtual experimentation inappropriate to test general eco-
150 nomic theories. They may have more hedonistic or short-term tendencies or
151 show less conformity than the average person. In addition, virtual behaviour
152 is not moderated by physical presence and may therefore not be comparable
153 to traditionally-generated results.

154 2.3 *Experimental Design*

155 The purpose of our study is to conduct experiments within SL to assess the
156 overall feasibility of virtual experimentation. Our approach is to gauge to what
157 extent the behaviour and values of virtual subjects conform to those of stan-
158 dard subjects. In the following, we outline the general working procedure that
159 we developed and deployed over the course of our experiments in terms of five
160 stages of which individual experimental sessions consisted. All our experimen-
161 tal sessions were conducted during standard GMT working hours between July
162 and November 2007. Experimental instructions are available upon request.

163 In the recruitment stage, we solicited participation by approaching online users
164 *in situ* immediately prior to a particular experimental session in the following
165 manner. Half an hour before a scheduled experimental session, we used a search
166 feature in the SL-inteface to identify the currently busiest locations in terms of
167 number of avatars present (excluding locations with an adult thematic focus).
168 Next, each of the three experimenters used their avatar to access one of these
169 locations and to address groups of avatars gathered there using public IM with
170 a standardised recruitment message. This message was in English and stated
171 our institutional affiliation and general information about the nature of the
172 task, its duration and incentivisation. Whenever interested users responded,
173 we answered any additional questions and informed volunteers of the time
174 and venue of the session. This process was repeated for a number of locations
175 and avatar groups in each until the recruitment of the desired number of
176 participants was complete. The thirty minute period was in almost all cases
177 sufficient to recruit between four and seven subjects.



Fig. 2. A typical experimental session in progress. The experimenters' avatars are standing.

178 Participants were transported to our virtual experimental laboratory in a ded-
 179 icated virtual building with controllable access rights and purpose-built labo-
 180 ratory furniture. In the briefing stage, subjects who have arrived (typically in
 181 groups between two and seven depending on the task) were given virtual doc-
 182 uments containing general information on experimental etiquette, anonymity,
 183 confidentiality and incentivisation. The two to three experimenters present at
 184 all times communicated with subjects using either public or private (i.e. one-
 185 to-one) IM. Once they have finished reading the briefing documents, subjects
 186 were asked to occupy cubicles that were purpose-built to restrict their vision
 187 and communication in order to prevent collusion between them. They were
 188 then given virtual documents containing the experimental instructions and
 189 a comprehension quiz. The decision task stage commenced after all subjects
 190 completed the quiz successfully. Experimenters instructed individually when
 191 subjects were initially unable to do so. Subjects communicated their decisions
 192 to the lead experimenter and received feedback via private IM. Next, in the
 193 survey stage, subjects were sent the URL of a webform containing a values
 194 survey as well as some demographic questions which they had to fill out. In the
 195 final, payment stage of the experimental session, subjects were paid earnings in
 196 \$L on the spot using the SL payment transfer feature. A typical experimental
 197 session in progress is shown in figure 2.

198 Table 1 provides some general information about the decision tasks of our ex-
 199 periments. Our choice of tasks was guided by our objective to assess whether a
 200 virtual subject pool may be appropriate in testing economic theories. In par-
 201 ticular, we wanted to examine whether virtual behaviour conforms to estab-
 202 lished results generated in conventional experimentation. As a result, we chose

203 the ultimatum (UG), dictator (DG), public good (PGG), guessing (GG) and
 204 minimum effort (MEG) games. Previous experimental results for all of these
 205 standard games abound for a variety of conditions as well as demographic and
 206 cultural groups and provide ready benchmarks for our own results. They also
 207 permit eliciting a broad spectrum of different types of strategic choice. In the
 208 following, we do not explain or analyse these standard games in detail, but
 209 report data from our and those previous studies most appropriate for compar-
 210 ison. We also report results from tests of differences in means, medians and
 211 overall distributions between them using t -tests, Mann-Whitney U (MW) and
 212 Kolmogorov-Smirnov Z (KS) tests respectively. While means tests can indi-
 213 cate differences between the overall behavioural propensities in two pools of
 214 subjects, distribution tests can also reveal differences in the incidence of a vari-
 215 ety of behaviours when average behaviour does not differ. For experimental
 216 tasks with multiple decision rounds, we also used regression analysis to test
 217 for differences with previous results. In particular, we pooled available data
 218 from our own and the previous study used as a comparator and estimated the
 219 following regression equation:

$$220 \quad Y_i^t = \alpha + \beta Y_i^{t-1} + \gamma X_i + \delta n_i \quad (1)$$

221 where Greek letters represent constant and parameters, Y is observed be-
 222 haviour, t the task round, n experimental group size and X a dummy variable
 223 for the comparator study. No differences between SL and comparator study
 224 behaviour exist to the extent that the coefficient for the latter variable is in-
 225 significant. The inclusion of the lagged variable on the right-hand side was
 226 intended to reduce omitted-variable bias in our model. In particular, it is well
 227 established that simple learning processes may explain some changes in be-
 228 haviour over time in specific game and choice contexts (see, e.g., Camerer
 229 1987, Erev and Roth 1998). As a result, we opted for a specification simi-
 230 lar to a partial adjustment model, where the behaviour in the current period
 231 is adjusted to that in the previous one. These kinds of dynamic model have
 232 been previously applied to the three games for which we seek to estimate be-
 233 haviour, i.e. the PGG (Healy, 2006), the GG (Kurz, 2008) as well as the MEG
 234 (Crawford, 1995).

235 It should be noted that our design makes no provision for establishing a control
 236 treatment by replicating our virtual experiments in a standard physical set-
 237 ting with otherwise identical experimental parameters. While this alternative
 238 has certain advantages, our approach was to rely instead on the replicability
 239 of existing studies and to design virtual experiments that mirror their task
 240 conditions such as to permit using their results as a comparator.

241 An additional avenue for testing subject pool suitability is to survey and com-
 242 pare our subjects' values and demographics to those of standard experimental
 243 subjects and general populations. Values provide a measurement of a respon-

Task	UG	DG	GG	PGG	MEG	ESS
Subjects (N)	64	60	31	32	31	113
Subjects per session (n)	4-5	4-5	3-7	4	5-6	n/a
Average pay (U.S. \$)	5.25	1.95	2.30	20.15	8.25	3.85
Duration (minutes approx.)	25	10	25	35	20	10
Rounds (r) or questions	1	1	10	10	10	21

Table 1

Summary statistics for experimental games and survey.

244 dent's cultural orientation and are known to affect behaviour (Rokeach, 1973;
 245 Chuah et al., 2006). We used the human values survey designed by Shalom
 246 Schwartz for the European Social Survey (ESS) project (Schwartz, 2002). Like-
 247 wise, a number of demographics such as gender, age, and nationality are known
 248 to affect behaviour (see Camerer 2003 for an overview). In the following sec-
 249 tions, we report the results we obtained from the game tasks and survey.

250 3 Experimental Results

251 3.1 Subject Demographics

252 Subjects' basic demographical data are summarised in figure 3. The average
 253 age of respondents was 32, with the youngest at 18 and the oldest at 64. Com-
 254 pared with the general population of the European Union (EU), the age range
 255 20-40 years was over represented, an expected result given the technological
 256 and cultural status of virtual worlds. In line with SL generally, most subjects
 257 were from populous Western nations, although UK and European countries
 258 were somewhat over-represented in our sample. The reason may lie in using
 259 the English language and our institutional affiliation in recruitment. Recruit-
 260 ing during GMT daytime hours further bias sample selection in terms of time
 261 zone. In terms of gender, exactly half of our respondents were male.

262 3.2 Ultimatum Game

263 Separate sessions with UG-proposers and responders were conducted on 6,
 264 25 and 26 July 2007. In the proposer sessions, subjects were given the task to
 265 decide how to share L\$3000 (U.S. \$11.50) with a randomly-chosen co-player
 266 from a responder session who had the choice to accept or reject the split,

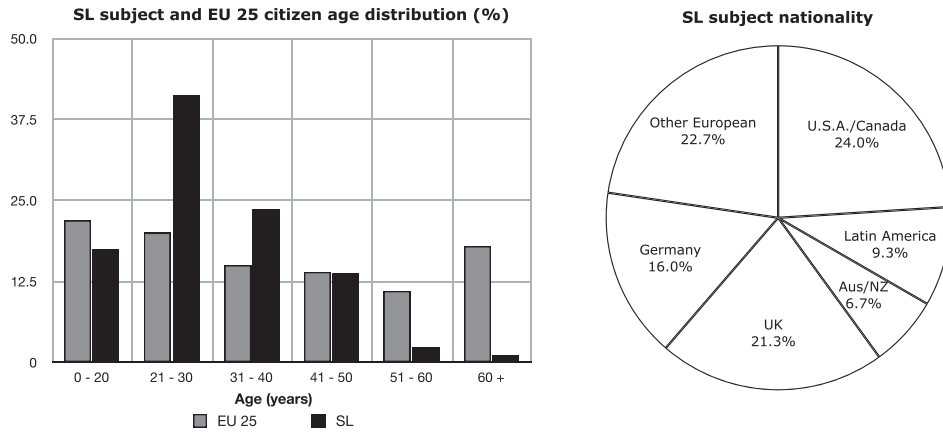


Fig. 3. Age and nationality distribution of SL-subjects.

267 resulting in the proposed shared being paid out or neither player receiving
 268 anything.

269 Although there is little evidence for stake size effects in the UG (see Camerer
 270 2003), we aimed for comparability by using a stake in the U.S. \$10-15 interval
 271 used in many previous studies, as well as for easy mental divisibility. Theory
 272 predicts that, because instrumentally-rational responders should accept any
 273 share of the stake, rational proposers should offer the minimum. However,
 274 proposers in previous studies offer in the region of 42-48% (see table 2.2.
 275 in Camerer 2003), reflecting a mixture of altruistic and strategic thinking on
 276 their part (Forsythe et al., 1994). In standard task conditions and subject pools
 277 recruited in industrialised nations, UG-results are relatively robust. Roth et al.
 278 (1991) (RPOZ) found little difference between offers made by urban subjects
 279 recruited in the U.S. (RPOZ 1), Tokyo (RPOZ 2), Yugoslavia and Israel.
 280 However, alternative cultural and demographic characteristics can generate
 281 differences (Camerer, 2003; Oosterbeek et al., 2004). Buchan et al. (1997) and
 282 Chuah et al. (2007) (CHJW) identified slightly but significantly higher offers
 283 of South-East Asian subjects potentially linked to their collectivist values.
 284 Henrich et al. (2004) found a much wider range of offers (between 25-57%)
 285 in a series of experiments with traditional, small-scale societies across the
 286 developing world.

287 Table 2 reports summary statistics of UG bargaining by SL-subjects compared
 288 with behaviour reported by RPOZ (1 and 2), by Hoffman et al. (1994) for
 289 U.S. subjects (HMSS) and by CHJW for UK subjects. The SL mean offer
 290 is 45.73% of the stake with a modal offer of half. These central tendencies
 291 in the proposals are very similar to those reported for comparable samples.
 292 Figure 4 shows the distributions of offers in all these experiments. With the
 293 exception of a small number of hyper-fair outliers among SL-subjects, the
 294 distribution we found is also very similar to those in the previous studies.
 295 Statistical tests bear these observations out. As the UK formed the largest

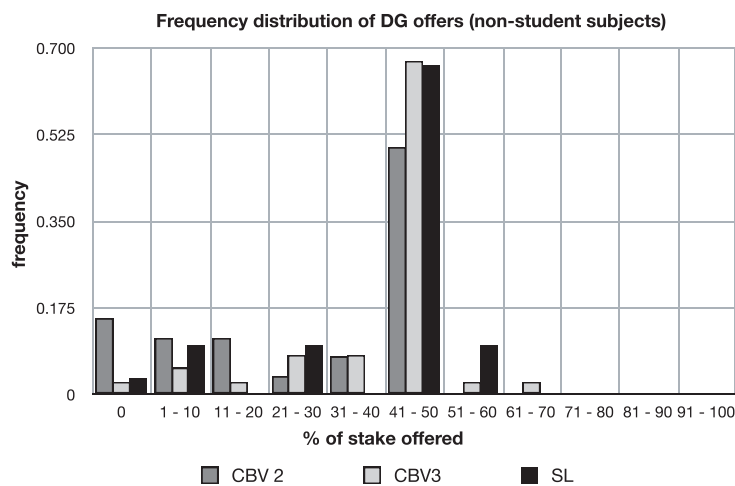
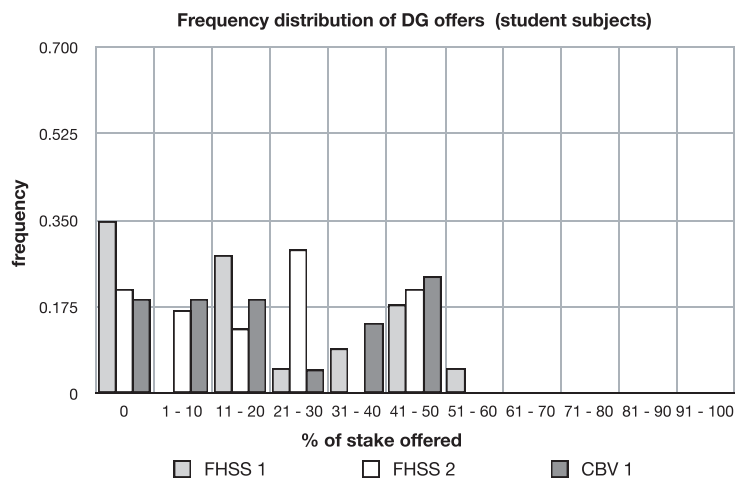
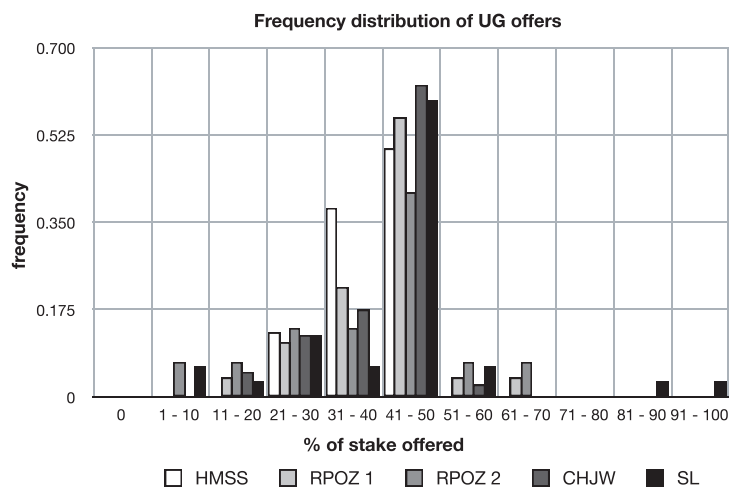


Fig. 4. Distribution of UG and DG offers in SL as well as in selected previous studies.

	SL	HMSS	RPOZ 1	RPOZ 2	CHJW
$N/2$	32	24	27	29	40
Stake	11.50	10	10	10	16
Offers					
Mean	45.73	44	45	45	44
Mode	50	50	50	50	50
St. Dev.	18.6	7.2	9.6	21.0	9.5
Rejections					
% of offers < 20%	33.33	-	-	50	-
% of all offers	6.25	8.3	22	24	15

Table 2

Summary statistics of ultimatum game offers (in % of the U.S. \$ stake) and rejections for $N/2$ subject pairs in SL as well as in selected previous studies.

296 national group among our subjects (see section 3.7), we used UK subject data
 297 from CHJW as a comparator for our findings. No differences in the mean
 298 ($t=0.216$, $p=0.829$), median ($U=2706.5$, $p=0.422$) or distribution ($Z=0.595$,
 299 $p=0.870$) of offers were found between their and SL subjects.

300 3.3 Dictator Game

301 DGs were also conducted in separate sessions for proposers and responders,
 302 except that responders were not given the opportunity to accept or reject
 303 offers. The sessions were conducted on 27 and 31 July 2007. As a number of
 304 previous studies employed stakes divisible by 10, and since stake size effects
 305 are not noticeable between studies with significantly different stakes (see table
 306 3), we opted for a stake size of 1000 \$L (U.S. \$3.90). The DG was originally
 307 conceived as a way of separating altruistic and strategic motives in UG-offer
 308 behaviour (Forsythe et al., 1994). While instrumentally rational players should
 309 keep all of the stake, experimental subjects offer in the region of 20-35% to
 310 responders, reflecting altruistic preferences. DG-behaviour is sensitive to a host
 311 of experimental conditions such as anonymity, source and destination of the
 312 stake (see Camerer 2003 for an overview). In addition, subject demographics
 313 influence offers.

314 Table 3 reports summary statistics of SL-dictator behaviour compared to sub-
 315 jects in comparable studies by Forsythe et al. (1994) (FHSS) and Carpenter
 316 et al. (2005) (CBV). Figure 4 displays the distributions of offers in the experi-
 317 ments reported there. The first two of these studies (centre panel of the figure)

	FHSS 1	FHSS 2	CBV 1	CBV 3	SL	CBV 2
$N/2$	24	45	21	26	30	37
Stake	10	5	100	100	3.90	100
Offers						
Mean	24	24	25	33	43	45
Mode	30	0	50	50	50	50
Median	25	20	20	45	50	50
St. Dev.	17.68	20.44	19	20	16.17	12

Table 3

Summary statistics of dictator game offers (in % of the U.S. \$ stake) for $N/2$ subject pairs in SL as well as reported in selected previous studies.

318 report offers made by standard college student subjects which tend to be in
 319 the region of 23-24% of the stake (see also Hoffman et al. 1996, Cason and
 320 Mui 1998), although some studies, such as Schotter et al. (1996), have found
 321 offers close to 40%. Of particular interest to us is the study by CBV, who
 322 identified marked differences in DG offer levels based on age and experimental
 323 location (bottom panel of figure 4). In their study, they compare offers made
 324 by students (average age: 19 years) in standard college settings (CBV 1), by
 325 older community college students (27, CBV 2) and by workers in a warehouse
 326 setting (37, CBV 3).

327 The data show the DG offers made by SL-subjects to be higher than those
 328 reported in standard college settings, but similar to those made by older sub-
 329 jects in CBV. These results reflect the greater average age of our subjects (see
 330 section 3.7) and the fact that DG-offers are sensitive to age (Harbaugh et al.,
 331 2003). Previous and current DG-results pertaining to older subjects are shown
 332 in the bottom panel in figure 4. It is also noteworthy that in our experiment,
 333 proposers communicated their offers to the experimenter directly using pri-
 334 vate IM rather than using forms collected and delivered in stacks by monitors
 335 as tends to be practiced in physical locations. Our treatment provides more
 336 scope for social influence and demand effects that would be expected to raise
 337 offers.

338 The age similarity between warehouse workers in CBV 3 to our own SL-
 339 subjects provides us with an appropriate benchmark for the comparison of
 340 DG-behaviour. No statistically significant differences were found between the
 341 means ($t=-0.700$, $p=0.485$) medians ($U=981.5$, $p=0.823$) and distributions
 342 ($Z=0.383$, $p=0.999$) of DG offer data in these two pools.

343 3.4 *Public Good Game*

344 The PGG sessions were conducted on 25 October and 2 November 2007. In
 345 them, subjects in groups of $n = 4$ were asked to divide a stake of L\$400 (U.S.
 346 \$1.50) between a private and a group fund and explained that their total earn-
 347 ings would be their private allocation plus $a = 0.4$ times the total of all group
 348 allocations. This was repeated $r = 10$ times. The parameter values for n , r
 349 and a were chosen with comparability with other studies in mind (see table 4).
 350 The PGG is a n -person version of the prisoner's dilemma and pits subjects'
 351 self-serving motives against their desire to further the benefit of the group.
 352 Instrumentally-rational play involves complete free-riding and allocating the
 353 whole endowment to the private fund. In repeated PGGs, players decisions
 354 may be guided both by strategic considerations of reciprocation and purely
 355 altruistic motives. A large literature exists that identifies the experimental
 356 conditions that elicit cooperative behaviour. In general, subjects contribute
 357 positive amounts to the public good that steadily decline as the game is re-
 358 peated. The studies reporting PGG games under standard conditions serve as
 359 benchmarks for the behaviour of our SL-subjects. We compare the behaviour
 360 of SL-subjects with those in experiments with comparable conditions reported
 361 by Andreoni (1988, 1995) (A (88) and A (95)) as well as Fehr and Gächter
 362 (2000) (FG), who used values for parameter a of 0.5, 0.5 and 0.4 respectively.
 363 Table 4 reports summary statistics of SL-PGG behaviour compared to sub-
 364 jects in these three.

365 The top panel in figure 5 shows the average contribution to the group fund
 366 subjects made in SL and in the three previous studies over ten rounds. SL-
 367 subjects contribute marginally more than subjects in the other pools in all
 368 rounds. The average contribution decays over rounds in similar ways in all
 369 studies. The higher average we find is not unusual within the context of find-
 370 ings made using variegated subject pools. For instance, Henrich et al. (2004)
 371 report on PGGs played with traditional society subjects in many continents
 372 and find mean contribution rates to vary between 22 and 65%. The SL sub-
 373 jects differ from standard college students in a number of ways, age being one.
 374 Our result may also be due to the apparent greater altruism of SL-subjects
 375 compared with students we observed in the DG.

376 For our statistical tests of PGG behaviour, we chose A (95)'s Western student
 377 subject data as a benchmark. It should be borne in mind that this experiment
 378 differs from our study in two ways; the differences in experimental platform
 379 we are assessing, and the differences in subject demographics. We performed
 380 mean, median and distribution tests between the offers for each of the ten
 381 rounds played by A (95) and SL subjects (see table 5). Only one of the resulting
 382 thirty test statistics was significant ($Z_{n=10} = 1.370, p=0.047$). As the repeated
 383 testing procedure amplifies the probability of Type I errors, we also estimated

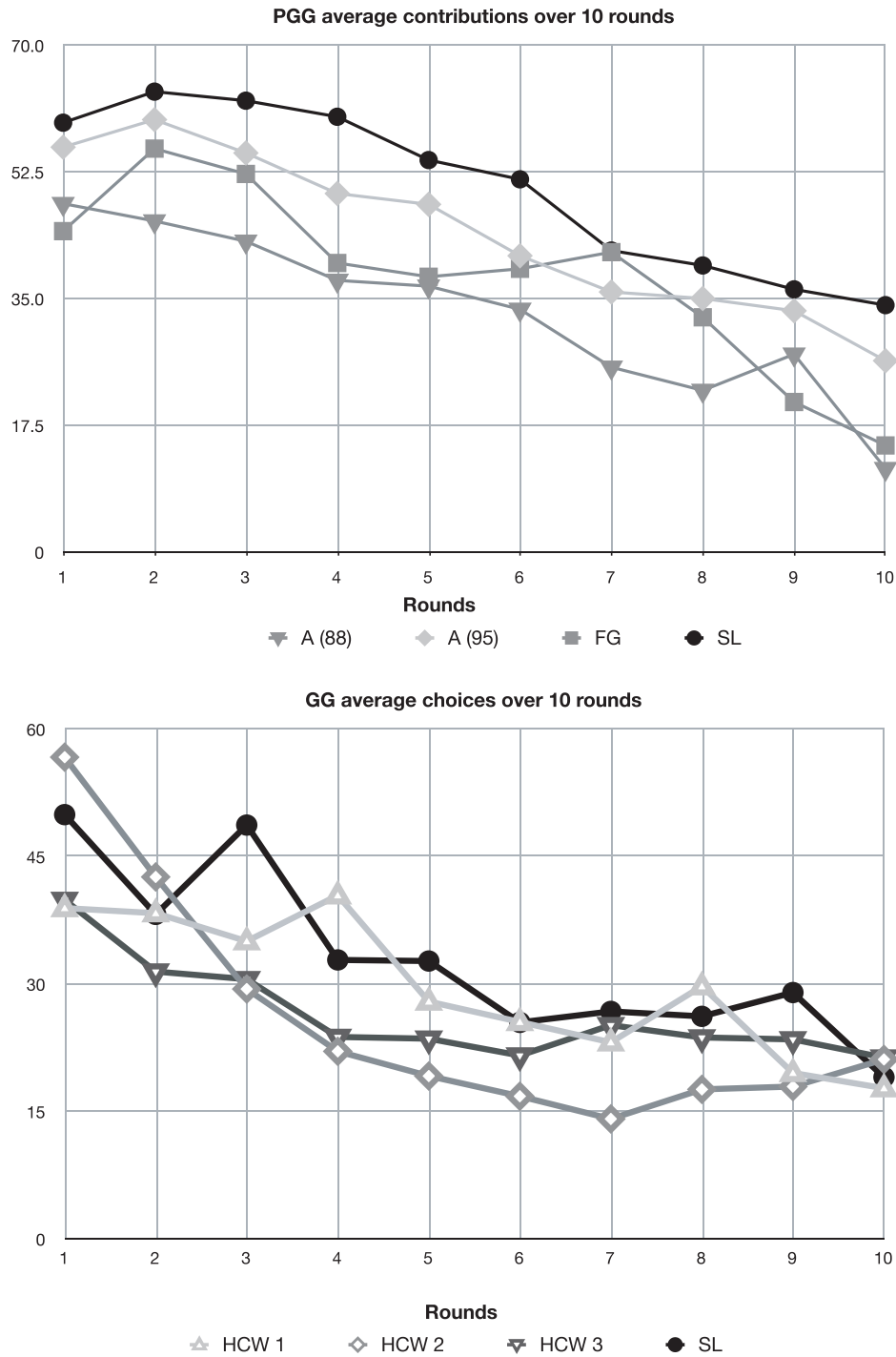


Fig. 5. Average subject decisions in GG and PGG over $r=10$ rounds in SL and selected previous studies.

384 equation 1 to compare the two data sets. The factor n could not be entered
 385 due its insufficient variation in the data set. The regression results are given
 386 in table 6 and show an insignificant coefficient for X , leading us to conclude
 387 that no behavioural differences are in evidence.

	A (88)	A (95)	FG	SL
N	30	40	24	32
n	5	5	4	4
r	10	10	10	10
Stake	0.50	0.60	0.86	1.50
α	0.5	0.5	0.4	0.4
Contributions				
Mean	33.20	44.09	37.94	50.34
Median	32.00	42.50	40.25	45.63
St. Dev.	21.65	27.47	16.89	22.54

Table 4

Summary statistics of public good game contributions (in % of the U.S. \$ stake averaged over r rounds) for N subjects playing in groups of n in SL as well as reported in selected previous studies. Stakes are given as U.S.\$-values of tokens subjects were asked to allocate per round.

388 3.5 Minimum Effort Game

389 The MEG sessions were conducted between 16 and 21 November 2007. In
 390 them, groups of $n = 5$ to 6 subjects were asked to choose an integer in the
 391 interval $[1, 7]$ and informed that payoffs would be determined by the smallest
 392 number chosen within the group according to the payoff matrix adapted from
 393 Van Huyck et al. (1990) (VBB) and shown in table 7. Each group played ten
 394 rounds of this game. Again, these parameter values are standard to the extent
 395 that they have been adopted by the majority of previous studies. The game
 396 has multiple equilibria in which all players make the same choice, which payoff
 397 dominate each other in turn with a unique Pareto-efficient equilibrium in every
 398 player choosing 7. The game represents situations where a group's ability to
 399 coordinate on the individually as well as collectively best outcome may be
 400 undermined by individuals' pessimistic expectations of others' reasoning. A
 401 typical example is punctuality (Camerer, 2003). While everyone arriving on
 402 time for a meeting is mutually the best outcome, an individual may arrive late
 403 to avoid a wait expecting others to also be late. After a number of meetings,
 404 such expectations may become increasingly self fulfilling as general punctuality
 405 disintegrates. Previous experimental evidence shows this kind of convergence
 406 on payoff-dominated outcomes to be dependent on the size of the group, the
 407 size of payoffs and information players receive about the choices of others.

408 Figure 6 shows the round-to-round changes in the choices and minimum

Task	r	t	MW U	KS Z
PGG	1	0.431 (0.667)	619.0 (0.808)	1.054 (0.216)
	2	0.499 (0.619)	611.5 (0.743)	0.949 (0.329)
	3	0.864 (0.391)	572.0 (0.436)	0.764 (0.603)
	4	1.231 (0.223)	536.5 (0.235)	1.001 (0.269)
	5	0.697 (0.488)	567.0 (0.403)	0.817 (0.517)
	6	1.231 (0.222)	534.5 (0.227)	0.870 (0.436)
	7	0.673 (0.503)	573.5 (0.446)	0.738 (0.648)
	8	0.568 (0.572)	544.5 (0.274)	1.370 (0.047**)
	9	0.372 (0.711)	567.5 (0.405)	0.817 (0.517)
	10	0.926 (0.358)	539.5 (0.240)	1.133 (0.153)
MEG	1	1.482 (0.141)	982.5 (0.139)	0.895 (0.452)
	2	1.218 (0.226)	1023.0 (0.236)	0.833 (0.491)
	3	1.927 (0.057)	931.0 (0.070*)	1.109 (0.171)
	4	2.660 (0.009***)	822.5 (0.011**)	1.353 (0.051*)
	5	1.449 (0.150)	986.0 (0.153)	0.713 (0.690)
	6	1.382 (0.170)	990.0 (0.162)	0.983 (0.289)
	7	1.571 (0.119)	955.5 (0.102)	0.888 (0.409)
	8	0.785 (0.435)	1059.0 (0.351)	0.951 (0.326)
	9	0.518 (0.606)	1073.5 (0.406)	1.042 (0.228)
	10	2.364 (0.020**)	841.0 (0.014**)	1.347 (0.053*)
GG	1	1.798 (0.078*)	219.0 (0.079*)	0.928 (0.355)
	2	0.091 (0.928)	305.0 (0.923)	0.478 (0.976)
	3	2.195 (0.033**)	212.5 (0.060*)	1.226 (0.099*)
	4	-1.090 (0.281)	268.5 (0.423)	0.821 (0.510)
	5	1.003 (0.321)	289.5 (0.692)	0.664 (0.771)
	6	0.032 (0.974)	280.5 (0.569)	0.703 (0.706)
	7	0.538 (0.593)	283.0 (0.602)	0.664 (0.771)
	8	-0.552 (0.583)	278.5 (0.543)	0.652 (0.788)
	9	2.107 (0.041**)	250.5 (0.250)	1.277 (0.077*)
	10	0.279 (0.781)	292.5 (0.735)	1.063 (0.209)

Table 5

Test statistics for differences in mean (t), median (U) and distribution (Z) of behaviour between SL subjects and those in selected previous studies for $r=10$ rounds. Corresponding p -values are given in parentheses. The symbols *, ** and *** denote significance at the 10, 5 and 1% levels respectively.

PGG	Estimate	<i>t</i> -value	<i>p</i> -value
Constant	14.24	6.39	0.000***
Y_{t-1}	0.67	22.63	0.000***
X	-2.47	-1.15	0.252
	R^2 (adj.)= 0.45	$F= 260.69$	$p=0.000***$
MEG	Estimate	<i>t</i> -value	<i>p</i> -value
Constant	1.22	1.36	0.17
Y_{t-1}	0.61	24.03	0.000***
n	0.06	0.43	0.664
X	-0.34	-2.12	0.034**
	R^2 (adj.)=0.39	$F=204.55$	$p=0.000***$
GG	Estimate	<i>t</i> -value	<i>p</i> -value
Constant	26.84	11.38	0.000***
Y_{t-1}	0.25	8.09	0.000***
n	-1.11	-2.82	0.005**
X	-1.36	-0.81	0.416
	R^2 (adj.)=0.09	$F=30.42$	$p=0.000***$

Table 6

Regression results for experimental behaviour across three tasks in SL and one comparator study respectively. The symbols *, ** and *** denote significance at the 10, 5 and 1% levels respectively.

409 choices averaged over experimental groups in SL and comparable previous
 410 studies of Knez and Camerer (1994) (KC), Bornstein et al. (2002) (BGN),
 411 Devetag (2005) (DT) and VBB. Table 8 reports summary statistics of SL-
 412 PGG behaviour compared to subjects in these studies. All these studies used
 413 VBB's payoff matrix and had groups between 5-7 subjects except VBB, which
 414 had groups of 14-16. The figure shows similar declines in choices in all these
 415 studies. On the other hand, there appears to be greater variability in the over-
 416 all level of average choices, with SL-averages appearing higher than those in
 417 other studies.

418 We used the data reported by DT for the comparison with SL-observations.
 419 In terms of means, medians and distributions for $r=10$ rounds, round four
 420 and ten behaviours were different in terms of all three at the 10%-level of
 421 significance (see table 5). With one exception ($U_{n=3} = 931.0, p = 0.070$), the
 422 other twenty-four tests were negative, suggesting no differences exist in the

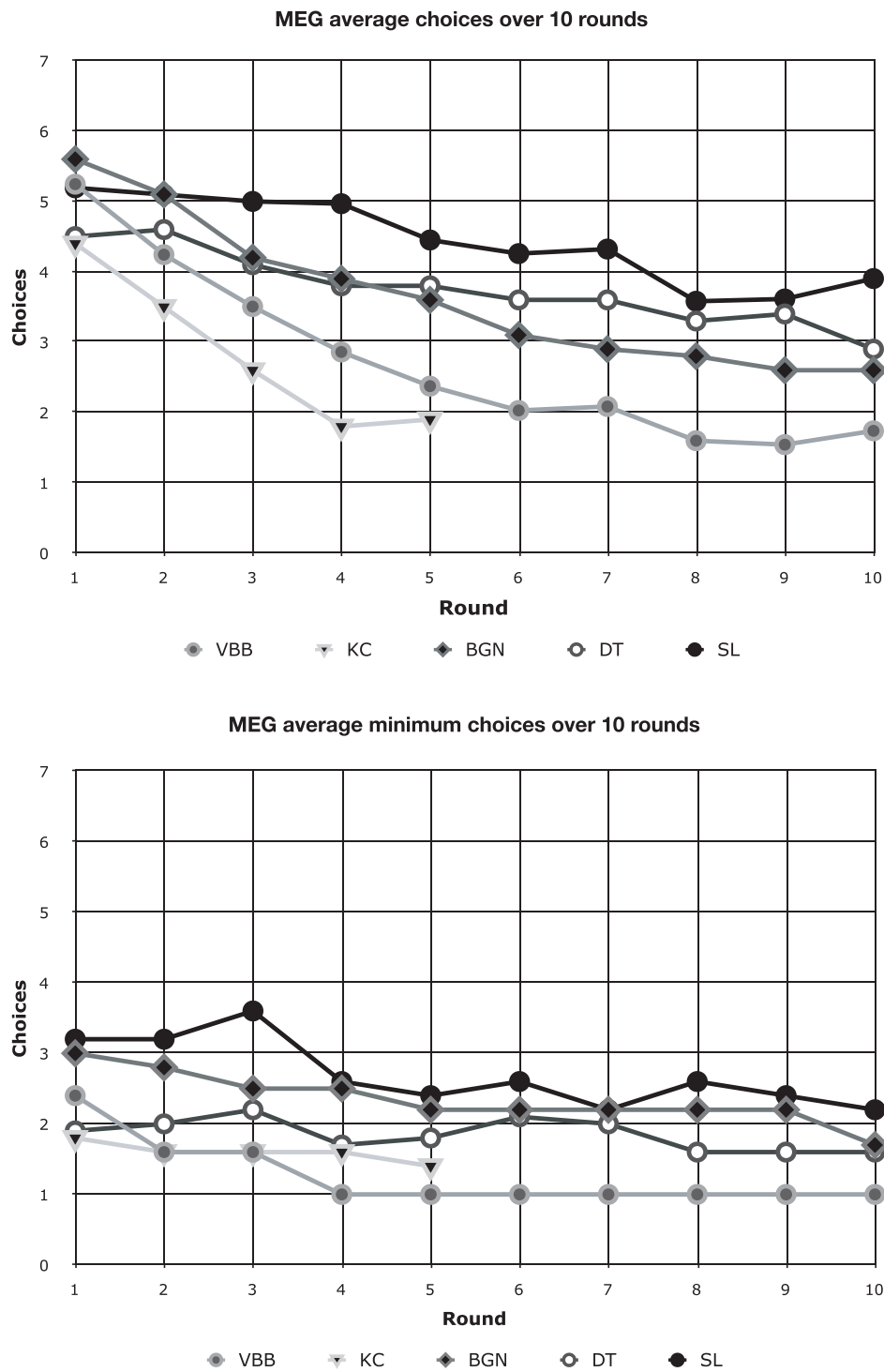


Fig. 6. Average and average minimum MEG choices over $r=10$ rounds in SL and selected previous studies.

	Smallest choice in group						
	7	6	5	4	3	2	1
7	390	330	270	210	150	90	30
6	-	360	300	240	180	120	60
5	-	-	330	270	210	150	90
4	-	-	-	300	240	180	120
3	-	-	-	-	270	210	150
2	-	-	-	-	-	240	180
1	-	-	-	-	-	-	210

Table 7

MEG payoff matrix (in L\$). The first column represents player choices which, combined with the smallest choice in the group, determines payoffs. Dashes denote logically impossible outcomes.

	VBB	KC	BGN	DT	SL
N	107	30	42	77	31
n	14-16	6	7	7	5-6
r	10	5	10	14	10
Stake	1.30	1.30	1.30	1.82	1.46
Choices					
Mean	2.72	2.87	3.65	3.75	4.44
Median	2.50	2.80	2.40	3.60	4.60
St. Dev.	1.30	1.07	1.34	1.57	1.51

Table 8

Summary statistics of minimum effort game choices over r rounds for N subjects playing in groups of n in SL as well as reported in selected previous studies. Stakes are given as U.S.\$-value of payoff associated with unique Pareto-efficient outcome.

423 rounds concerned. Again, we regressed equation 1 for the combined data set
 424 (table 6). The results show that at the 95% significance level, our data are
 425 different to those of DT as the coefficient for X is significant ($p = 0.034$). It
 426 should be noted that the same model also yields differences between the data
 427 of DT and BGN ($p = 0.084$) as well as between SL and BGN ($p = 0.002$). As a
 428 result, for the MEG, these findings do not provide firm conclusions about the
 429 ability of virtual world experimentation to replicate laboratory results. The
 430 two comparator experiments differ from ours in an additional, demographical
 431 dimension and also differ from each other in terms of results. The reason may

432 lie in greater general variability in MEG-behaviour due to the presence of
433 multiple equilibria.

434 3.6 *Guessing Game*

435 The GG sessions were conducted on 8 and 15 November 2007. In them, $n=3$
436 to 7 subjects were asked to choose integers in the interval $[0,100]$ and informed
437 that the subject with a response closest to $g = 0.7$ times the average of all
438 choices would receive L\$200 (U.S. \$0.75). Ties were resolved by dividing this
439 sum among the winners. Each group of subjects played $r = 10$ rounds of this
440 game.

441 The GG (sometimes known as the beauty contest game) is used as a tool to
442 identify what levels of reasoning subjects employ in strategic thinking (Nagel,
443 1995; Duffy and Nagel, 1997; Camerer, 1997). A zero-order (i.e. unstrategic)
444 player may choose randomly or use a focal point such as the median of the
445 interval (50 in our case). First-order choosers may take others into consider-
446 ation but assume these to be of order 0. An optimal first-order choice would
447 be in the interval $[0,70]$ accounting for the impossibility of the group average
448 to exceed 70. In particular, a choice of 35 (0.7×50) may reflect a belief that
449 zero-order guessers choose 50 on average. Second-order players who assume
450 others to use order 1 will not choose above 49 (0.7×70), and may opt for 25
451 (0.7×35) believing order 1 choices to average 35 and so forth. The iterative
452 application of increasingly higher levels of reasoning will eventually yield an
453 equilibrium choice of 0.

454 The average and distribution of GG-choices therefore provides insights not
455 only to what levels of reasoning subjects use, but also what levels they at-
456 tribute to others. Equilibrium choices may reflect higher orders of reasoning
457 but be ineffective when other players operate at lower levels. In addition, re-
458 peated GGs show to what extent subjects learn to adjust their choices on the
459 basis of previous rounds' results. Table 9 shows statistics concerning subjects'
460 choices in single or first rounds of repeated games played in groups of differ-
461 ent sizes with a parameter $g = 0.7$. The Singaporean student data are from
462 10-round GG-experiments reported in Ho et al. (1998) (HCW). The HCW 1
463 pool consisted of 3-player groups playing the game for the first time. Subjects
464 in HCW 2 also played in 3-player groups but had experience of one previous
465 game with a different g -value. Finally, HCW 3 was composed of inexperienced
466 7-subject group players. In all HCW-treatments, the winning subject received
467 50 Singapore cents (ca. U.S.\$ 0.34). The U.S. study of Kovalchik et al. (2005)
468 (KCGPA) compares one-round choices by college students (KCGPA 1) with
469 those of mentally healthy senior citizens with an average age of 82 (KCGPA
470 2). Our experimental settings of group size, g -value and repetition are the

Subjects	Mean	Median	St. Dev.	% 0	<i>N</i>
Caltech students	21.88	23.00	10.35	0.07	27
Portfolio managers	24.31	24.35	16.15	0.08	26
Economics PhDs	27.44	30.00	18.69	0.13	16
U.S. high school students	32.45	28.00	18.61	0.04	52
College students (KCGPA 1)	35.00	35.00	12.86	0.00	51
Singaporean students (HCW 1)	36.45	35.00	24.28	0.00	21
German students	36.73	33.00	20.21	0.03	67
Senior citizens (KCGPA 2)	37.00	33.00	17.46	0.00	50
University CEOs	37.81	36.50	18.92	0.03	73
Wharton students	37.92	35.00	18.84	0.00	35
Singaporean students (HCW 3)	39.78	35.00	25.46	0.02	49
SL	50.00	56.00	27.10	0.00	31
Singaporean students (HCW 2)	58.27	50.00	26.98	0.05	21

Table 9

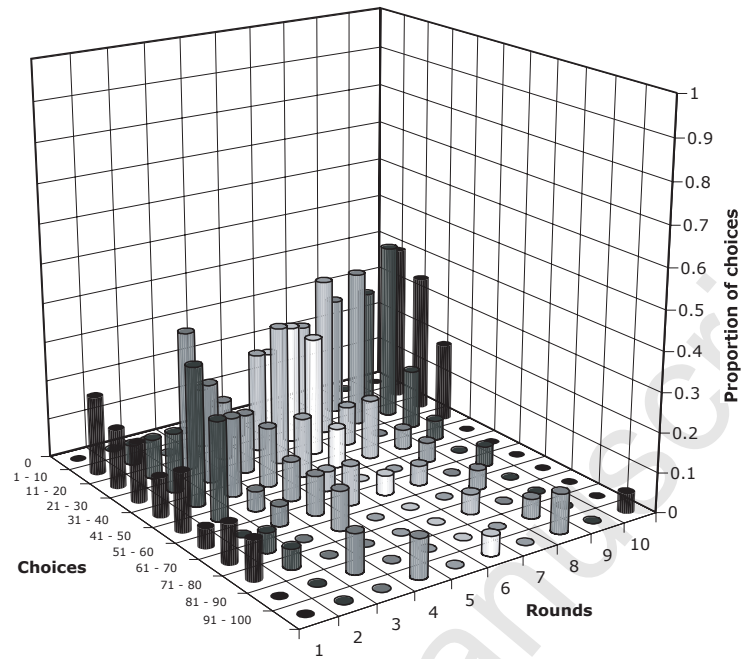
Summary statistics for round 1 GG choices in n -subject pools in SL as well as reported by Camerer (2003), Camerer (1997) and Kovalchik et al. (2005). The percentage of subjects choosing 0 is given by %0.

471 same as in HCW 1, which is most useful for a direct comparison.

472 SL first round choices are relatively high (especially compared to our bench-
473 mark HCW 1) but by no means outside the range of previous results. The bot-
474 tom panel in figure 5 shows mean choices over ten rounds among SL-subjects
475 and Singaporean students (HCW). Table 9 reports summary statistics of SL-
476 GG behaviour compared to subjects in this study. Our subjects did appear to
477 converge towards the equilibrium at similar rates to the latter. The frequency
478 distribution of individual SL-choices over all ten rounds is displayed in figure 7,
479 along with the corresponding data for HCW 1 reported in Ho et al. (1998)(p.
480 955, figure 2E). Both distributions are similar in that a greater proportion
481 of choices are low in later rounds. The SL-data appear different mainly in
482 the more equal distribution in early rounds. However, towards the end of the
483 game, the distributions are more similar, reflected in the convergence of curves
484 in figure 5.

485 GG data generally show divergence in first-round average choices. Part of the
486 reason may be the role that players' common knowledge of rationality has in
487 equilibrium reasoning. Lower choices are not merely associated with greater
488 strategic sophistication among players, but also with greater expectations con-

Frequency distribution of GG choices (HCW 1)



Frequency distribution of GG choices (SL)

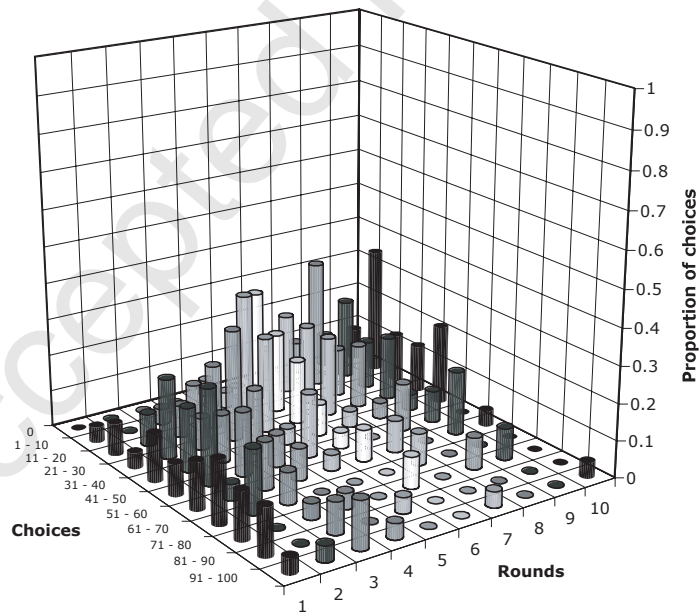


Fig. 7. Subject choice frequency distributions over $r=10$ rounds (group size 3, $p=0.7$) in HCW 1 and SL.

cerning the sophistication of others. Groups that are more sophisticated as well as more uniformly so, such as Caltech students, may therefore be expected to exhibit lower choices than comparatively heterogeneous groups such as SL where little is known about others who take part. Our first-round results may have not been much different had our pool consisted of anonymous and mutually unaware game theorists disguised by avatars. The fact that SL-subjects' learning resulted in similar final-round choices supports this possibility. The anonymity of SL, potentially subverting the common knowledge of rationality, may therefore partly explain any differences in round one choices in SL.

We compared the means, medians and distributions of SL choices with HCW 1 over $r=10$ rounds (see table 5). Rounds 1, 3 and 9 show differences in all three dimensions. In total, seven of the thirty tests were positive, most only at the 10%-significance level. Table 6 shows the regression results for equation 1 pooling SL data with HCW 1 and 3. The latter study was not used for the tests as its larger subject group size rendered it inappropriate for a direct comparison; however, we were able to control for that difference using variable n in the regression. The results show an insignificant coefficient for X ($p = 0.416$). We conclude differences are not in evidence between the data sets.

3.7 *Universal Human Values*

In order to assess whether an idiosyncratic cultural environment exists within SL, we administered the ESS human values survey. This survey is based on Schwartz's portrait values questionnaire, a well-tested instrument for identifying ten universal value dimensions (listed in figure 8). An individual's scores are calculated on the basis of responses on a 6-point Likert scale indicating own similarity with 21 hypothetical value portraits. Subjects completed the survey on a webform immediately after the decision task stage of the session. Upon completion, each subject was paid L\$1000 (ca. U.S. \$3.85) for the survey in addition to the pay-outs from the decision task.

Again, a host of existing data for this survey generates scope for comparing SL-subjects with standard populations. Cultural and demographic factors may have an influence on economic behaviour as they shape an individual's social interaction and socialisation into particular values. Values are therefore an important indicator of how representative particular subject pools are of the underlying population to which economic theory relates. We conducted the human values survey in order to ascertain to what extent SL-residents resemble standard experimental subjects culturally. Figure 8 shows the average value orientations of our subjects compared with those of respondents of the 2002-2003 ESS, as well as a standard sample of thirty-six UK university stu-

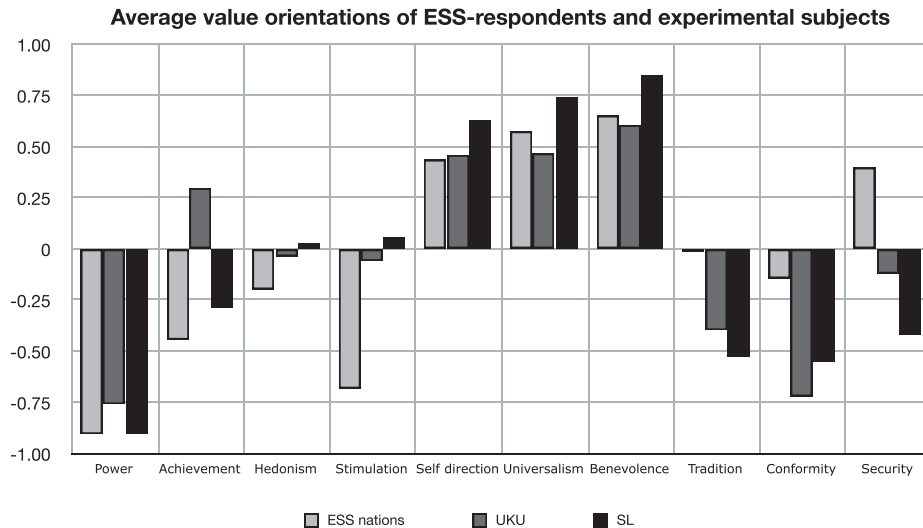


Fig. 8. Average orientations of ESS-respondents (ESS), SL and UK student subjects (UKU) according to Schwartz' ten value dimensions.

528 dents (UKU) we also administered the questionnaire to. The ESS randomly
 529 samples more than 1500 adults from each participating nation's resident popu-
 530 lation. The students were UK nationals invited randomly by automated email
 531 from the experimental subject database maintained by the *Centre for Decision*
 532 *Research and Experimental Economics*. For comparative purposes, we follow
 533 the ESS practice of presenting averages of ipsative scores, i.e. an individual's
 534 Likert-scale responses standardised in terms of his or her overall response av-
 535 erage and variance. Ipsatised scores for different value dimensions have the
 536 advantage of being comparable in terms of relative strength.

537 Schwartz' ten human values are shown along the horizontal axis of figure 8.
 538 They have established empirical interrelationships that are commonly used
 539 to reduce them to two basic dimensions shown along the two respective axes
 540 in figure 9. The first dimension, *self-transcendence v. self-enhancement*, en-
 541 compasses six values: hedonism, stimulation and self direction relative to tra-
 542 dition, conformity and security. The former three values express underlying
 543 motivations such as pleasure, sensuous gratification, excitement, novelty and
 544 independence, while the latter express respect and acceptance of norms, self-
 545 restraint and harmony. The remaining four universal values are contained in
 546 the the second dimension, *openness to change v. conservatism*. It weighs the
 547 values of universalism and benevolence against those of power and achieve-
 548 ment. The former two values express motivations including tolerance and care
 549 for the welfare of others, while the latter two encompass social status, personal
 550 success and dominance over others. Figure 9 plots nations and subject pools
 551 according to the two overall dimensions.

552 Our survey data indicate that while SL-users' value orientations differ from
 553 those of ESS-respondents, they do so to a lesser extent than those of the UK

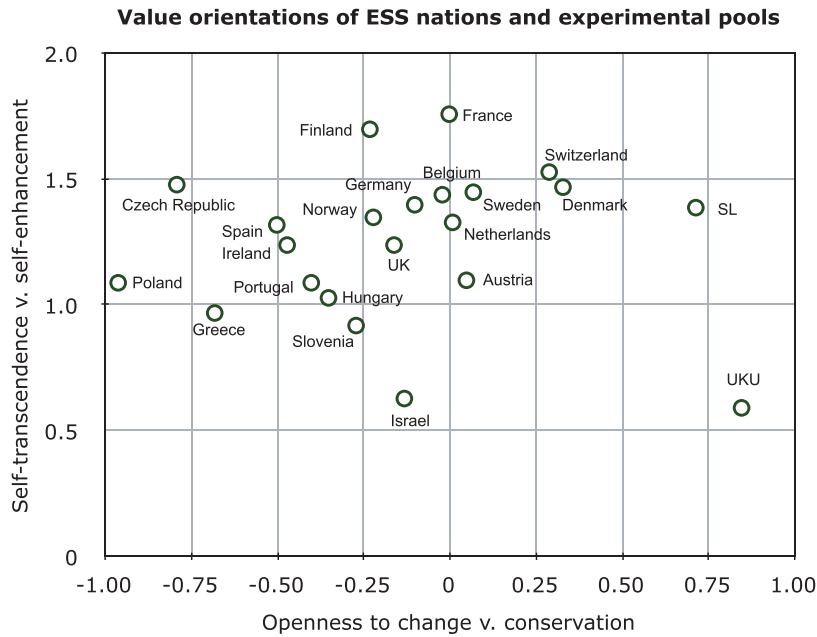


Fig. 9. Average orientations of ESS-respondents by nationality, SL and UK student subjects (UKU) according to Schwartz' two composite value dimensions.

554 student subjects. The SL and student average value orientations correlate at
 555 90% with each other, and respectively at about 70 and 64% with the averaged
 556 overall ESS-orientation of EU respondents. By comparison, individual national
 557 samples within ESS correlate with the average EU-values profile at about
 558 94%. The graph shows a relatively small distance between randomly-sampled
 559 individuals from European nations to SL-users and UK students. The students
 560 place a greater importance on the factors underlying self-enhancement, as can
 561 be verified in figure 8. This is consistent with age effects found in previous value
 562 surveys comparing students and teachers (Schwartz, 2001). Another reason for
 563 the difference may lie in a slightly higher relative socio-economic background
 564 and educational potential of students. However, caution has to be exercised
 565 due to our small sample size.

566 4 Methodological Discussion

567 Our experience of conducting experiments in SL suggests a number of ad-
 568 vantages and disadvantages of virtual experimentation generally as well as
 569 practical steps to adapt the platform for experimental purposes.

570 It was possible, with little organisation and preparation, to recruit subjects
 571 *in situ* in the numbers we could manage within the SL-interface. SL's fea-
 572 tures make it simple to create and maintain a database of subjects for future
 573 use. On the other hand, this procedure is prone to biased sample selection

574 on the basis of choosing busy recruitment locations, of solicitation, in the
575 recruitment language, time and institutional affiliation we used. In addition,
576 the relative anonymity that avatars confer on subjects makes it difficult in
577 practice to prevent financially-motivated repeat participation or the recruit-
578 ment of unfit (tired or intoxicated) or non-eligible or non-targeted subjects.
579 While these issues may not be completely resolvable, we attempted to miti-
580 gate both repeat and unsuitable participation by disqualifying avatars using
581 the following criteria. First, to avoid repeats, we excluded avatars who partici-
582 pated previously, who were created after the first experimental session or who
583 made unsolicited approaches to us. To avoid unsuitable participants, we also
584 excluded avatars less than a month old and potentially insufficiently familiar
585 with the SL-environment, avatars referred by previous subjects who may have
586 prior knowledge of the task, and avatars representing users who appeared to
587 be in an unfit state. An additional identity issue both in our and in other
588 virtual world studies concerns the potential for a disparity between user and
589 avatar characteristics. For many users, the attraction of SL consists of the
590 potential for using an avatar to assume a new and different identity. While
591 our study was designed to elicit the behaviour and values of users and not
592 avatars, we cannot be certain to what extent this was practised by subjects
593 responding through their avatars.

594 Our demographical and values survey shows that virtual worlds provide oppor-
595 tunities for recruiting subjects who are demographically more representative
596 than university students. In addition, targeting particular types of individuals
597 is possible within those groups represented in virtual worlds, such as partic-
598 ular nations. Clearly, some groups are currently not sufficiently represented
599 in virtual worlds, including individuals from smaller and traditional societies.
600 However, the bias of SL towards industrialised nations is likely to change as
601 economic development provides greater access to the Internet to more people
602 worldwide.

603 The relatively sophisticated SL-economy provides some scope for appropriate
604 incentive mechanisms. In particular, SL has developed informal labour and
605 product markets which generate incentivisable subjects as well as money or
606 in-kind rewards that can be delivered easily. Many users regularly participate
607 in paid online activities for returns which are modest compared with those of
608 standard economic experiments. In addition, the developed markets for virtual
609 objects provide alternative in-kind incentives.

610 While the computerised interface of SL provides an economical experimental
611 environment that is well suited for data generation, collection and storage, it
612 also has certain disadvantages. Communicating with subjects using IM makes
613 it difficult to deal with more than a handful per session. In addition, private IM
614 makes it hard to detect collusive behaviour or conferring amongst subjects.
615 While is it not possible to override the communication mechanisms of SL,

616 we developed virtual laboratory furniture that alerts the experimenter to the
617 potential for clandestine communication between subjects (visible in figure
618 2). In particular, upon entering the virtual laboratory, subjects were asked
619 to sit in cubicles and to enter *mouselook*, a SL-mode under which avatars
620 are restricted to frontal vision and where private IM is suspended, in line
621 with standard experimental conditions. Once activated, the furniture indicates
622 whenever a subject suspends the mouselook mode and is therefore able to use
623 private IM. While this furniture assured discipline in practice, it is in theory
624 possible for experts to circumvent such mechanisms. On the other hand, this
625 requires not only significant expertise on the part of a subject, but matching
626 skills of and prior collusion with another subject present in order to establish a
627 clandestine communication channel. Another problematic issue is establishing
628 subject trust in the experimenters. Because of the nature of virtual worlds,
629 it is difficult to convince subjects of the genuine nature of the experiment
630 and incentivisation. A further problem involves the potential for disruption
631 of experimental sessions by other users. This, however, may be controlled by
632 restricting access to the virtual laboratory.

633 The absence of physical signals and presence in virtual worlds creates clear
634 differences between virtual and physical experimental conditions. Virtual ex-
635 periments preclude physical presence that may influence behaviour through
636 involuntary non-verbal communication that reveals emotional states. In ad-
637 dition, the potential for anonymity means that the social consequences of
638 virtual behaviour are different to those in physical laboratories. These factors
639 may limit the comparability of virtual and physical experimental results in
640 many cases. Virtual experimentation is clearly not appropriate when physical
641 phenomena are part of the experimental treatment, such as when the effect of
642 face-to-face interactions is tested.

643 5 Conclusion

644 Despite the non-standard nature of the SL-subject pool and certain imper-
645 fections of the experimental environment that it provides, we were unable to
646 detect significant and systematic overall differences between their behaviours
647 and those observed in traditional settings. In particular, given SL-users' demo-
648 graphics in terms of age and cultural background, behaviour closely matched
649 expectations based on a host of existing experimental evidence for a range
650 of five important games. These results suggest tentatively that virtual world
651 economic phenomena are based on similar behavioural regularities observed
652 in standard economic settings and can be tested experimentally within the
653 virtual environment.

654 In addition, there is a slightly lesser cultural and age bias within SL than

655 at the average university campus. Users' values are more in line with those
656 of general populations of economic agents. There was little evidence of users'
657 niche interests or motivations generating an unsuitable subject pool. Our work
658 therefore supports Yee (2006), whose study of virtual world demographics
659 dispels the popular notion that they are predominantly the domain of a male,
660 adolescent sub-culture with niche interests. His data indicate that usage and
661 appeal are equally strong over gender and age groups as well as based more
662 on general social motivations (such as relationship building) than escapism.

663 It should be noted that our study was not designed to provide support for
664 or against virtual world experimentation as a method in absolute terms. In-
665 stead, we adopted a less ambitious research question regarding its ability to
666 reproduce the results of traditional experimentation in physical laboratories
667 with standard subjects. As a result, the absence of observed behavioural dif-
668 ferences between the two environments does not necessarily make a case for
669 virtual experiments *per se*, but rather suggests they may be a valid alterna-
670 tive to traditional method, subject to similar methodological advantages and
671 limitations. Conversely, the presence of such differences would not necessarily
672 invalidate virtual experimentation to the extent that the standard physical
673 laboratory method is not without imperfections. As a result, these method-
674 ological issues remain and may benefit from renewed debate in the context of
675 virtual experimentation.

676 While the above suggests that virtual world experimentation has potential as
677 an economical and practical alternative to standard laboratory experiments,
678 there are certain disadvantages associated with virtual worlds as experimental
679 platforms which suggest that their suitability depends on the type of experi-
680 ment planned. For instance, studies that consider the effects of physical sig-
681 nals or depend on recruiting specific types of subjects will find little value in
682 virtual experimentation. On the other hand, suitably adapting experimental
683 procedures to the virtual world environment makes it possible to effectively
684 and cheaply recreate many standard decision tasks. In addition, virtual world
685 users appear to constitute suitable subject pools to the extent that they dis-
686 play many of the economic behaviours associated with standard subjects. The
687 future development of this technology will further increase the sophistication
688 of the virtual experimental platform.

689 References

- 690 Anderhub, V., Müller, R., Schmidt, C., 2001. Design and evaluation of an
691 economic experiment via the internet. *Journal of Economic Behavior and*
692 *Organization* 46 (2), 227–247.
- 693 Andreoni, J., 1988. Why free ride? Strategies and learning in public goods
694 experiments. *Journal of Public Economics* 37, 291–304.

- 695 Andreoni, J., 1995. Cooperation in public-goods experiments: Kindness or
696 confusion? *American Economic Review* 85 (4), 891–904.
- 697 Bainbridge, W. S., 2007. The scientific research potential of virtual worlds.
698 *Science* 317, 472–476.
- 699 Bloomfield, R. J., 2007. Worlds for study: Invitation - virtual worlds for study-
700 ing real-world business (and law, and politics, and sociology, and....). SSRN
701 eLibrary.
- 702 Bornstein, G., Gneezy, U., Nagel, R., 2002. The effect of intergroup competi-
703 tion on group coordination: An experimental study. *Games and Economic*
704 *Behavior* 41, 1–25.
- 705 Buchan, N. R., E., J., Croson, R. T. A., 1997. Culture, power and legitimacy:
706 Contrasting influences on fairness beliefs and negotiation behavior in Japan
707 and the United States, University of Wisconsin Department of Marketing
708 Working Paper.
- 709 Cagnina, M. R., Poian, M., 2007. How to compete in the metaverse: The
710 business models in second life. Tech. rep.
711 URL <http://ssrn.com/abstract=1088779>
- 712 Camerer, C. F., 1987. Do biases in probability judgment matter in markets?
713 *American Economic Review* 77 (5), 981–997.
- 714 Camerer, C. F., 1997. Progress in behavioral game theory. *Journal of Economic*
715 *Perspectives* 11 (4), 167–188.
- 716 Camerer, C. F., 2003. *Behavioral Game Theory*. Princeton University Press,
717 Princeton, NJ.
- 718 Carpenter, J. P., Burks, S., Verhoogen, E., 2005. Comparing students to work-
719 ers: The effect of social framing on behavior in distribution games. In: Car-
720 penter, J. P. (Ed.), *Field Experiments in Economics*. Vol. 10. Elsevier, Ch. 7,
721 pp. 261–290.
- 722 Cason, T. N., Mui, V. L., 1998. Social influence and the strategy method in the
723 sequential dictator game. *Journal of Mathematical Psychology* 42, 248–265.
- 724 Castranova, E., 2005. *Synthetic Worlds: The Business and Culture of Online*
725 *Games*. University of Chicago Press, Chicago, IL.
- 726 Castranova, E., 2006. On the research value of large games: Natural experiments
727 in *Norrath and Camelot*. *Games and Culture* 1, 163–186.
- 728 Chuah, S. H., Hoffmann, R., Jones, M. K., Williams, G. A., 2006. An economic
729 anatomy of culture: Attitudes and behaviour in inter- and intra-national
730 ultimatum game experiments, *Occasional Papers* 13, Industrial Economics
731 Division, Nottingham University Business School.
- 732 Chuah, S. H., Hoffmann, R., Jones, M. K., Williams, G. A., 2007. Do cultures
733 clash? Evidence from cross-national ultimatum game experiments. *Journal*
734 *of Economic Behavior and Organization* 64, 35–48.
- 735 Crawford, V. P., 1995. Adaptive dynamics in coordination games. *Economet-*
736 *rica* 63 (1), 103–143.
- 737 Devetag, G., 2005. Precedent transfer in coordination games: An experiment.
738 *Economics Letters* 89, 227–232.
- 739 Duffy, J., Nagel, R., 1997. On the robustness of behaviour in experimental

- 740 ‘beauty contest’ games. *Economic Journal* 107, 1684–1700.
- 741 Erev, I., Roth, A. E., 1998. Predicting how people play games: Reinforce-
742 ment learning in experimental games with unique, mixed strategy equilibria.
743 *American Economic Review* 88 (4), 848–881.
- 744 Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods ex-
745 periments. *American Economic Review* 90 (4), 980–994.
- 746 Forsythe, R., Horowitz, J. L., Savin, N. E., Sefton, M., 1994. Fairness in simple
747 bargaining experiments. *Games and Economic Behavior* 6, 347–369.
- 748 Guest, T., 2007. *Second Lives*. Random House, London, UK.
- 749 Harbaugh, W. T., Krause, K., Liday, S. G., 2003. Bargaining by children,
750 working paper.
- 751 Harrison, G. W., List, J. A., 2004. Field experiments. *Journal of Economic*
752 *Literature* 42 (4), 1009–1055.
- 753 Healy, P. J., 2006. Learning dynamics for mechanism design: An experimen-
754 tal comparison of public goods mechanisms. *Journal of Economic Theory*
755 129 (1), 114–149.
- 756 Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. (Eds.),
757 2004. *Foundations of Human Sociality: Economic Experiments and Ethno-*
758 *graphic Evidence from fifteen Small-Scale Societies*. Oxford University
759 Press, Oxford, UK.
- 760 Ho, T., Camerer, C., Weigelt, K., 1998. Iterated dominance and iterated best
761 response in experimental ‘p-beauty contests’. *American Economic Review*
762 88 (4), 947–969.
- 763 Hoffman, E., McCabe, K., Shachat, K., Smith, V. L., 1994. Preferences, prop-
764 erty rights and anonymity in bargaining games. *Games and Economic Be-*
765 *havior* 7, 346–380.
- 766 Hoffman, E., McCabe, K., Smith, V. L., 1996. Social distance and other-
767 regarding behavior in dictator games. *American Economic Review* 86 (3),
768 653–660.
- 769 Knez, M., Camerer, C., 1994. Creating expectational assets in the laboratory:
770 Coordination in ‘weakest-link’ games. *Strategic Management Journal* 15,
771 101–119.
- 772 Kovalchik, S., Camerer, C. F., Grether, D. M., Plott, C. R., Allman, J. M.,
773 2005. Age and decision making: A comparison between neurologically
774 healthy elderly and young individuals. *Journal of Economic Behavior and*
775 *Organization* 58, 79–94.
- 776 Kurz, M., 2008. Beauty contests under private information and diverse beliefs:
777 How different? *Journal of Mathematical Economics* 44 (7-8), 762–784.
- 778 Linden-Labs, 2008. Second life website.
779 URL <http://secondlife.com/>
- 780 Nagel, R., 1995. Unraveling in guessing games: an experimental study. *Amer-*
781 *ican Economic Review* 85 (5), 1313–1326.
- 782 Noveck, B. S., 2004. Introduction: The state of play. *New York Law School*
783 *Law Review* 49 (1), 1–18.
- 784 Oosterbeek, H., Sloof, R., van de Kuilen, G., 2004. Cultural differences in

- 785 ultimatum game experiments: evidence from a meta-analysis. *Experimental*
786 *Economics* 7, 171–188.
- 787 Rokeach, M., 1973. *The Nature of Human Values*. Free Press, New York, NY.
- 788 Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., Zamir, S., 1991. Bargaining
789 and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An
790 experimental study. *American Economic Review* 81 (5), 1068–1095.
- 791 Schotter, A., Weiss, A., Zapater, I., 1996. Fairness and survival in ultimatum
792 and dictatorship games. *Journal of Economic Behavior and Organization*
793 31, 37–56.
- 794 Schwartz, S. H., 2001. Value hierarchies across cultures: Taking a similarities
795 perspective. *Journal of Cross-Cultural Psychology* 32 (3), 268–290.
- 796 Schwartz, S. H., 2002. A proposal for measuring value orientations across
797 nations. In: *European Social Survey Questionnaire Development Report*.
798 European Social Survey, London, UK.
- 799 Van Huyck, J., Battalio, R., Beil, R., 1990. Tacit coordination games, strategic
800 uncertainty, and coordination failure. *American Economic Review* 80, 234–
801 248.
- 802 Yee, N., 2006. The demographics, motivations and derived experiences of users
803 of massively-multiuser online graphical environments. *Presence: Teleopera-*
804 *tors and Virtual Environments* 15, 309–329.