Journal of Economic Psychology 33 (2012) 482–493

Contents lists available at SciVerse ScienceDirect

Journal of Economic Psychology

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

Truck, barter and exchange versus the endowment effect: Virtual field experiments in an online game environment

Yannick Ferreira De Sousa [a], Alistair Munro [b,*]

[a] Royal Holloway, University of London, UK
[b] National Graduate Institute for Policy Studies, Roppongi 7-22-1, Tokyo 106-8677, Japan

ARTICLE INFO

Article history:
Received 23 December 2009
Received in revised form 29 December 2011
Accepted 31 December 2011
Available online 5 January 2012

JEL classification:
C91
C93
D03

Keywords:
Endowment effect
Virtual field experiment
Runescape
MMORPG

ABSTRACT

We examine the feasibility of using a massively multiplayer online role-playing game (MMORPG) to test economic theories. As a test vehicle we use the well-known endowment effect. Even though our goods are entirely virtual, our results confirm earlier results that individuals with more trading experience are less likely to exhibit status quo behaviour in trade. However, we also find evidence that highly experienced individuals are more likely to swap the item rather than keep it – i.e. there appears to be a propensity to 'truck, barter and exchange'. A further experiment suggests that this feature is robust and is unlikely to be due to subject misperception or experimenter demand effects. However we are unable to eliminate selection effects as the source of our correlation between experience and propensity to trade. We conclude that virtual economies may be useful venues for field experiments.

1. Introduction

Online role-playing games have become a significant venue for social interaction drawing millions of players daily to virtual worlds full of conflict, puzzles and trade. Multi-player games such as Runescape, Second Life, Farmtown or World of Warcraft have claims to be real economies (Castronova, 2002). In them, individuals labour to produce goods which are then exchanged with the products of other peoples' time within well-developed trading institutions. As with real-world societies, experience acquired through the investment of time and energy changes productivity, while the freedom given to players to define their own strategies means that institutions and associations emerge and disappear spontaneously.

As well as being small-scale economies, virtual game environments have the potential to be venues for economics experiments, a point well-demonstrated by Chesney, Chuah, and Hoffmann's (2009) use of Second Life.[1] In this paper we report on an experiment conducted within Runescape, a Java-based MMORPG (massively multiplayer online role-playing game). Our aim is threefold: first, to deepen understanding of behaviour in online economies. Second, in common with Chesney et al. (2009) or Castronova (2008), our aim is methodological: to explore the feasibility of using multiplayer gaming environments for (virtual) field tests of economic theories. An obvious point of contrast between our work and Chesney et al. (2009), or Fiedler and Haruvy

* Corresponding author.
E-mail address: alistair-munro@grips.ac.jp (A. Munro).

[1] The same environment is used by Fiedler and Haruvy (2009) for their trust game experiments

0167-4870/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.joep.2011.12.011

(2009), is that their set-up is designed to mimic a laboratory environment. In that regard our intervention more closely resembles a field experiment, albeit one where subjects still know that they are taking part in an experiment. Our third and final goal is to extend the domain of experimental research on the 'endowment effect'. This phenomenon is said to occur when the minimum compensation which individuals are willing to accept (WTA) in return for giving up a good is greater than the amount which they would be willing to pay (WTP) to acquire it in the first place (Kahneman, Knetsch, & Thaler, 1991). It has been recorded by researchers many times in laboratory experiments (Bateman, Munro, Rhodes, Starmer, & Sugden, 1997; Kahneman et al., 1991; Shogren, Seung, Dermot, and James, 1994) and investigated in a number of field settings (e.g. List, 2004). In one typical format for the experiment, subjects are randomly endowed with one of two goods, say A and B. At a later stage in the experiment, they are invited to exchange the endowment for the other product. According to standard consumer theory, preferences are independent of endowment. As a result the proportion of the subjects who prefer A to B should be equal in the two sub-samples. If choice reveals preference, then this means that the proportion that chooses A over B should be equal across the sub-samples or equivalently that the sum of the proportions of subjects who accept the invitation to swap should equal 1. In practice in many experiments (e.g. Kahneman et al., 1991; Knetsch, 1989), the sum falls well short of one, indicating that a significant proportion of subjects show a reluctance to give up the endowment.[2] In his widely-cited field experiments, List (2003, 2004) extends the results by leaving the laboratory and using subjects from US memorabilia markets. In these markets collectors and traders exchange collectibles such as sports cards and limited edition pins (badges).[3] As in the laboratory, each subject is randomly endowed with one good of two goods and later offered a chance to exchange it for the other item. List finds endowment effects for amateurs who trade less intensively, but no effects for professional traders and high intensity amateur traders. If they are widely true, then these results have strong implications for the general applicability of laboratory experiments on the endowment effect. Thus there is value in seeing whether List's influential results are replicable in other arenas or whether they are confined to the particular US memorabilia markets he studies. One particular feature of role playing games (RPGs) is that, like the sports card and pin markets, virtual markets have variation between experienced and inexperienced individuals. Furthermore, the level of experience in the virtual market is typically reflected in player scores that make it easier to categorise experienced and inexperienced market players. Runescape in particular has many opportunities for trading, so it seems like a natural venue for a field test.

In terms of methodology, we find it relatively easy to conduct an experiment in an online environment, though the level of control over subject behaviour is closer to that typically observed in a field experiment or a street-based survey rather than a laboratory experiment. For instance, though nearly every subject invited to take part agreed and completed the experiment successfully, three potential subjects 'ran away' with the endowed good, a problem not typically faced in the laboratory.

The major results are as follows: first, we find evidence of a large endowment effect for inexperienced players. In other words, online players act like subjects in many other experiments on the subject. Second, in keeping with other field studies we find that the higher the level of experience the lower the endowment effect. Our third main result is that highly experienced players over-trade, in the sense the average person is more likely to swap their endowment for the alternative than to stick to the status quo. This is where our quote from Adam Smith originates: for experienced players a propensity to truck barter and exchange[4] appears to dominate any endowment effect.

The remainder of this paper is as follows: Section 2 provides background material on our virtual laboratory and describes the experimental design and results of our first experiment; Section 3 presents the design and analyses the empirical evidence from a second, follow-up experiment and Section 4 concludes with some discussion of methodological issues and some speculation on the implications of our results for the interpretation of field experiments on the endowment effect.

2. Experiment one

2.1. Design

We begin with some brief but necessary background about the possibly unfamiliar environment of the experiment, Runescape which has around one million paid-to-play accounts (P2P) and nine million free-to-play accounts (F2P).[5] The game takes place in the fantasy realm of Runescape which is divided into 18 different kingdoms or regions. Players are shown on the screen as playable avatars (i.e. online characters). Compared to some other MMORPGs, the game gives extensive freedom to players in term of setting their own objectives, and deciding which of the available skills and activities to pursue. There is no linear path that must be followed in Runescape: players can engage in fights with or against others or with monsters. They

[2] The original experimenters, surveyed in Kahneman et al. (1991), devote some efforts towards removing misunderstanding, perceptions of scarcity, transactions costs or other frictions as explanations of the inertia.

[3] For instance, a sports card might be a small card showing a famous baseball player along with some key statistics from his career. A pin or badge might depict a Disney character such as Mickey Mouse. What makes the item collectible varies, but it typically involves some element of scarcity.

[4] "This division of labour, from which so many advantages are derived, is not originally the effect of any human wisdom, which foresees and intends that general opulence to which it gives occasion. It is the necessary, though very slow and gradual consequence of a certain propensity in human nature which has in view no such extensive utility; the propensity to truck, barter, and exchange one thing for another." (Smith, 1776, Book 1, chap. 2).

[5] From www.jagex.com.

are free to choose to complete *quests* – set tasks designed to teach players to trade and develop new and old skills – or to follow a certain story line or to just spend their time playing with others in cooperative mini-games.

In Runescape accumulated experience is reflected in a number of published skill scores, chief amongst which is the *combat level* of a character – a basic measure of the ability to deal with monsters and aggressive players that can be observed by other players. Some specific skills are directly relevant to combat (and therefore feed into the aggregate combat level score) and some are more important for the survival of players – such as providing food, armour and weapons to the character. In total, there are 23 skills: seven are combat skills (e.g. attack, defence, magic) that determine the combat level score through a weighted average and the remaining 16 are non-combat related (e.g. mining, smithing, crafting, fishing and cooking). Some eight skills are not available to F2P players. Players raise their non-combat skill levels through engaging in production which produces 'experience points' for that skill. In other words, a higher skill score for fishing is obtainable through spending more time fishing and so on. Higher skill levels raise productivity, make new activities feasible and may enhance social status within the game. Players can also choose particular tasks or quests that require combat and these lead to a rise in the associated combat skill levels. Combat with other players or with non-playing characters such as dragons, risks dying and while this is not terminal for a user, there is a significant cost associated with the loss of the precious items dropped or destroyed when a player is killed. To avoid death, players are encouraged to choose domains and combat challenges that are appropriate to their current skill level, but there is a trade-off between the risk of dying and the rewards from winning in combat. Players can also consult a wide variety of official and unofficial on-line guides and then buy or produce the most suitable weapons and armour in preparation for combat.[6]

Items to enhance skills, such as particular magic spells or armour can be acquired through killing specific monsters, through production or through trade. Items can also be banked for future use or exchange. To lubricate trade, the game has its own currency variously described as GPs or gold pieces. One way to earn money is to attack non-playing characters in the game or other players both of which will drop valuable items when killed. An alternative means of earning money is to engage in production and subsequent trade. For instance, a player might mine or grow crops or add value to existing goods through the skills of smithing or cooking. Essentially, to be effective in combat a player must obtain protective armour, weapons, food and various spells. These specific goods are usually unobtainable unless the player trades.

The player to player trading market is very close to a real world market and shares the same basic microeconomic fundamentals of demand and supply. For instance, the shock introduction of a macro (a script that allows actions in the game to be reproduced without the intervention of the player) once illegally introduced into the game to automate the harvesting of flax led to a sharp decrease in its price.[7] Trading is straightforward. For some goods there are specific shops, but trading can also take place either through random encounters with other players, or in response to forum messages, or in a number of specific and well-advertised locations in the virtual world where individuals who wish to trade congregate. Once contact has been established, the game has a trading system in which both players can inspect the items offered for trade. If both click in a box to signal their assent, a trade is made. We used this system to organise potential swaps.

There are no fixed rules to determine what constitutes an experienced player, but in the Runescape community it is generally believed that combat skills of 70–75 or above represent a high level player – anyone lower than that is seen as a "noob" or "new comer".[8] In what follows, we label subjects with a combat level of 75 or above as highly experienced and if below this limit then they are called low experience players. However, it is worth noting that our results are not sensitive to this precise division and indeed when we run regression results we treat skill levels as a continuous variable. Now, in some cases, higher level players can be difficult to reach as they prefer to disable their public chat – so as to avoid being disturbed by lower level players asking questions or begging for items. We wished to have a range of trading experience amongst subjects, so we recruited participants from two subject pools. The Runite Legion is a clan based community of several 100 members, many of whom have high combat levels. To obtain a sample with a decent proportion of higher level players we therefore recruited members from this community using the private chat system for clans. Individual clan members are not necessarily known to one another, but when approached (or more accurately, after receiving a message) a clan member would know that the researcher was a member of the same clan. Our second recruitment 'venue' was "Lumbridge Castle", the location within the game where new players appear after undergoing their compulsory preliminary training. The second group of players typically therefore had very low levels of experience, although in fact Lumbridge Castle is a busy, central location with players of all levels present. We waited at this place and made an initial approach when a player appeared in the on-screen vision of view of the researcher.

The design of these experiments is based on field trials of the endowment effect reported in List (2003, 2004). In January–March 2007, subjects were approached individually within the virtual world and invited to take part in a survey. Because of the busy nature of the Lumbridge Castle location there was typically a choice of whom to approach. So, to limit unconscious bias, in that venue a random number generator was used to determine whether to offer an invitation to take part. At this stage, the invitee would be able to observe the (relatively high) combat level of the researcher.[9] If a potential subject accepted

[6] See http://runescape.wikia.com/wiki/Combat_level for the formulae in use.

[7] Source: www.zybez.com. Macros are actually banned from the game, as is trading goods for real money outside the game, but there is nevertheless an external market using third party websites that provides another risky way for players to buy GPs for their characters. Players caught using goods or macros acquired in this way are banned or suspended from play.

[8] See http://runescape.wikia.com/ for example.

[9] Runescape subjects are encouraged by the design of the game to interact for trade, friendship and joint enterprise (http://runescape.wikia.com/). There is very little material reason for experienced players to kill newly created characters. Nevertheless some wariness amongst potential subjects might be expected.

Table 1
Characteristics of participants, experiment 1.

	High level players mean (std. dev.)	Low level players mean (std. dev.)
Combat level	89.66 (10.48)	34.79 (27.3)
Fraction female	0.17	0.07
Item estimations	3061.70 (435.04)	2275.86 (1852)
Years of playing	1.88 (0.99)	0.68 (1.00)
High trader	0.66 (0.48)	0.28 (0.45)
Age range	1.79 (0.55)	1.28 (0.45)
Fraction from clan	0.55	0.26
Fraction endowed with B	0.49	0.58
Fraction USA based	0.69	0.40
Fraction UK based	0.19	0.44
Fraction, non-English speaking countries.	0.04	0.11
N	47	43

Notes:
1. Item estimations *denotes how much the subject thinks the endowment is worth.*
2. High trader *denotes 1 if the subject trade 5 times or more per day, 0 otherwise.*
3. Age range equals *1 if the subject is between 13 and 16 years, 2 if 17 to 21 years old and 3 if the subject is 22 years or older.*

the invitation then he or she was endowed with one of the two possible goods using a coin toss to select the item. He or she then took part in a short question and answer session and was then finally offered the chance to swap the endowment for an alternative. In the interview phase we collected the following information: combat level, total level (the total level of all skills, including combat), and quest points.[10] For number of trades per day we used four ranges: none, 0–4, 5–9 and 10 or above. Because of privacy rules governing the game, we did not ask for any personal information other than age and gender.[11] Moreover, in order to respect the rules on privacy, we decided to use ranges of values for the age question and explicitly warned the player that they were not obliged to answer the two personal questions (in fact all did so). The next question asked was whether the subject knew about the item and how much it was worth. Finally, the subject was asked how long she or he had played the game. Once these questions were completed the alternative object was introduced for inspection and the player was invited to swap.

In this experiment we used moderately expensive bundles of goods as follows:

Good A was 150 mind runes (players from any level can use this item), a magic artefact that allows the user to cast basic attack spells including teleporting.

Good B was 150 iron arrows (players from any level can use this item). Arrows are used for ranging which allows a player to attack enemies at distance.

The goods were selected on the grounds that they were well-known and potentially useful to a wide-range of players, as well as being of approximately equal value.[12] We used questions on user forums and direct contact through the clan to select possible goods and ran pre-tests with a small number of Runescape members. Goods A and B traded for around 3,000 GP at the time of the experiment. To put this in perspective, a novice player begins the game with 32 GP and might take up to 20–30 h of work to acquire sufficient funds to afford one of these bundles of goods.[13] A higher level player might typically take only 1–3 h to achieve the same income (Bilir, 2009).

2.2. Experiment one: results

In total in the first experiment we had 90 participants, 47 with high combat levels and 43 with low levels. Table 1 shows a summary statistical description for both groups. The pattern of ages and gender revealed by this data is perhaps expected and similar to that depicted in Bilir's (2009) survey of Runescape players. Subjects are typically male and in their late teens (i.e. similar ages to the university students often used in laboratory experiments). Most subjects are from the USA or the UK, with eight coming from European countries in which the main language is not English. Higher combat level players tend to be older than the lower level players, they spend twice more time playing the game and there are relatively more female subjects within the high level sample. Estimates made by high level players are closer to the real price range value for both treatments (2250–3750 GP for each bundle) and have a lower standard deviation, showing perhaps a better understanding of

[10] As with combat level, total level is a weighted average of skill levels across all skills. Quest points are a further measure of experience based on the number and difficulty of quests completed.
[11] Although the experiment did not violate any of the rules of the game, it was not officially sanctioned by the company involved.
[12] Data on current trading prices can be obtained at a number of official and unofficial websites including, for instance, www.zybez.com. There has been extensive inflation within the game since we conducted by the experiment, largely it seems due to the illicit purchase of gold piece.
[13] With some experience, iron arrows for instance can be manufactured out of iron bars which in turn can be made from iron ore. Alternatively there is a probability that goblins will drop a few iron bars when killed. The point is that as skill levels rise, players can kill goblins more easily or smith iron more efficiently.

Table 2
Summary trading statistics for experiment 1.

Variable	Percent traded	*P*-values for Fisher's exact test (2 sided)
Total combat levels sample ($n = 90$)	54.4	
Good A for Good B	47.9	0.398
Good B for Good A	61.9	
High combat levels ($n = 47$)	78.7	
Good A for Good B	78.3	<0.001
Good B for Good A	79.2	
Low combat levels ($n = 43$)	27.9	
Good A for Good B	20	0.01
Good B for Good A	38.9	
Medium low levels ($n = 24$)	45.8	1
Very low levels ($n = 19$)	5.3	<0.001

Note: Medium low levels are defined as scores from 74 down to 30. Very low levels are those below 30 and includes 13 complete novices with a combat level of just 3 and 0 months of experience. The 'percent traded' figure is, where relevant, the percentage of the sub-sample that trades.

the market price. Thirty-seven of our subjects are from the clan, while the remainder were obtained from the second venue. In both our sites we have a range of combat levels and as the table shows, the treatments are evenly distributed across the sub-samples: 49% of high levels and 58% of low combat levels were endowed with good B.

Mean trading intensities are higher for the high combat levels than low combat levels. Twenty-three subjects (but only one high level player) reported that they never or rarely traded objects. Nearly all of these players were concentrated amongst the novices. The remainder reported trading at least once a day on average, but as Table 1 indicates around half of the sample trade more than five items on average *per day*. It is worth noting that in List (2003), the most intense private trading category was 11 or more items *per month*. In other words, as measured by the volume of trades, most subjects here are trading at intensities several times higher than those encountered in the sports memorabilia markets. It is also worth noting that the five measures of experience and trading intensity were highly correlated (correlation with combat level in parentheses): combat level points; quest points (0.792); total points (0.939); years of experience (0.640) and high trading intensity (0.450).

The top row of the Table 2 depicts the summary trading statistics for the pooled data. It shows that 54.9% players decided to trade their item: 47.9% traded Good A for B and 61.9% gave up Good B for A. Given that all participants were randomly assigned good A or B, the proportion choosing A over B should be equal in the two treatments (Kahneman et al., 1991; Knetsch, 1989). Using Fisher's exact test, we accept the hypothesis of no endowment effect ($p = 0.398$).

However, when we disaggregate by combat level the pattern of trading is very different; a picture illustrated by Fig. 1. It points to the fact, also shown in Table 2, that high levels of combat are associated with a level of trading that typically exceeds 50%. This conclusion is robust to changes in the combat level of 75 as the dividing point between high and low experience players. Along with the rows in Table 2, Fig. 1 supports the notion that highly experienced players tend to over-trade (meaning that they are more likely than not to trade). When we conduct a two-sided Fisher's exact test that the proportion of high combat level players who choose A over B is independent of the endowment, this is firmly rejected with a p value below 0.001.

Conversely, when the low combat level group is tested we see that only 27.9% of players swap their endowment for the alternative. For this group, we reject the null hypothesis that final choice is not linked to endowment with $p = 0.01$, suggesting that there is strong endowment effect within the low combat level sample (column 2, Table 2). Fig. 1 suggests that the endowment effect is most strongly associated with very low combat levels. We split the low level data into two and test again. The medium low combat levels (combat levels of 30–74) sub-sample shows a trading rate of 45.8%, compared to 5.3% of very low levels (combat level below 30). Not surprisingly therefore the null hypothesis of no link between choice and endowment is accepted for the 24 medium low level players but strongly rejected ($p < 0.001$) for the 19 very low combat level players.

There was no significant difference in the pattern of results between the two sampling sites: low combat level clan members were resistant to trade in the same way that non-clan members were and higher combat level individuals behaved similarly in the two locations. Specifically, 19 wished to swap out of the 26 high-level combat score clan members while 18 out of 21 non-clan members with a high combat score wished to exchange their endowment for the alternative. Amongst the 11 lower level clan members, 2 wished to swap and for non-clan members 10 out of 32 chose to swap.[14] This difference is not statistically significant.

[14] Paying users have access to a wider set of skills and sites compared to free users, though we selected our goods so that they were useful for all types of users. We do not have clear data on whether subjects were free or paying users of Runescape. All we can say is that, across the two experiments, thirty-two subjects declared favourite skills that can only be utilised by paid members. Within this group, 76% of high combat level users swap and 27% of non-high combat level users swap, so we still get a positive association between skill level and willingness to swap. We tried adding a dummy for 'clearly paying users' skill-type' to the probit models in Table 6 but it did not change other results.

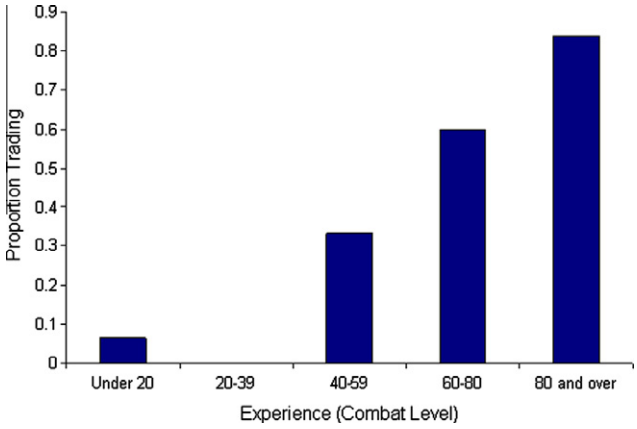

Fig. 1. Trading and combat levels in experiment 1.

There were differences in trading patterns conditional on familiarity: subjects who said they were familiar with their endowed good were more likely to swap it compared to subjects who were not familiar (31 out of 46 in the first category, 18 out of 44 in the second category, $p = 0.019$ for a Fisher's exact test). It may be that subjects who were familiar might also be owners of the object in which case (through a diminishing marginal utility argument) they might be more willing to trade than non-owners. We can say that experienced subjects were more likely to be familiar with their endowment. Indeed the familiarity variable is not significant once behaviour is conditioned on combat level via a probit model.

Our results are consistent with previous indications that market experience is correlated with a lower incidence of the endowment effect, but even for players with higher measures of experience, clear evidence of over-trading is not usual. Plott and Zeiler (2007), do find some overtrading with one treatment, reporting that 67% of students endowed with a pen end up with a mug (the other possible endowment), while 54% of students endowed with a mug end up with a mug ($p = 0.12$, two-sided Fisher's exact test, $p = 0.073$, one sided test). However, using a similar protocol, Knetsch and Wong (2009), find 49% mug-endowed subjects end up with a mug, whilst 50% of subjects endowed with the other good choose to take the mug (see also Isoni, Loomes, & Sugden, 2011). Amongst intense private traders, List (2003), finds 56% swap their endowment for the alternative, but this is not significantly different from 50% given the sample size. Similarly in List's (2011) field experiment also with private traders, 16 out of 29 subjects swap in the final trading session.

There is some relevant non-experimental data on the issue from financial markets. Dhar and Zhu (2006), use personal trading data from 7965 US households who hold stock portfolios, in order to examine field evidence for the disposition effect. The disposition effect is defined as the tendency for individuals to hold stocks that fall below their purchase price longer than stocks that rise higher than the purchase price. Now, the disposition effect is not the same as the endowment effect. Nevertheless, the same underlying explanation is often given for both of them (e.g. Kahneman et al., 1991): that individuals have reference dependent preferences and that a specific change in consumption is given more weight in decisions when it is viewed as a loss than when it is coded as a gain. In this theoretical framework, in a trading experiment the endowment establishes the reference point. Meanwhile in the disposition effect story, the buying price establishes the reference point. Dhar and Zhu (2006), note that "However, we find that despite the significant disposition effect on average, about one fifth of investors in our sample exhibit behaviour opposite to the disposition effect." p. 728. In fact, as their regression results show, while the household at the median level of trading intensity has a disposition effect, the opposite is true for traders at the mean level of trades, or higher. In a similar vein, Feng and Seasholes (2005), offer a very careful longitudinal analysis of the trading habits of 1511 individual stock traders from the People's Republic of China.[15] They show that within the 2-year window of observations, individual disposition effects decline. For a fraction of the traders with the most experience, a significant reverse disposition effect is observable.[16] To summarise, the phenomenon of overtrading is not unique to our dataset, but the scale and significance of overtrading we see is not in line with past evidence from experiments. There is some field evidence on the disposition effect that is compatible with our data, but it is from a context that is not exactly parallel to our environment.

Let us consider four possible reasons for the unexpected aspect of our results. One class of explanations of the endowment effect is based on the notion that individuals miss-forecast their own preferences (Loewenstein & Adler, 1995). Specifically, the attachment to a new good is underestimated and this makes subjects more inclined to hang onto their endowments. To

[15] An important feature of their data is that there is no tax reason for a disposition effect or its reverse in China. Moreover, since citizens may legally hold only one trading account, they claim that the trading records of the individuals are comprehensive.

[16] Within the set of experienced traders it tends to be younger, male individuals who are most likely to have a reverse disposition effect by the end of the sample period.

explain our results using this type of model, highly experienced players would have to *underestimate* the value of the endowment relative to the alternative good. This is by no means an unfamiliar notion in everyday life. In English, 'the grass is greener on the other side'[17] is a well-known saying summing up the pull of objects that are not possessed. Meanwhile, in Shakespeare's *Much Ado about Nothing* Friar Francis laments "That what we have we prize not to the worth." (Act 4, Scene 1). In keeping with Friar Francis, we call this possible relative underestimation of the value of the currently-held object, the 'prize not to the worth' effect. We did not ask players in experiment 1 to estimate the value of the alternative good, so we have no direct evidence on it.

A second possibility is that the high levels of trading represent experimenter effects of some kind. In particular some subjects might view trading as the point of the experiment and seek to conform. We noted earlier that conditional on combat level, clan members and non-clan members show a similar disposition to trade, so it is not a simple matter of differences in the reactions of the sub-samples to the experiment. Rather, the experimenter effect would have to vary according to the experience of the subjects. Levitt and List (2007), put forward a theory in which experimental behaviour will vary according to the monetary size of the rewards. For relatively low value rewards, experimenter effects loom larger in the minds of subjects, diminishing as the value of the goods at stake increases. In our experiment, we might therefore expect to see larger experimenter effects amongst high productivity subjects for whom the goods were of relatively low value. Since highly experienced subjects are also subjects with higher productivities, then this would explain the relationship between experience and overtrading. We cannot rule out this explanation on the basis of the evidence from experiment 1, but we note that it suggests that, if the goods were of higher value, overtrading would tend to diminish. We examine this possibility in experiment 2 below.

An alternative hypothesis is that there are in fact two competing drives amongst players: one drive is a reluctance to trade, sometimes denoted as 'caution' (Bateman et al., 1997). The second is Adam Smith's propensity to 'truck, barter and exchange'. Models explaining caution in trading have been put forward by a number of authors, including Huck, Kirchsteiger, and Oechssler (2005) and Carmichael and MaCleod (2006). A common theme of these works is that understating willingness to pay (or overstating minimum compensation required) may raise the expected payoff for individuals in some forms of bilateral bargaining. Such behaviour is not incentive compatible in simple exchange games such as the one used in this experiment, but the idea is that caution may become a default heuristic used in trading situations. However, in such models overtrading cannot be accommodated. Moreover in some contexts overtrading appears perverse, since with myopic agents it might imply cycles of exchange.

Theories that produce expected gains from cautious trading do so in the context of exchange. But in an economy where agents both produce and consume, being over-attached to the endowment may mean a person undervalues the gains from specialisation. In Smith's original argument, truck barter and exchange is necessary to exploit the gains from the division of labour:

> "But without the disposition to truck, barter, and exchange, every man must have procured to himself every necessary and conveniency of life which he wanted. All must have had the same duties to perform, and the same work to do, and there could have been no such difference of employment as could alone give occasion to any great difference of talents." (Smith et al., 1776, chap. 2, p. 21).

Specialisation is especially useful in Runescape for the reasons that Smith alludes to: concentrating on one or two skills raises individual productivity and thereby raises the ability to buy all sorts of useful goods. This feature of the virtual world might be responsible for a reduction in cautious behaviour, but it does not imply that overtrading is a good, long-run heuristic. In some cases, though, overtrading might be useful if it raises the probability of future transactions. For instance in Runescape, players often add other players to their private chat list when they engage in successful trade. An extensive address book is useful for future trading and cooperation in quests and so gives an extra incentive to make an initial contact end in an exchange.

Finally we might wonder if there is a selection process at work. Individuals attracted to trading environments, such as memorabilia markets, eBay, second-hand goods markets, collectors fairs and online role playing games may obtain utility from trade itself (for sociological evidence in support of this view see Crewe & Gregson, 1998; Yee, 2007) thereby having a higher propensity to truck and barter compared to the average person. Subjects with a low level of enjoyment from trade would tend to self-select out of the market, leading to a higher measured propensity to trade amongst experienced traders. In this interpretation, experienced players would not be overtraders because of what they have learnt from within the game, but because of some propensity that was innate or at least, learnt outside of the Runescape environment. We know that out of the 13 complete newbies in our sample, only 1 was willing to exchange their endowments, which suggests that new subjects do not enter the game with a high propensity to trade. Fundamentally though, in the absence of data on attrition rates amongst new players, we cannot eliminate the selection argument.

[17] Sunstein (1993, p. 224), writes that 'The popular notion that "the grass is always greener" suggests that preferences may be strongest for things to which people do not have entitlements, but there is apparently no empirical support for this intuitively plausible view.' This experiment seems to provide such empirical support.

3. Experiment two

3.1. Design

In the format we have here we cannot distinguish between the final two explanations (experience and selection), but we can modify the basic design to test the prize not to the worth hypothesis and to see if higher valued goods eliminate the incidence of overtrading. To make some headway we therefore ran experiment 2, in which the basic design is unchanged, but we make three adjustments. First, we use the clan subject pool to recruit only experienced players, since it is amongst them that overtrading seems to occur. The lowest combat level amongst the subjects is 50. Secondly, we use considerably higher value goods, selected after brief trials with a set of possible prizes. The goods selected were as follows:

Good A: A Warrior Ring, a decorative ring that adds four slash bonuses to any weapon when equipped.

Good B: An Uncharged Black Mask, a decorative mask that gives a 15% bonus in strength and attack during slaying tasks.

These rare items are valued between 400,000 and 500,000 gold pieces – i.e. about 150 times the value of the items in experiment A. This places them in the top 10% of most expensive items in the game, while their rarity limits their substitutability with money. They cannot usually be bought within the in-game shops and can only be obtained from trading with other players[18] or through quests. Even for highly experienced and productive players these products can only be obtained through many hours of endeavour. In other words, for nearly all experienced players 400,000 gold coins represent a significant proportion of their total wealth within the game.

For the third change, we extended the questionnaire to ask about the subject's perceived valuations of both goods (not just the endowment). Of course, questionnaire answers represent 'cheap talk' in that they are not incentivized, but responses may provide insight into whether the 'prize not to the worth' story is plausible.

3.2. Experiment 2: results

Table 3 provides the basic results. As can be seen, the null-hypothesis of no endowment effect cannot be rejected against a two-sided alternative. Indeed exactly half of the endowments are traded.

However, as Fig. 2 shows, the pattern of trading is still markedly different between different levels of experience in the game. For players with a combat level below 70, 9 keep their endowment and 3 swap. For the 18 players with a combat level at 70 or above 12 swap and only 6 keep. This difference is statistically significant at the 10% level (Fisher's exact test, $p = 0.060$, two-tailed test) and suggests that overtrading is not simply an artefact of using low value items. This result is also robust to changes in the definition of experience. Moreover if a probit regression is run with years of experience as the explanatory variable, then it is statistically significant at the 1% level. Of course we still do not know whether overtrading would disappear if we had used items in the top 1% of the value distribution or higher, but the results do not support a simple story that overtrading was due to the value of the items being traded.

To examine the 'prize not to the worth' hypothesis, we test to see if knowledge of the objects plays a role in behaviour. Thirteen of our subjects claim to know both objects, while nine profess ignorance of both. The remaining 8 know one object, but there is no link between the pattern of claimed knowledge and the decision to swap. Similarly, there is no statistically significant link between the relative price placed on the two objects and the decision to sell. Now, we have six data points where subjects placed extremely high or extremely low valuations on one of the objects (but not both). Given the large number of zeros that need to be entered to give an accurate price, there is a possibility of a typing error in these data points. When we add or subtract zeroes to bring the valuations into line with other estimates, the pattern of decisions is as shown in Table 4.[19] While this data is more suggestive of the 'prize not to the worth' hypothesis the results are not statistically significant at standard levels of significance. Thus there is no good evidence for 'prize not to the worth'.

We can use the data from both experiments to examine the relationship between favourite skill and propensity to swap. We code each person's favourite skill as either 'intermediate' or 'final'. For instance, combat is a final skill while woodcutting and farming are intermediate because individuals who specialise in these skills normally do so to sell the results of their production. The results are in Table 5, where we can see that individuals who favour intermediate production skills are more likely to swap their endowments. Some skills (e.g. cooking on the one hand or slayer and combat on the other) are easy to classify, while others are more ambiguous (e.g. thieving), so we redo the analysis, dividing skills between fighter skills (a subset of final production skills) and non-fighter. We find a similar, statistically significant difference between categories. This is consistent with a world in which subjects with a higher propensity to trade are more willing to specialise in production activities that have only indirect benefits. However, we shall see below in Table 6 that this correlation is not robust when controls for combat level and other measures of experience are brought into the statistical model.

[18] For experiment 2, all the items selected were purchased from other players at a typical price of 500,000 gold coins per item. Funds for the purchase were generated by drawing down existing savings and through several weeks of intensive production and trade within the game by one of the authors.

[19] When we use the uncleaned data the respective fractions are 5/12 and 10/18 – a similar pattern but still further from significance.

Table 3
Trading (experiment 2).

Variable	Percent traded	*P*-values for Fisher's exact test (two-sided)
All ($n = 30$)	50.0	
Good A for Good B	46.0	$P = 1.00$
Good B for Good A	54.0	

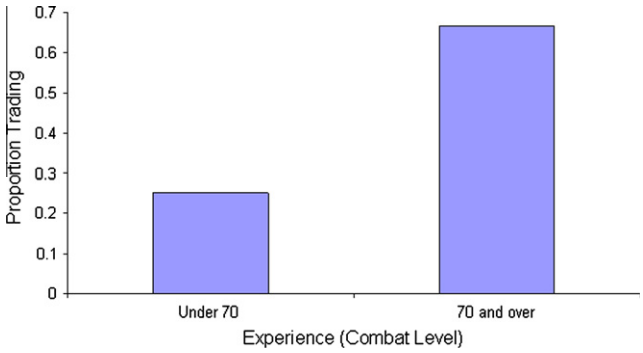

Fig. 2. Trading in experiment 2.

Table 4
Relative valuations and trading.

Relative valuation	Trading (*x* goods swapped out of *y* subjects in the relevant category)	*P*-value for Fisher's exact test (one sided)
Alternative priced lower	4 from 11	0.225
Alternative priced equal or higher	11 from 19	

Table 5
Traders versus fighters.

Favourite skill	Sub-sample size	Percentage traded	*P*-values for Fisher's exact test (two sided)
Intermediate	69	60.9	0.041
Final	51	43.1	

Table 6
Probit model estimation results.

Dependent variable	Trade	Trade	Trade	Trade
Combat level	0.025*** (0.007)	0.028 (0.043)	0.028 (0.041)	0.021*** (0.0077)
Years of playing	0.149 (0.155)	3.160*** (0.963)	3.201** (1.31)	0.454** (0.208)
Endowed with good A dummy	−0.020 (0.303)	−0.604 (0.621)	−0.751 (0.731)	−0.162 (0.262)
Experiment 2 dummy	–	–	–	−0.118 (0.296)
Relative price of endowment	–	–	−0.564 (0.920)	–
Tradeskill dummy	–	–	–	0.008 (0.278)
Hightrader	–	–	–	−0.124 (0.309)
Constant	−1.732*** (0.460)	−4.48 (3.12)	−3.91 (2.71)	−1.74*** (0.469)
Experiment	1	2	2	Combined
Sample size	90	30	30	120

1. The Trade dependent variable equals 1 if the subject chose to trade; 0 otherwise.
2. The robust standard errors are in parentheses.
** Significant at 5% level, one tailed tests.
*** Significant at 1% level, one tailed tests.

Finally we use the results from both experiments and check for other factors that might affect the propensity to trade. The results are reported in Table 6. To find correlations with trading behaviour we use a probit model with the following underlying specification:

$$\text{Trade} = \beta_0 + \beta_1 \text{Item dummy} + \delta Y + u$$

where the dependent variable, Trade, equals 1 if the subject chose to trade, and 0 otherwise; Item dummy takes the value 1 if the player is endowed with good A and zero otherwise and Y is a vector of within game characteristics such as combat level and years of experience playing the game. The error term is u.

None of the player characteristics (gender, age, location) are remotely significant in this specification, so we drop them from the reported equations. The close correlation between tradeskill, hightrader, age, etc. means that in finding predictors of trading behaviour there is an issue of multicollinearity. In fact, as the final column of Table 6 reveals, the variables for favourite skill and level of trading[20] are not significant in the presence of the combat level and experience variables. We have to be careful about over-interpreting the difference between the effects of trading intensity and combat level and other measures of experience. Not only are these correlated variables, the trading intensity variable is based on interval data which may produce measurement error. The error may be compounded by the fact that the combat level score is publicly observable, whereas the trading intensity data is based on each subject's recall of his or her own history. With these reservations, it does seem as though trading intensity per se does not affect the incidence of the endowment effect. On the other hand, the equations estimated in Table 6 confirm the importance of both experience and combat. For instance, using the probit equation of the final column and setting all other right-hand side variables to their sample means, adding one more year of experience and altering combat level to 75 from 0 raises the estimated propensity to swap by 0.78.

Within experiment 2, the years of playing variable continues to be statistically significant in its association with the probability of trading. However, perhaps because the range of combat levels is smaller than that for the whole sample, the combat level variable is not significant. The relative price of the endowment (i.e. perceived market value of the endowment divided by the perceived market value of the alternative) does not have a significant effect on the probability of trade, although the sign is in line with theory. As we noted above, there are question marks over the measurement of perceived prices in some cases. Whether we use the data cleaned of 'extraneous' zeros or not does not alter the conclusions of this paragraph. It might be viewed as surprising that relative price does not have a significant effect, but it is worth pointing out two things. First, 10 out of 30 subjects report equal valuations for the two goods, limiting the variation in the relative price variable. Secondly, a simple model with just "years of playing" as the independent variable correctly classifies all but 4 of the decisions. So adding extra variables could add little to the explanatory power of the model.

4. Discussion

We conduct an experiment on the endowment effect in a virtual field – an online role-playing game. We show that this is feasible and indeed relatively straightforward. In terms of the actual endowment effect experiment, much of the underlying pattern of results is consistent with previous work on the endowment effect. Magical runes and imaginary jewellery produce effects similar to those obtainable from mugs and pens. In line with List (2003), individuals with lower levels of trading experience exhibit a great reluctance to trade, but subjects who have traded in the market for longer show no endowment effect. In fact, the overall results of this study showed an unusual high level of trade – 54.5% for all sample groups and this is where our results depart significantly from those obtained in prior research. For higher levels of experience there was robust evidence of overtrading – in experiment 1 over 80% of subjects with combat levels above 80 chose to swap their endowment for the alternative. Players were more cautious in their behaviour when the goods were of significantly higher value, but in experiment 2 mean swap rates were still significantly above 50% for experienced players. Rather than day-to-day levels of trading (which were high for most players) it was years of experience playing the game and a player's combat level which were most closely correlated with swapping behaviour.

The evidence is compatible with both selection and experience-based explanations of trading behaviour. An experience-based argument – and one in keeping with Adam Smith, is that a propensity to truck, barter and exchange may be a beneficial heuristic in contexts where the individual is both producer and consumer. In this situation the cautious trading heuristics of inexperienced players may be gradually eradicated by the benefits obtained from specialisation and subsequent trade. This explanation is entirely speculative, but it is compatible with our evidence. An alternative, selection-based explanation could centre on the notion that trading can be a leisure activity and that certain kinds of markets can attract individuals who have a high propensity to truck, barter and exchange. Yee (2007) and Kallio, Mäyrä, and Kaipainen (2010) suggest that the motives of online MMORPG players are not so different from motives of individuals in other social situations. Choi et al. (2007), offer survey evidence that players in Mabinogi (a MMORPG popular in South Korea) have a mix of trading motives that includes the desire to maximise rewards along with (for some), intrinsic pleasure from trading. High levels of buying and selling can be seen in many other consumer activities (Crewe & Gregson, 1998), including collector's memorabilia markets, 'car-boot' sales and, possibly, online auction sites. Under this reading, field experiments conducted in such markets

[20] It is probably not surprising that the level of trading is not significant given that the vast majority of our subjects were trading at a rate far more intense than the level of the memorabilia swappers in List's (2003) study. Forty-five percent of our subjects for instance, claimed to be trading at least 150 items per month.

would provide a poor guide to the stability of the endowment effect in markets where there no players who obtain utility from the act of trading.

Whatever the conclusion drawn about the causes of overtrading amongst highly experienced subjects this research, like that reported in Chesney et al. (2009), has shown that virtual worlds may be useful platforms for testing economic theories and provide a further venue for economic experiments. We therefore end with some final observations about the methodological questions that arise in running an experiment of this kind. One pertinent issue for replication and extensions to experiments concerns the permanence of MMORPGs. Yee (2006), notes that virtual worlds as whole are becoming more popular, particular games rise and fall in popularity. Several apparently successful worlds have closed[21] and Second Life laid off 30% of its labour force in 2010.[22] A relatively high turnover of platforms would limit the ability to replicate experiments.

We had few problems recruiting subjects, but as noted earlier there were isolated examples of subjects disappearing with the endowments before the end of the experiment. Data recording was cumbersome, compared to a typical computer-based experiment since answers and questions had to be cut and pasted from a chat system within the game. However this also meant that the events of the game were genuinely sequential – unlike paper-based surveys or experiments subjects could not look ahead to subsequent questions. As with Chesney et al. (2009), we note that communicating through instant messaging or chat eliminates the ability to communicate through non-written means such as tone of voice or body language. This may in some cases limit subject understanding but at the same time it leads to a greater uniformity in the interactions between researcher and subject compared to a typical real-world (field) experiment.

In this particular game, the fact that exchanging virtual goods or virtual money for real money is banned meant that one of the researchers had to work inside the game to earn the endowments before the experiment could be conducted. The feature created some potential problems of sample size for experiment 2, where several weeks of prolonged and mundane work were required to fund the goods. It also meant that it was particularly useful to have one researcher with extensive experience of the game. Other on-line environments vary in their restrictions on exchange between virtual and real worlds and so might not pose the same challenges. However, from the viewpoint of an endowment effect experiment, easy trading between virtual goods and real money would have been a disadvantage, since this would have encouraged subjects to view the goods as close substitutes for money.

Acknowledgements

The Economics Department at Royal Holloway College, University of London provided financial support for the experiments reported here. We are grateful to the comments of Robert Hoffmann and other conference participants at the Asia Pacific Economic Science Association meeting in Kuala Lumpur, February 2010. We are especially thankful to the editor of this journal and three referees whose numerous and thorough suggestions led to many improvements to the manuscript.

References

Bateman, Ian J., Munro, Alistair, Rhodes, Bruce, Starmer, Chris, & Sugden, Robert (1997). A test of the theory of reference-dependent preferences. *Quarterly Journal of Economics, 112*(2), 479–505.

Bilir, Tanla E., 2009. *Real economics in virtual worlds: a massively multiplayer online game case study*, Runescape, SSRN. <http://ssrn.com/abstract=1655084>.

Carmichael, H. Lorne, & Macleod, W. Bentley (2006). Welfare economics with intransitive revealed preferences: A theory of the endowment effect. *Journal of Public Economic Theory, 8*(2), 193–218.

Castronova, Edward, 2002. *On virtual economies*. Cesifo working paper series no. 752.

Castronova, Edward, 2008. *A test of the law of demand in a virtual world: Exploring the petri dish approach to social science (July)*. CESifo Working paper series no. 2355.

Chesney, Thomas, Chuah, Swee-Hoon., & Hoffmann, Robert (2009). Virtual world experimentation: An exploratory study. *Journal of Economic Behaviour and Organization, 72*(1), 618–635.

Choi, Boreum, Lee, Inseong, Lee, Kiho, Jung, Seungki, Park, Sunju, Kim, Jinwoo, 2007. The effects of users' motivation on their perception to trading systems of digital content accessories: Focusing on trading items in online games. In *Hawaii international conference on system sciences* (pp. 1530–1605). doi: {http://doi.ieeecomputersociety.org/10.1109/HICSS.2007.527}.

Crewe, Louise, & Gregson, Nicky (1998). Tales of the unexpected: Exploring car boot sales as marginal spaces of contemporary consumption. *Transactions of the Institute of British Geographers, New Series, 23*(1), 39–53.

Dhar, Ravi, & Zhu, Ning (2006). Up close and personal: An individual level analysis of the disposition effect. *Management Science, 52*(5), 726–740.

Feng, L., & Seasholes, M. S. (2005). Do investor sophistication and trading experience eliminate behavioural biases in financial markets? *Review of Finance, 9*(3), 305–351.

Fiedler, Marina, & Haruvy, Ernan (2009). The lab versus the virtual lab and virtual field – An experimental investigation of trust games with communication. *Journal of Economic Behaviour and Organization, 72*, 716–724.

Huck, Steffen, Kirchsteiger, Georg, & Oechssler, Jörg (2005). Learning to like what you have – Explaining the endowment effect. *Economic Journal, 115*(3), 689–702.

Isoni, Andrea, Loomes, Graham, & Sugden, Robert (2011). The willingness to pay willingness to accept gap, the endowment effect, subject misconceptions, and experimental procedures for eliciting valuations: comment. *The American Economic Review, 101*(2), 991–1011.

Kahneman, Daniel, Knetsch, Jack L., & Thaler, Richard H. (1991). The endowment effect, loss aversion, and status quo bias: anomalies. *Journal of Economic Perspectives, 115*(1), 193–206.

Kallio, Kirsi P., Mäyrä, Frans, & Kaipainen, Kirsikka (2010). At least nine ways to play: Approaching gamer mentalities. *Games and Culture, 6*, 327–353.

Knetsch, Jack L. (1989). The endowment effect and evidence of non-reversible indifference curves. *American Economic Review, 79*(5), 1277–1284.

Knetsch, Jack L., & Wong, Wei-Kang. (2009). The endowment effect and the reference state: Evidence and manipulations. *Journal of Economic Behaviour and Organization, 71*(2), 407–413.

[21] The forum, http://www.globalmmo.com/mmorpg-list/mmorpg/cancelled lists over 20 cancelled sites as of 2010.

[22] Source: http://www.globalmmo.com/mmorpg-list/mmorpg/cancelled.

Levitt, Steven D., & List, John A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives, 21*(2), 153–174.
List, John A. (2003). Does market experience eliminate market anomalies? *Quarterly Journal of Economics, 118*(1), 41–71.
List, John A. (2004). Neoclassical theory versus prospect theory: Evidence from the marketplace. *Econometrica, 72*(2), 615–625.
List, John A. (2011). Does market experience eliminate market anomalies? The case of exogenous market experience. *American Economic Review, 101*(May), 313–317.
Loewenstein, George, & Adler, D. (1995). A bias in the prediction of tastes. *Economic Journal, 105*(43), 929–937.
Plott, Charles R., & Zeiler, Kathryn (2007). Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory? *The American Economic Review, 97*(4), 1449–1466.
Shogren, Jason F., Seung, Y. Shin, Dermot, J. Hayes, & James, B. Kliebenstein (1994). Resolving differences in willingness to pay and willingness to accept. *American Economic Review, 84*, 255–270.
Smith, Adam, 1776. The wealth of nations (Vol. 1). Penguin.
Sunstein, Cass R. (1993). Endogenous preferences, environmental law. *The Journal of Legal Studies, 22*(2), 217–254.
Yee, Nick. (2006). The demographics, motivations and derived experiences of users of massively-multiuser online graphical environments. *Teleoperators and Virtual Environments, 15*, 309–329.
Yee, Nick (2007). Motivations of play in online games. *Journal of CyberPsychology and Behaviour, 9*, 772–775.