

# Trusting the avatar

25-10-2004

8700 words

DRAFT

Jonas Heide Smith ([smith@itu.dk](mailto:smith@itu.dk))

PhD candidate

Center for Computer Games Research

Dept. of Digital Aesthetics and Communication

The IT University of Copenhagen

## ABSTRACT

Playing online entails far more than dragon-slaying, identity experimentation, communication and elaborate synchronized aggression. Letting other players affect your experience means placing yourself – if ever so slightly – at their mercy. In other words you need to trust them, evoking a concept which has divided philosophers for millennia.

This article describes how trust is an important factor in multiplayer gaming, how it can be approached theoretically and how it may be favourably affected by game design.

## INTRODUCTION

History, I'm sure you'll agree, has not been kind to Plato's brother. In fact, it would be a testament to your above-average memory if you even appreciate the fact that Glaucon, Plato's older brother, espoused a theory of trust. Since philosophers have so overwhelmingly favoured the younger brother, let us hear what the older one had to say. Considering what would happen if a man (whether normally considered just or not) suddenly did not face the threat of punishment, Glaucon was certain that

No man would keep his hands off what was not his own when he could safely take what he liked out of the market, or go into houses and lie with any one at his pleasure, or kill or release from prison whom he would, and in all respects be like a God among men. ...a man is just, not willingly or because he thinks that justice is any good to him individually, but of necessity, for wherever any one thinks that he can safely be unjust, there he is unjust. (Plato, 360 BC)

Glaucon believes that no man is inherently just. In his view, we can only trust others not to steal our possessions or otherwise transgress against us to the extent that those others fear punishment.

We can think of this as the cynical perspective (although this colloquial label should not lead us to dismiss it at this point). Not surprisingly, prominent philosophical figures have disagreed with Glaucon. For instance, David Hume and Immanuel Kant believed in an innate human morality that made us take pleasure in the wellbeing of others. Before looking more closely at the concept and philosophy of trust let us look at a few examples of what is at stake regarding multiplayer gaming. We will return to those examples continuously throughout this text and end with a more general analysis of how trust

(and distrust) manifests itself in multiplayer gaming.

### Game examples

To understand multiplayer games we may want to compare it to something else that it is not. Let us begin, then, with *Moon Patrol* (Irem, 1982) a single-player arcade game. In *Moon Patrol* the player assumes the role of Captain Thorgon in charge of a buggy-like armed vehicle patrolling (as the title hints) a lunar landscape (see Figure 1).



Figure 1 – Moon Patrol (Irem, 1982)

Although we have yet to define trust precisely, in the standard sense of the word you have little need for trust when playing the game on your own. Like many other arcade games of the time, however, *Moon Patrol* allows two players to play in a turn-based fashion. When one player loses a life the other takes over and so on. In terms of the game world the two games that are being played are unrelated; one player's actions do not influence the other player's situation. But here the game situation is irreducible to the game world (something which is fact always the case as we shall

see). Since the players are physically right next to each other one may of course affect the other, whether by distracting, obstructing or helping. We shall return to this issue below.

These dynamics were quite different for players of *Gauntlet* (Atari, 1985). Here up to four players – each choosing one of four avatars with different strengths - could cooperate against the monstrous enemy hordes (see Figure 2).



Figure 2 – Gauntlet (Atari, 1985)

Although sharing the same grand goal each player had an individual score and therefore a certain incentive to gobble up coins and healing potions conveniently littering the monster-filled dungeons. Players were not able to hurt each other directly, however (in modern game parlance there was no ‘friendly fire’<sup>1</sup>).

Also supporting four simultaneous players, *Top Spin* (Power\_and\_Magic, 2003) is different from *Gauntlet* in the sense that players or teams of players are entirely

<sup>1</sup> A somewhat confusing term referring to the fact that fire from allies does not “count” in terms of the game. Such fire does no damage.

opposed. In this tennis game for the Xbox, players are still in the same physical location<sup>2</sup> but have no game-internal reason for cooperating with the opponent(s) (see Figure 3).



Figure 3 – Top Spin (Power\_and\_Magic, 2003)

The same is true for players of the highly popular team-based first-person shooter *Counter-Strike* (Valve, 2000). Again, the team wins or loses but the game’s strong focus on individual achievement, something prominently displayed on the score sheet, also creates a certain tension between team members. More importantly, however, the



Figure 4 – Counter-Strike (Valve, 2000)

game is usually played online with allies and opponents in different physical locations (see Figure 4).

Much the same features characterize the real-time-strategy game *Age of Kings* (Ensemble\_Studios, 1999). In *Age of Kings* players vie for control of the game map in a constant trade-off between technological advancement and military strength (see Figure 5). Once inside the game trust issues are highly dependent on scoring conditions. In this particular title, however, player matching opens a range of trust issues not related to the core game rules (J. H. Smith, in review) nor found in *Counter-Strike* which relies less on pre-game conversation and negotiation.



Figure 5 - Age of Kings (Ensemble\_Studios, 1999)

Finally, players of *Star Wars Galaxies* (Sony\_Online\_Entertainment, 2003) face quite a different social situation. *Galaxies* is a massively multiplayer online role-playing game (an MMORPG) in which hundreds or thousands of players inhabit the same fantasy themed world (see Figure 6). As in *Gauntlet*, players are not in direct opposition and are unable to directly hurt each other. Nevertheless, various factions

<sup>2</sup> The game can also be played online, but we’ll concentrate on the offline game mode here.

have opposing interests and certain in-game activities (like transferring items from one character associated with an account to another) often depend on trusting other players.



Figure 6 – Star Wars Galaxies  
(Sony\_Online\_Entertainment, 2003)

### What is trust?

There is a strong tradition prescribing that work on trust must begin by emphasising the elusive, super-complex nature of trust. But while surely a non-trivial topic worthy of deep thought there is nothing particularly elusive about the phenomenon itself. A few ambiguities must be untangled, however.

First of all, there is less than total agreement as to the relationship between trust and certain institutions. Consider the following. You are a game developer who has been asked to produce a game for the website of a private company. Your contact at this company makes a verbal promise to pay you upon delivery of the game. Just to make the case clear-cut let's assume that no written document exists which describes your business relationship. Should you

choose to keep your end of the arrangement, delivering the game in time without any security that the customer will live up to his side of the bargain you are surely displaying a great deal of trust. But imagine a different scenario in which you and the customer draw up a contract specifying the details of the order (price, delivery date, sanctions related to breaking the contract etc.). In this case the relationship is quite safe for both parties (or rather: one party is not safer than the other). But does the contract increase the level of trust or should we rather see it as a means to eliminate the need for trust altogether? Is a contract-regulated relationship one of high trust or one where trust plays no role?<sup>3</sup>

In a sense this is merely a question of definition. On the other hand, the position that trust can only exist when people truly place themselves at the mercy of others run the risk of ignoring that even without contracts or other institutions we do not trust indiscriminately. There are those into whose hands we would not let ourselves fall backwards. Some do not display the right signs, an issue which we will return to below. Here it will suffice to note that trust in terms of this article should be understood in the broader way – it can be strictly interpersonal or it can be supported by institutions.

But what does it mean to trust? To trust means, in the terms of this article at least,

---

<sup>3</sup> We ignore here the fact that contracts can in certain situations actually destroy the fabric of a relationship. A prenuptial agreement (unless imposed by external authorities) is a case in point.

to expect future cooperation. Cooperation here should be understood as behaviour (or choices between courses of action) which is beneficial to the person who trusts. This expectation is not measured in the way that, say, blood pressure is measured. Rather, it is measured indirectly through observed behaviour. To be more precise, then, trust is present to the extent that somebody acts in a way which implies an expectation of future cooperation from someone else. When you deposit money in the bank you show trust. When you lend money to a friend you show trust. And when you leave your child in somebody's care you show trust.

The more interesting question, of course, is *why* you show this trust.

#### **Why we trust**

Glaucon's perspective on this matter, sceptical or not, has attracted many. Famously, Niccoló Machiavelli felt that people could only be trusted to follow their own selfish interests (Machiavelli, 1532/1950). Agreeing with this principle, political philosopher Thomas Hobbes expressed his belief that mankind was motivated by greed and fear of others (who were considered equally greedy) and needed a powerful sovereign to ensure trust between members of a society (Hobbes, 1651/1997). Although often phrased less dramatically, the idea that trust is a function of perceived personal consequences is one highly compatible with the rational agent model of human behaviour residing at the heart of classical economics. In this model, you would only trust a *Counter-Strike* opponent to play fairly and not violate local rules (such as the

common rule against "camping"<sup>4</sup>) to the extent that this opponent could be punished (or rewarded for cooperating). Let us call this perspective the *Homo Economicus* approach to trust.

*Homo Economicus* is sometimes contrasted to *Homo Psychologicus*, a creature which is either more kind-hearted or whose behaviour is not a consequence of cold-blooded rational analysis of a situation. Although this glosses over individual differences we can think of David Hume and Immanuel Kant as belonging to this tradition. Both had a far more positive view of human nature than Glaucon and his followers. In this perspective you should expect your *Counter-Strike* opponent not to cheat because it is not the right thing to do and human beings have a shared sense of morality.

In the recent decades a somewhat different approach has been suggested and continuously strengthened by a substantial amount of research. In an important sense, this perspective has been able to bridge the gaps between *Homo Economicus* and *Homo Psychologicus* as it has explained the obvious observation that people trust beyond what classical economics might predict but clearly have a pattern to their trust. We can call this the *Homo Biologicus* perspective. Evolutionary biologists and their colleagues have argued (and to a large extent documented) that risk-filled cooperation which requires trust can sometimes be a sensible strategy even if you're facing the harshness of natural selection (see for

---

<sup>4</sup> "Camping" is the less-than-brave strategy of hiding in a safe spot to safely dispatch with the enemy as he or she walks by.

instance Vogel, 2004). If you look at the world from an evolutionary perspective (that is from the perspective of genes, not individuals), consider the fact that much human interaction is shaped by the advantages of cooperation, and consider ways in which non-cooperative behaviour may be kept under control<sup>5</sup> many patterns of trust become explicable. For instance, from a genetic point of view it makes perfect sense that you should trust close family members more than strangers and that seems to very much be the case for most of us.

In this perspective, the question of why you should trust your *Counter-Strike* opponents is a complex one. But generally we should expect people to be risk-averse although they may well be willing to give strangers a chance to prove their good intentions (Axelrod, 1984).

Both evolutionary psychology and experimental economics are providing indications that basic patterns of human cooperation (and trust) should be considered in the light of our evolutionary past. While the importance of this theoretical development is wide-ranging we may be justified in continuing this discussion on trust by appealing to the approximation that people are selfish in something approaching the *Homo Economicus* sense. In the following, in other words, we'll be conceptualizing trust in terms that are more mechanical and

rational (in the sense of selfish calculation) than those justified by the evidence.

More generally, of course, we trust because we must. We cannot achieve perfect information about those we interact with. You cannot run a comprehensive background check on every employee in your bank and you certainly cannot investigate the criminal record of everyone who comes close to you on the street. But of course you usually need not worry since a multitude of features (law enforcement, contract law, our ability to remember faces) usually help ensure that people's interests are quite aligned (you don't *want* to steal my money since it would probably mean going to prison). This brings us back to games, since games are structures built to establish opposing interests in some form.

### Games and trust

Before we go on, we need to distinguish between two aspects of games: *the rule system* and *the gaming context*. The former concerns winning conditions and attribution of game points (the ideal game, if you will). Here, we invoke an idealized player who may be classified as a *greedy gentleman*<sup>6</sup>. This idealized player is completely in line with the winning conditions specified by the game, wants to win but also acknowledges the importance of sportsmanship. Another way to phrase the criterion for sportsmanship would be to acknowledge that the idealized player respects the spirit of a game. Sitting down

---

<sup>5</sup> Recent studies have shown that punishing trust-breakers implies a neural reward (Quervain et al., 2004).

---

<sup>6</sup> At least for highly competitive games. In games such as *The Sims* (Maxis\_Software\_Inc., 2000) or *EverQuest* (Verant\_Interactive, 1999) our idealized player may be less greedy.

for a game of chess he or she does not shake the table continuously in order to distract the opponent even if this possibility is not addressed by the rules.

On the other hand, the *gaming context* includes all aspects of how the game is actually played (e.g. whether players are in the same room, whether players can communicate before the game starts, whether someone cheats, how long single games take).

### **Trust and the rule system**

If we stick with the *rule system* for now, three or four of the six games introduced above have no trust issues. The single *Moon Patrol* player has nobody inside the game which he or she would consider trusting. The aliens are honest in the sense that they don't break any promises – they merely attack without prejudice. The very same goes for one's *Top Spin* opponent(s). The rules of tennis do not introduce any temptation to cooperate with the other side; one side's gain is the other side's loss<sup>7</sup>. You would never trust your opponent to do anything but his or her best to defeat you.

The issue is slightly more complicated in *Counter-Strike* and *Age of Kings*. *Counter-Strike* teams are entirely opposed and in this sense are comparable to the two sides in *Top Spin*. As regards the rule system, a player can never benefit from assisting the other side. Within the team, however, the situation is equivalent to that of *Gauntlet*. One is unlikely to win the larger battle without cooperating but as the game

specifically ranks individuals there is competition between team-mates. In situations where the emphasis is clearly on team achievement (such as tournament play) inter-team competition becomes less of an issue.

Interestingly, for *Age of Kings* players, trust is very much a setting. The game sports a large number of variations each affecting in-game trust dynamics. Mainly, players can choose to play alone or in teams. Teams are subject to the dynamics affecting *Counter-Strike* teams, since player-specific ranking is constantly displayed (to be further elaborated upon as the game ends). Teams, however, may be locked or unlocked. In the latter case, alliances may shift throughout the game while in the former the player is spared the constant worry about the scheming of current allies. Also, the winning condition known as "Last man standing" introduces the unpleasant prospect that once one team is victorious, the alliance will disband and former team-mates must fight each other for ultimate victory. Clearly, this makes for extremely uneasy alliances as no single player will be likely to sacrifice position or wealth to further the interest of his or her team unless fairly certain that team-mates are making similar sacrifices or facing near-certain team defeat.

Finally, *Age of Kings* players may also share a civilization. In this case, two or more players share full control of an in-game "player". Here, the cooperating players share the exact same incentives and have no reason not to trust each other fully.

One thing connects the trust-dependent game aspects mentioned above. They fall

---

<sup>7</sup> This is only true for stand-alone matches.

into the category of interaction forms that economic game theory classifies as zero-sum games<sup>8</sup>. Zero-sum games are situations (modelled in the language of games) of conflicting interests in which one person’s loss is another person’s gain. To some degree competing companies are engaged in a zero-sum game. The interests of phone company 1 (get more customers) is directly opposed to the interests of phone company 2 (get more customers)<sup>9</sup>. *Top Spin* is most clearly the same and very easily modelled:

|  |               |
|--|---------------|
| Player 1= Win  | Player 1=Lose |
| Player 2=Lose  | Player 2=Win  |
| The two possible end conditions of <i>Top Spin</i> . |               |
| “Player” can be substituted for “Team”               |               |

We’re assuming here that losing the match is as “bad” as winning the match is “good”. This means that we could assign the value 1 to a victory and a -1 to a defeat. Any one match would then have a total sum of zero.

This is not the case in *Gauntlet* and *Star Wars Galaxies*. In *Gauntlet*, players (and we’re still dealing with idealized ones) try to maximize their individual score and the progress of the adventuring party. But

since the game characters have different strengths and weaknesses and since some enemies are simply very difficult to kill without cooperation, one player’s score is dependent on the actions of the other player(s). Technically, one player would be able to gobble up more coins and potions than the others but if everybody played selfishly, the total score would probably be less than it could have been. For the sake of argument, let’s assume that any one player has two options – she can either play altruistically or selfishly. Also, let’s begin with a scenario with only two players, for instance the valkyrie and the wizard.

|                 |            |                            |                            |
|-----------------|------------|----------------------------|----------------------------|
|                 |            | <b>Wizard</b>              |                            |
|                 |            | Altruistic                 | Selfish                    |
| <b>Valkyrie</b> | Altruistic | Valkyrie = 2<br>Wizard = 2 | Valkyrie = 0<br>Wizard = 3 |
|                 | Selfish    | Valkyrie = 3<br>Wizard = 0 | Valkyrie = 1<br>Wizard = 1 |

Numbers indicate points earned.

In the table above points have been assigned to the two playing styles in a way which merely illustrates the following: Playing selfishly is tempting since it yields the highest payoff but if both players play selfishly monsters are not dealt with optimally leading to a suboptimal payoff. If you play selfishly and I play altruistically, you get all the points.

Quite possibly, you do not agree that this is an entirely reasonable depiction of the *Gauntlet* point attribution. You will probably agree, however, that there is a tension between selfish and altruistic play (brought on by a temptation to be greedy) and that when both players act selfishly they are not going to get very far in the game. The great advantage of the table

<sup>8</sup> For a discussion of the relationship between economic game theory and computer games see (J. H. Smith, in review).

<sup>9</sup> In reality this is only approximately the case – both companies, for instance, have an interest in bettering the operating conditions for phone companies. Interestingly, legislation ensures that the relationship stays zero-sum. Without anti-trust laws dominant companies could make deals raising the total score by coordinating prices.



above is that it introduces the most famous phenomenon to emerge from economic game theory: *The Prisoner's Dilemma*.

The Prisoner's Dilemma is a classic way of modelling a basic trust problem – we (me, you, our neighbours) may be tempted to choose a course of action that would lead to disaster if everybody made that choice (Axelrod, 1984; J. H. Smith, in review). Informally, the two players involved (who cannot communicate) are each presented with two options. They can choose *to cooperate* or choose *not to cooperate* (usually referred to as “defection”) with one another. If they both play it nice they both walk away with a decent reward (corresponding to both *Gauntlet* players playing altruistically). But from the perspective of either player he or she can do better by not cooperating no matter what he or she thinks the other player will do. In the tragic logic of this thinking, both players will choose the selfish path and their combined score will be far lower than it could have been (in our *Gauntlet* example it would be 2 instead of 4).

The Prisoner's Dilemma, in contrast to tennis, does not have a fixed sum. The total number of points depend on the behaviour of the players, making the game belong to the (unimaginatively named) category of *variable sum games* or *non-zero sum games*. Figuratively speaking, in zero-sum games, players share a single cake while in a prisoner's dilemma the size of the cake varies (among a total set of three sizes) according to the actions of the players. In *Gauntlet*, of course, the size of the cake is allowed to vary almost indefinitely. The same is true for *Star Wars Galaxies*. Here, the score of each player within the game

world does not detract from the total points available to others. To encourage collaboration (which in this case is seen as an end in itself) the character classes have specializations that make it rational (even for a selfish player) to group with others.

In sensibly formed groups, the whole is greater than the sum of the parts; another way of saying that the relation is non-zero sum<sup>10</sup>. But how can you know that group members will not take advantage of you? In fact, the game resolves such trust problems by automatically distributing experience points based on individual contributions and by allowing group leaders to switch on a feature which evenly distributes any loot acquired among the group members.

Before proceeding to the wider gaming context we should acknowledge that the perspective introduced above does not preclude an understanding of single-player games as building on trust. For instance, a player may well wonder if the AI in a given game “cheats” (for instance by knowing more than seems realistic) and in such a case trust becomes an issue (even if primarily an issue between the player and the game designer).

---

<sup>10</sup> This is a stronger version of the technique known from *Gauntlet* which in turn builds on the non-zero sum inter-character class relationship known from role-playing games like *Dungeons & Dragons*. In economic terms, this is surplus value derived from specialization; a concept famously explored by Adam Smith.

### Trust and the gaming context

So far, we have looked at idealized versions of the games in which highly disciplined uniform players contend themselves with the core rules while seeking to comply with the spirit of the game. Reality, we should not be surprised, is a little different.

First of all, real players are not uniform. Sticking to the economics terminology, they have different utility functions. For instance, a novice *Counter-Strike* player might consider herself quite successful to score but a few frags against experienced tournament players; in a sense she wins even if the game rules specify that she loses. Similarly, if you wish to introduce someone (say, a child) to *Age of Kings* you may *want* the child to win in order to encourage its interest. But this objection to the limited rule system view goes further. You could, easily in fact, argue that while technically *Top Spin* is a zero-sum game, the *act of playing Top Spin* is not. Although most people may prefer winning, actually playing the game, even if you lose, may be highly entertaining. Few people leave gaming sessions in which they didn't win feeling that the activity was a complete waste of time.

Secondly, trust issues very often result from features of the game environment, the wider context in which the game is played. As we shall see, whether a game is played online or not can have a tremendous influence on the social fabric of the player relationship.

Clearly, then, taking into account the gaming context complicates things greatly. However, in order to speak sensibly about trust we need to limit our perspective by



Figure 7 – Star Wars Galaxies (Sony Online Entertainment, 2003). A group shares the experience reward for completing a mission

distinguishing between two phenomena. One is any specific instance of a specific game. In a concrete situation players may not acknowledge the game objectives in any standard sense of the word. Perhaps they want to lose or perhaps they apply an aesthetic perspective as is often the case when games are used for artistic expression. In such cases, trust is usually not interesting because players are not dependent on the behaviour of each other. What interests us here is the gaming context in a more general sense. It is the features of game which are general to all or most instances of the game. Thus, it is interesting that *Age of Kings* is (often) played online and that *Counter-Strike* players can choose to play on servers which demand that players use the anti-cheating device Punkbuster (of which more below). But it is not interesting that on a specific day, a specific person played a specific game of *Age of Kings* in which he concentrated on building an aesthetically pleasing cityscape rather than on beating the opposition.

Now, let's look at the gaming context of our six specimens revisiting some of the examples introduced above. As mentioned

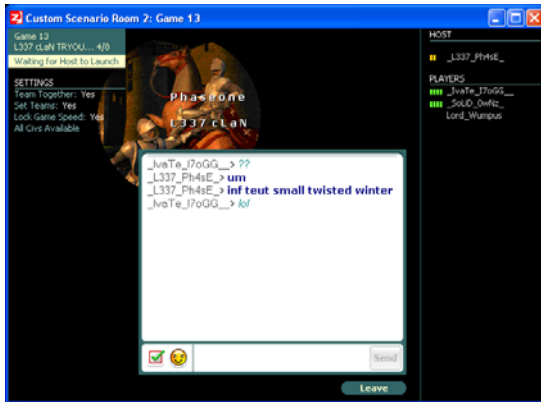


Figure 8 – Age of Kings chat screen

earlier, *Moon Patrol* players are not influenced by each others' actions in terms of the game world. But other aspects of the game arguably depend on trust. First of all, there's money involved. It is conceivable that, if the players did not have exact change, one player would pay for the first two-player game expecting the other player to pay for the next. Also, one player may distract the other while playing. When playing *Moon Patrol* does not in fact feel like an exercise in trust it is due to other features of the gaming context. For instance, the stakes are extremely low. The investment in a single game in terms of time and money are negligible. Also, one player does not really place himself at the mercy of the other. The players are in the same physical space; one cannot suddenly log off and disappear without a trace.

The same features make it much less likely that *Gauntlet* players end up in a situation of prolonged mutual selfishness (the result of a standard, one-shot Prisoner's Dilemma). The *Gauntlet* players can communicate and

face a continuous series of choices and the knowledge that they have to interact in some sense after one choice is made. In the phrasing of political scientist Robert Axelrod, the *Gauntlet* players (since the interaction is continuous) operate under the heavily trust-inducing "shadow of the future" (Axelrod, 1984).

In online gaming, players must largely do without such features. When playing *Age of Kings* online against opponents not in the same location, players face a series of potentially worrisome issues. Some of these are consequences of the platform. Since players have quite easy access to the game files and since their PCs can simultaneously

run other applications, *Age of Kings* players are able to cheat. Importantly, this is a problem in itself even if nobody should choose to boost his or her chances in unfair ways. From a perspective of trust, the very possibility that someone might cheat is of course detrimental to the social fabric of the game space.

Other issues spring from the concrete architecture of the game space. For instance, since player matching to some degree relies on communication through various chat windows, players are capable of making promises that they don't keep. For instance, a player may invite others to play on a custom-made scenario claiming that it is just as fair as the built-in maps. If fact, of course, it may not be (J. H. Smith, 2004)<sup>11</sup>.

<sup>11</sup> For a thorough analysis of trust issues in the *Age of Kings* gamespace see my article *The Games Economists Play* (J. H. Smith, in review).

The promises made by chatting players may or may not be trustworthy. Consider an analogy. In *Star Wars Galaxies* you may wish to transfer some items between two characters associated with your account. To do this you need help from another player (unless you have multiple game accounts). This player must hold the items until you return with your other character to reclaim them. But how do you know that he or she won't simply vanish taking your hard-earned items with him or her? Or let's reverse the perspective. *You* are that other player. Noble as you are, you wish only to help someone to make an item exchange. But how can you make that someone trust you? How can you make a credible promise?

### How we trust

Now we're approaching the question of how we trust or how we make others trust us. The slightly unsatisfactory answer is that we do this in a variety of different ways. All of these ways, however, involve signals. When considering whether to put ourselves (or our money, our children, our reputation) in the hands of others we examine the signals that the other party sends. Neither part of this process needs to be conscious. Our consideration needs not involve high-level cognitive functions (it needs not be an abstract deliberation) and signals do not have to be planned and transmitted on purpose. In fact the process does not have to involve thinking at all.

This point has been driven home quite forcefully. To see this, however, we need to turn briefly to birds. In evolutionary biology the male peacock's tail long proved something of an enigma. Simply put, it

seemed at odds with the general principle of natural selection that this impressive phenomenon had evolved. It did not, after all, confer any obvious survival advantages to its carrier who, lugging around the large tail, in fact seemed disadvantaged in terms of escaping predators. Israeli evolutionary biologist Amotz Zahavi, in the mid-1970s, suggested that the peacock's tail was a solution to a trust problem. The tail was a signal, aimed at peacock females, that its carrier was high quality mating material. Only the genetically well-endowed would be able to incur the cost imposed by the large tail. Zahavi called the more general phenomenon "the handicap principle" stating that in essence it is the theory that "*the reliability of communication (or advertisement)*



Figure 9 – Final chance to change settings before an Age of Kings game starts

*is increased in relation to the investment in the advertisement*" (Zahavi, 1977: 603).

The handicap principle solves the problem that someone told to trust someone else may be sceptical. Arguably, the male peacock could "tell" the female that he was a suitable father to her future offspring. But all males could say this, whether in fact suitable or not. The tail is a statement

which carries its own proof. It cannot be faked. In Zahavi’s own terms it is an *assessment signal* as opposed to a *conventional signal* which is the mere statement (without proof) that something is true (see also Donath, 1999; J. M. Smith & Harper, 1995).

The handicap principle is more pervasive in human life that may be immediately apparent. American sociologist Thorstein Veblen in 1899 described the practice of *conspicuous consumption*; rich Americans advertising their fortunate financial position in a way that could not (easily) be faked – by spending large amounts of money (Veblen, 1899/2000). Merely telling someone “I’m rich” would be a conventional signal.

The principle also lies at the heart of Thomas Hobbes’ analysis of how human beings could trust each other (Hobbes, 1651/1997). They could do so only by giving up their autonomy (incurring a costly handicap) in a way that could not be faked.

Finally, Google uses the peacock solution. While previous search engines had trusted websites to signal their own relevance (or at least what they were about) Google’s page-ranking method instead judged a website’s importance by the number and “weight” of incoming links. Having other websites, preferably important ones, link to you is seen by Google as an assessment signal.

Before we examine how this principle applies to computer games we can express it more formally by again turning to economic game theory. Although less famous than the Prisoner’s Dilemma,

another two-person non-zero sum game is rightly well-known. The Chicken Game is named after the dangerous game of “Chicken” in which two drivers speed towards each other. The first to swerve loses the game, although if neither swerves collective disaster ensues (dying is worse than merely losing the game). Collective defeat (both players swerving) is not as bad as single defeat (see Kollock, 1998).

|          |           | Player 2                     |                              |
|----------|-----------|------------------------------|------------------------------|
|          |           | Cooperate                    | Defect                       |
| Player 1 | Cooperate | Player 1 = 2<br>Player 2 = 2 | Player 1 = 1<br>Player 2 = 3 |
|          | Defect    | Player 1 = 3<br>Player 2 = 1 | Player 1 = 0<br>Player 2 = 0 |

Numbers indicate points earned. Cooperate in the driving game means swerving while defect means going straight.

Rather famously, it has been suggested that a truly effective strategy for a Chicken player would be to throw his steering wheel out the window for the other player to see. This loss of autonomy (this handicap) makes the player unable to swerve thus putting weight behind the threat that he will keep going straight. In the logic of the game losing is better than dying and the other player will be compelled to accept defeat by swerving (if Player 1 parts with his steering wheel, Player 2 can only achieve one of the two bottom row outcomes of which the left is better than the right)<sup>12</sup>.

<sup>12</sup> It would spell trouble, however, if the other driver was similarly unable to “chicken” out – for instance, if he or she had fainted from fear. This problem with the handicap logic is parodied in Stanley Kubrick’s movie *Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb*.

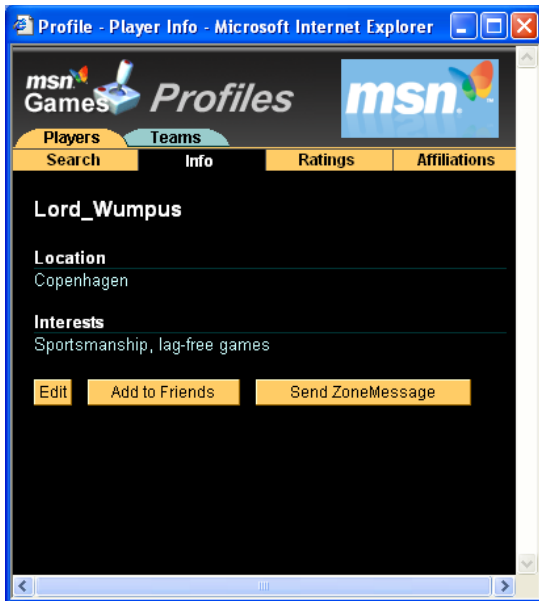


Figure 10 – Opening a player profile on zone.com

Can gamers make such credible promises (or threats)? Before attempting a more general answer let us look at our examples. The *Age of Kings* player may wish to claim one of the following statements to be true:

- I am finishing a download, my ping will improve soon
- My custom scenario is fair for all involved
- I will not attack anyone before the agreed-upon time limit has passed

The sceptical other player may choose to look at the statement maker's profile. Besides the actual ratings, however, the profile contents are chosen (written) by the player himself (see Figure 10). Letting a player write her own profile corresponds to a Chicken player telling the opponent that she will not swerve. It might be true, but then again it might not be.

Much the same is true for our *Star Wars Galaxies* player wanting only to help someone else make an item transfer. Essentially, he can communicate his good intentions using conventional signals based on writing and the limited actions available to his avatar (see Figure 11). Here too, players write their own profiles.

In *Counter-Strike* a player may wish to convince others that he or she does not cheat and will not violate any local norm such as the one against camping (J. H. Smith, 2004). If these promises are made in the form of text (say, on a web forum) they are highly conventional signals without much power to convince a sceptical opponent.



Figure 11 – A Star Wars Galaxies avatar

The inability to make credible promises in these situations can partly be explained with reference to the communication modality. It is the low bandwidth, computer-mediated communication which

grants the sender almost complete control over the communication which renders that very same communication a conventional signal. This said, we should not make the mistake of assuming that text-based communication renders the sender capable of exercising full control of the signals (or cues) being transmitted (Baym, 1996) but in terms of trust we can see *why* high signal control would correlate with low trustworthiness and we can note that it actually *is* the case (Jensen, Farnham, Drucker, & Kollock, 2000).

But of course, this line of argument is mostly relevant in the somewhat contrived situations imagined above. If computer-mediated communication tends towards relative low bandwidth communication (relative to face-to-face communication) and thus towards trust problems, computers also make such problems quite solvable. The reputation management system developed by eBay ([www.ebay.com](http://www.ebay.com)) is a case in point (for a brief discussion of reputation management systems see Zacharia, Moukas, & Maes, 1999). On eBay others essentially write your profile as they express their verdict on your quality as a trading partner. Suddenly, your profile morphs from being a conventional signal to being an assessment signal<sup>13</sup>. In terms of bandwidth the system is extremely crude but the simple feature ensures a remarkable level of trust (further supported by other features of the trading system).

Reputation managers of the eBay variety are rarely seen in online games. This is somewhat puzzling although part of the

explanation is probably the fear that gamers will strategically abuse the system in order to further their in-game goals. Particularly in zero-sum game situations, gamers will have an incentive to unfairly badmouth one another, an incentive not immediately present on eBay<sup>14</sup>.

Bandwidth, however, may not be the most important difference between playing with others online and offline. In all likelihood opponents sharing a physical space, a living room couch for instance, will have interacted before. And we can say for certain that they will interact in the future relative to any concrete in-game decision since they will still be present in the room when this decision has played out. This is different from some types of online play where a player might suddenly disconnect from the game never to be seen again (with the same “nick” at least). In online games, then, you often don’t know if you’ll ever see the interaction partner again. You are not interacting under the long shadow of the future which Axelrod posited as one of the keys to ensuring cooperation in a Prisoner’s Dilemma (see page 11).

Again we are dealing with a difference of degrees. First of all, group size matters. In a gamespace with small player groups or where the game can only be played in a single space (as opposed to different servers) the probability of future interaction rises. A similar effect may be achieved if changing one’s user name incurs a cost on the player. This is clearly

---

<sup>13</sup> Even if not a completely unforgeable one.

---

<sup>14</sup> There are ways to counter this incentive through design (for instance, badmouthing could be made costly to the player doing the rating).

the case in *Star Wars Galaxies* in which the name of an avatar cannot be changed and where an avatar cannot travel between servers. Escaping a bad reputation, then, means abandoning the character along with the investment made in skill improvement etc.

For the *Age of Kings* player, the situation is slightly more complex. Only if a player cares to improve his or her rating within the gamespace – and only if the player has a positive rating – will abandoning a user account mean parting with a costly investment. However, the system may inspire some scepticism towards player with new accounts since this may be a cue that someone is trying to escape a bad reputation. Having a pristine profile may also, of course, be taken as a signal of possible incompetence in which case the player may find locating willing co-players difficult. Thus, even for unrated play, leaving an account behind is not entirely unproblematic.

Lawrence Lessig, in *Code*, makes the related point that switching between two online communities (in some ways comparable to switching between two accounts within the same community) is far more costly than is often assumed (Lessig, 1999: 202)

#### **Player-driven trust mechanisms**

We have seen that certain multiplayer gamespaces do not offer the players specific mechanisms to deal with distrust. We have also seen that some mechanisms, such as continuously rating the players, may in fact affect trust favourable even if that was hardly their main intended

function.<sup>15</sup> But this whole discussion favours the perspective of the game designers. Players themselves are anything but passive when it comes to improving the social fabric of a gamespace by implementing trust-inducing features. Before finally turning to a discussion about the pros and cons of increasing trust in online gamespaces I will consider two player-generated trust mechanisms: *clans* and the application known as *Punkbuster*.

Why do players form clans (guilds, player associations etc.)? This is another question without a snappy one-line answer. Players surely join or form clans for a multitude of reasons, to socialize with likeminded, to achieve a sense of in-game group identity etc. In MMORPGs, players also join clans because of the benefits bestowed on members. Clans may have pools of equipment and other collective resources like guild halls. Also, being a clan member may be the only way to feasibly strive for certain in-game objectives (such as large castles etc.).

But we should not miss the fact that both Thomas Hobbes and Amotz Zahavi would nod in understanding. Despite all the other reasons a player may have for joining an in-game association clan membership makes a player more trustworthy since clan membership is an assessment signal. To join a clan (or remain a member) you need

---

<sup>15</sup> I have argued elsewhere that rated games are a source of *distrust* (J. H. Smith, in review). I believe both statements to be true – playing rated games increases trust as the probability of future interaction rises but it also decreases trust (it makes people more sceptical) since the personal stakes are higher.



to prove yourself worthy. In return you are vouched for by an institution with which other players within the game are likely to have a relationship affected by the long shadow of the future (even if they don't expect to meet *you* again).

Of course, much depends on the actual clan. Not all institutions can provide the same level of backing. In the real world, having a platinum Master Card is different from having an obscure credit card issued by a small local bank. Thus, the trustworthiness gained from joining a clan is a function of the general respect enjoyed by the clan itself. And the larger the benefit, the larger the handicap (or cost) associated with the signal. Jakobsson and Taylor report that in *Everquest*

...reputation plays a significant role in a gamer's success. In über guilds this lesson is doubly important and indeed it might be said that reputation is everything. At a very basic level ones reputation forms an important component in even being admitted into a high level guild. Potential members generally undergo a process in which they petition to join, often listing their equipment and skills. Sponsorship scenarios are common and applicants are often only considered for guild membership after being vouched for by a current member (Jakobsson & Taylor, 2003)

Once one passes the barriers to entry, of course, one still has to respect the restrictions put upon one's autonomy by the group. The *Star Wars Galaxy* player's association (PA) Knights of the Force list the following member rules (among others):

- There is by common sense a code of ethics within KotF, No member shall bad mouth another member or shall be given a demerit,

two demerits shall warrant a vote of dismissal and three demerits warrants automatic dismissal.

- Remember that you are a member of the Knights of the Force and each action you do is a mirror of our PA. With luck and a good group we can become a great PA but each member must be willing to help make this a great PA. (Knights\_of\_the\_Force, 2002)

The clan, in this perspective, is a mechanism which (among other functions) enables players to each surrender personal autonomy in return for mutual trust within the clan and the ability to send an assessment signal of trustworthiness to non-members.

It is true, of course, that in modern games clan systems are not always player-generated. But this is merely an instance of the general trend that new games often include features in which players of older games displayed an interest by constructing themselves.

Another feature arising from players' desire to regulate their games even if it means incurring a personal cost is the anti-cheating application Punkbuster (and similar anti-cheating measures). Punkbuster which today is a commercially produced program, is installed on the client and then continuously monitors the player's machine for signs of cheating applications (see Figure 12). This is Hobbes' Leviathan and Zahavi's handicap principle all in one practical download. Players limit themselves because they are given a guarantee that others limit themselves in the same way. Albeit less all-encompassing this gesture echoes Hobbes' "social contract":

I authorize and give up my right of governing myself, to this man, or to this assembly of men, on this condition, that thou give up thy right to him, and authorize all his actions in like manner.” (Hobbes, 1651/1997: 132)

Only those who limit themselves in a similar way enjoy the fruits of your self-limitation as a Punkbuster-running client will only connect to clients also running Punkbuster.

Punkbuster, in other words, is a mechanism by which players limit their autonomy (they cannot cheat) by incurring a cost (they have to download and install the program) in order to form a subgroup within which members can trust one another. In contrast to the clan system, running Punkbuster is not so much a way of gaining credibility among non-members as it is a way of simply keeping these non-members away.

Like the clan system, however, Punkbuster has increasingly been integrated into the actual game architecture as modern shooters in particular let players filter out game servers which do not require Punkbuster and may even come bundled with the application<sup>16</sup>.

## DISCUSSION ON THE VALUE OF TRUST IN GAMES

This article has focused on the status and nature of trust in multiplayer gaming. I find it evident that many online game spaces display signs of trust problems as players worry greatly about cheating and spend

great amounts of time making sure that others can be trusted to behave constructively within a game. So, should we do our utmost to increase levels of trust within gamespaces?

The brief answer is “no” while the more thoughtful answer is that it depends on what we want to achieve.

First of all, it follows in no way from the discussion above that anyone should want



Figure 12 – The website of Even Balance who produce Punkbuster ([www.evenbalance.com](http://www.evenbalance.com))

to increase trust on the *rule system* level (of course). There is no reason to assume that games would be better, or gamers happier, if chess, *Counter-Strike*, and *Top Spin* were less competitive and more non-zero sum. But even when focusing on the larger gaming context, we should be aware of two important caveats.

First of all, it is often the case that safety (an important side of trust) is in opposition to freedom. There may be pleasure in big city anonymity, perhaps even an advantage

<sup>16</sup> At the time of writing this only 450 out of 2427 (18,5%) *Battlefield 1942* servers do not require the player to use Punkbuster.

to getting lost ending up in unexpected neighbourhoods once in a while. Such pleasures would be compromised by electronic GPS-enabled guidance systems and the ability to look up anyone's name and occupation by pressing a button. Similarly, the safety produced by widespread surveillance of public spaces may conflict with our desire to act without this act being registered.

In games too trust can be bought expensively. *The Dark Age of Camelot* player (on most servers) can trust others to not ever attack her since the game simply does not allow a player to attack another player. Undisputedly, this is a type of trust but it is achieved in a very heavy-handed way.

This leads us to another concern which is tied to the status of gamespaces. In the sense that online games are seen as entertainment and perhaps performance we cannot dismiss that conflict, to a large extent, is of course the essence of drama. A truly utopian game world with complete inter-player trust would probably appeal to very few. Might there not, for instance, be a certain appeal to the inability to fully trust the person assisting to switch objects? Could his betrayal not lead to a deep personal motivation and an epic world-spanning chase for the culprit? Revenge is sweet, after all, and without offence it is hard to conceptualize revenge.

Players who play for rating, tournament players for instance, and players who have other ulterior motivations (like real-life money) are unlikely to share this view, however.

Finally, *some* deviance may surely help strengthen a community. Without deviants

to define one's own group in relation to it is difficult to imagine a community becoming anything but a rather loose association of individuals (although external enemies may provide much the same benefits).

These considerations do not tell us that distrust is good. They merely show us that trust in gamespaces is a complicated affair and that trust can be bought at the expense of other virtues if handled carelessly.

If this description reminds us of something that something may be the real world. Our states, our collective institutions are testament to one possible balance between personal freedom and mutual trust. To the extent that democracy works they are, arguably, a reflection of our aggregate attitude towards the question of which value should be assigned to trust in a world of limited resources. But while it would be game design folly not to seek inspiration in millennia of experimentation with these delicate balances we should also acknowledge that those real life balances are based on assumptions that may not hold in online gaming. While classical political philosophers may be ambitious they do not tanker with the basics, they do not, for instance, propose a rebellion against the laws of nature. In this sense, game designers are like gods when compared to their real life counterparts in the business of social structure building. The makers of multiplayer game spaces, virtual worlds in particular, for all their professed (and actual) helplessness in the face of complex social dynamics are powerful to an extent that would astonish both Glaucon and his famous brother. The consequences of their choices make for

remarkable experiments. Interesting developments surely await us.

## REFERENCES

- Atari. (1985). Gauntlet (Arcade ed.).
- Axelrod, R. (1984). *The Evolution of Co-operation*. London: Penguin Books.
- Baym, N. K. (1996). The Emergence of Community in Computer-Mediated Communication. In S. Jones (Ed.), *Computer-Mediated Communication and Community*. London: Sage Publications.
- Donath, J. S. (1999). Identity and deception in the virtual community. In P. Kollock & M. Smith (Eds.), *Communities in Cyberspace*. New York: Routledge.
- Ensemble\_Studios. (1999). Age of Empires II: The Age of Kings (PC ed.): Microsoft.
- Hobbes, T. (1651/1997). *Leviathan – Or the Matter, Forme and Power of a Commonwealth Ecclesiasticall and Civil*. New York: Touchstone.
- Irem. (1982). Moon Patrol (Arcade ed.).
- Jakobsson, M., & Taylor, T. L. (2003). The Sopranos Meets EverQuest: Social Networking in Massively Multiplayer Online Games. *FineArt Forum*, 17(8).
- Jensen, C., Farnham, S. D., Drucker, S. M., & Kollock, P. (2000). *The effect of communication modality on cooperation in online environments*. Paper presented at the SIGCHI conference on Human factors in computing systems, The Hague.
- Knights\_of\_the\_Force. (2002). Rules (Vol. 2004).
- Kollock, P. (1998). Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology*, 24, 183-214.
- Lessig, L. (1999). *Code and Other Laws of Cyberspace*. New York: Basic Books.
- Machiavelli, N. (1950). *The Prince And the Discourses*. New York: The Modern Library.
- Maxis\_Software\_Inc. (2000). *The Sims*. Version: Electronic Arts Inc.
- Plato. (360 BC). *The Republic*. Online version: <http://classics.mit.edu/Plato/republic.html>.
- Power\_and\_Magic. (2003). Top Spin (Xbox ed.): Microsoft.
- Quervain, D. J. F. D., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., et al. (2004). The Neural Basis of Altruistic Punishment, *Science* (Vol. 305, pp. 1254-1258).
- Smith, J. H. (2004). Playing Dirty - Understanding Conflicts in Multiplayer Games, *5th annual conference of The Association of Internet Researchers*. The University of Sussex.
- Smith, J. H. (in review). The games economists play - implications of economic game theory for the study of computer games, *In review*.
- Smith, J. M., & Harper, D. G. C. (1995). Animal Signals: Models and Terminology, *Journal of Theoretical Biology* (pp. 305-311).
- Sony\_Online\_Entertainment. (2003). *Star Wars Galaxies - An Empire Divided* (PC ed.): LucasArts.
- Valve. (2000). *Half-Life: Counter-Strike* (PC ed.): Sierra On-Line, Inc.
- Veblen, T. (2000). Conspicuous Consumption. In D. B. Holt (Ed.), *The Consumer Society Reader*. New York: The New Press.
- Verant\_Interactive. (1999). *EverQuest*. Version: 989 Studios.
- Vogel, G. (2004). The Evolution of the Golden Rule, *Science* (Vol. 303, pp. 1128-1131).
- Zacharia, G., Moukas, A., & Maes, P. (1999). Collaborative Reputation Mechanisms in Electronic Marketplaces, *Thirty-second Annual Hawaii International Conference on System Sciences*. IEEE Computer Society.
- Zahavi, A. (1977). The cost of Honesty (Further Remarks on the Handicap Principle), *Journal of Theoretical Biology* (Vol. 67, pp. 603-605).