Volatility in Game Design: Lecture Notes from Gen Con 2012

By James Ernest, Sept. 25 2012

Description: *This article is a discussion the use of randomness in game design. It will focus on tabletop games, such as board game and dice games. These are essentially the lecture notes from my talk about Volatility at Gen Con 2012.*

Background: I wrote this document before Gen Con and added some material a few weeks after. This builds on some of the work of Richard Garfield and Dave Howell in game design analysis.

Introduction:

High volatility in games can level the playing field and widen the range of potential opponents. In games with high luck and high skill, any player can win in the short term, but more skilled players will win more in the long term. To design a game that uses luck correctly, you must balance the interactions of biased and fair random events.

Randomness is an aspect of game theory that most game designers understand poorly. This talk is not about the mathematics of probability and statistics, which you can research elsewhere. Rather, it's about the application of those principles to create a good game experience.

Assumption: We are talking primarily about multiplayer tabletop games.

What is Luck?

Three basic elements affect a player's chances of winning a game. I'll call them **Aptitude** (the player's intrinsic ability or skill at the game); **Play** (the player's actions and decisions within the game) and **Luck** (factors that are beyond the player's control, including random events).

Luck vs. Skill: Games are often described on a single axis with "luck" at one end and "skill" at the other. This is an oversimplification of three independent variables. "Luck" and "Skill" can both be very high or very low in a game, and "Skill" is really a loosely defined conflation of "Aptitude" and "Play."

In reality, games can have a little or a lot of all three aspects. Tic-Tac-Toe requires very little Aptitude, very little Play, and no Luck. Poker has high Aptitude and high Luck, while still requiring very little Play. Games like Warhammer 40k and Magic: The Gathering are high in all three.

Using randomness correctly in strategy game design requires *adding luck without removing play*. The degree to which you are allowed to let luck overwhelm play is inversely proportional to the amount of play in the game. That is to say, in extremely simple games (slot machines, for example) the work is low enough that the game can be entirely luck. This is why most adults can't play Candy Land.

Adding Randomness without Removing Play, or "I Hate Chess."

I hate Chess because, these days, it's all about who knows the most opening moves, and not about who's actually better at playing Chess.

I'm paraphrasing, but who said that, besides me? Bobby Fischer. He's talking about the distinction between Aptitude (what you bring to the game, in this case memorized moves) and Play (what you do in this game, in this case making strategic choices, or "playing" chess).

Aptitude trumps Play in a nonrandom game that's too well understood. Blackjack is another example of such a game, where aptitude (memorizing a chart, or counting cards) trumps play (making decisions on the fly) for nearly all players, but at least it's more exciting than chess because the best move does not always produce the best result.

What was Fischer's solution? Something hobby designers know well: randomize the setup. If you've figured out one set of opening moves, rearrange the pieces and start again! Fischer worked hard, and against the chess establishment, to promote "Chess 960," a random way to arrange the starting pieces. The pieces are the same, black mirroring white, for a game which is equally fair, but which takes memorization out of the picture. This randomized arrangement shifts the burden of victory from Aptitude to Play, without raising the amount of Play required.

Purposes of Randomness:

Random events can make games more engaging, through variability and surprise, and they can broaden the field of potential opponents, by reducing the impact of Aptitude. Casual players respond well to random events and enjoy the variable outcomes of random games, even though they sometimes lose. However, hardcore players are often frustrated by games in which their better decisions do not lead to success.

Unfortunately, many "game designers" fall into the "hardcore" category, and they deal poorly with random events, if they use them at all. You can tell when a hardcore designer is forced to use randomness, because he inevitably uses it terribly, and then covers for it by adding extra anti-volatility rules. (Catchup features, for example.)

Usually when a game has the right mix of luck and skill, it appeals to the broadest range of possible players: hardcore players have something that will give them an edge, and casual players don't feel like their chances are too low. Poker is the best widespread example of this, though popular hobby games also fit the mold: Settlers of Catan, Ticket to Ride, Munchkin.

As a designer, if you want your game to appeal to a broader audience, you need to let go of any personal bias you might have towards a more hardcore framework, and build a game that is more generally appealing.

Yet many designers fear randomness, implement it poorly when they do use it, and are always providing avenues to avoid purely random results. This is based on an unrealistic presumption on the part of the designer that all players desire a high degree of flexibility and control.

Applied Randomness:

Randomness in game design can be broken down into three general categories relating to game design: I call them Cosmetic, Biased, and Fair.

- **"Cosmetic" randomness** has no bearing on the strategy or outcome of the game. An example would be a change in background color, or the difference (in poker) between a heart and a spade.
- **"Biased"** randomness gives resources unfairly to one player over another; this includes most random events in most games. It's the easiest type of randomness to design, but it has pitfalls.
- **"Fair"** randomness is the hardest to employ. It challenges players to think strategically in a randomized environment, but does not favor one player over another. For example, allocating different but equally valuable resources based on a die roll, or randomizing the starting setup in a way that is the same for all players.

It is possible to have random elements in a game that fall somewhere in between biased and fair, but in those cases, different aspects of each mechanic can usually still be described through this filter.

Some Examples:

Cosmetic: We don't talk much about Cosmetic randomness, but its purpose is to provide variability without any impact on the game mechanics. It's easier and more prevalent in computer games. Cosmetic randomness might describe the different suits in Poker (each is functionally equivalent) or different artwork for the same Magic card (Land cards, for example).

Biased: This is the easiest form of randomness to introduce to your game design, and the most common. Here are some examples.

- A slot machine spin. Some spins are good, some are bad. This is purely biased randomness, when judged by the dollar value of each result. Different reel symbols can have the same value, which could be described as a cosmetic element.

- A dice roll in Settlers, when it gives resources to some players, and nothing to other players. However, this mechanic might be considered partially “fair” when it gives different resources to different players, if those resources are of equivalent value.

- A movement roll in Monopoly. This can land the player on a good or bad spot, strictly based on the dice and the state of the board.

- Something-or-Nothing rolls, in games too common to mention.

Fair: This type of random event is hard to create. Fair randomness has a meaningful impact on the flow of the game without specifically favoring one player over another. A strictly fair random event favors no one, while slightly less fair events might only favor a certain player by a small margin. Some examples:

- A random setup in Chess (Chess 960). Each side is a mirror image of the other.

- A random starting layout in Settlers of Catan.

- This-or-That rolls (get Resource A or Resource B) in a context where both resources are equally useful but to different strategies.

Grit: Grit takes its name from Dwarven Dig, but it’s a design concept that I use a lot. When you fail a roll in Dwarven Dig, you don’t get nothing. You get “Grit,” which is a resource that you can trade in for better rolls, more power, and so on. It’s a simple way to change a “this-or-nothing” roll into a “this-or-that” roll.

Hybrid Biased-Fair: This is any random event that falls between a purely biased and purely fair random event. A lot of game randomness can be argued to fall into this category, neither purely biased nor purely fair, such as the Settlers resource roll described above.

- Grit could be seen as a middling type, since presumably you don’t want the grit as much as you want whatever you were rolling for. But it really depends on the game; sometimes the consolation prize is just as good as the prize.

- The shuffle in poker. While this seems like a purely random distribution of cards, poker allows for a player option that many games don’t have: folding. Thus, there are correct decisions with bad cards, making this random distribution slightly more than just a purely biased event. Be careful that you do not try to build Poker without folding, because sticking someone with a bad hand is the epitome of biased randomness.

Collectible Card Games and Fair Randomness: One of the “fair” mechanics to which collectible card games aspire is the individual value of each card. When you open a booster pack and receive a Land, a Goblin, and a Lord of the Pit (or any other assortment of common, uncommon, and rare cards), each of those resources has a “fair” value when compared to the others from extremely far away.

This doesn’t mean they are all worth the same money, or do the same damage in combat, or anything that specific. What’s meant by “fair” is that the common card may be weaker, but is more useful to many different strategies, while the uncommon card is more useful but in a more limited way, and the rare card is extremely useful but only in a very specific situation.

You can think of resources in your own game as having this kind of distributed utility, and use a booster-pack mentality when randomly distributing these resources to the players. Someone who is prepared for a Lord of the Pit can make good use of it, but everyone else will find it useless.

The Crazy Train:

One strong way to use randomness in game design is to make random choices optional. Games often give players a choice between a slow and steady path, and a highly volatile one. In the best games, neither of these strategies is dominant. We call choosing the volatile path "hopping on the Crazy Train."

A simple Crazy Train mechanic is in Deadwood, a game about acting on a Western backlot. When working on a movie scene, players have a choice to play it safe (rehearse) or take a chance (act). If you rehearse, you earn a permanent +1 to subsequent die rolls on that scene. If you act, you roll the die. Other game elements notwithstanding, each strategy has roughly the same success rate; but players can always choose between conservative and aggressive strategies.

There is a comparable mechanic in Lords of Vegas. When you own a lot, you can choose to develop it, or leave it empty. When empty, it earns a dollar every turn. If you build, it has a roughly 1/5 chance of paying a lot more, roughly $4 and a point.

In both Deadwood and Lords of Vegas, Players can jump on the crazy train when they want, but they are not penalized for playing it safe (for a while). With multiple lots, of course, you can balance your risk between both strategies.

Summary of Main Topic:

You should be using randomness in strategy games to keep the game interesting and variable, without unfairly favoring one player over another. Biased randomness can make casual games more chaotic and family-friendly, but it ought to be used sparingly if at all in games where players want to think.

Further Discussion:

The Gambler's Fallacy:

Biased Randomness can't be fixed with more biased randomness.

All game designers are the victim of the gambler's fallacy at some point. Even when they know what it is. In brief, it's the belief that long series of random numbers tend to even out, so if you have just flipped "heads" ten times, you are now more likely to flip "tails." When you say it like that, it's easy to refute, but designers frequently add more dice rolls and more card draws with the intent of balancing out purely biased random events. It doesn't work.

In a dice-driven horse race where each player will roll a 6-sided die 50 times, suppose the results after turn 1 are 1 versus 6. This early in the game, with 49 rolls to go, you would hope that the game is not already tilted heavily in one player's favor. But in fact, the player who is in the lead is already a 66% favorite to win, and it only gets worse as you move deeper in.

On the other hand, when the random events are "fair," adding more events gives players more choices, which means more chances to make good and bad moves. Fair randomness can really only exist in the context of strategic decisions, where the die roll happens before the choice. Adding more of these events will certainly give players a chance to catch up from bad situations.

Roll Before You Think:

Lots of the best Fair Random events are random setups. That is, it's better to roll the dice before you think, rather than thinking before you roll the dice. Here's why.

Dice Chess is an example of a bad way to add biased randomness to a game. Actually, two ways: a bad way, and a worse way.

Let's play Dice Chess, version one. Here's how it works. At the beginning of the game, we will roll a 6-sided die. On a 1, I win. On a 6, you win. On any other result, we play chess to see who wins.

While this is a horrible game, it serves to illustrate that if you put the randomizer before the strategy, you can at least prevent me from feeling like I've wasted effort in playing the game.

To make the example more clear, let's play Dice Chess version two. In this game, we play chess first, and THEN roll the die.

(Thanks to Robert Gutschera and Richard Garfield for the example of Dice Chess.)

Catchup Features: Why they are terrible, and what you should do about them.

A "catchup feature" is any rule that intentionally favors players who are losing, or hurts players who are winning. Also called a "headwind," if it hurts the winner, or a "tailwind" if it helps the loser. For example, the taxation in Settlers, punishing players who have generated a lot of resources.

Unless they are baked into the fundamental structure of the game, Catchup features are usually terrible. They cannot, by definition, solve the problem they address. And they are indications of a fundamental flaw in the core design.

Why do I say that catchup features can't work? Dave Howell's Theorem of Fun states that a game is only fun as long as every player thinks he has a reasonable chance to win. But "reasonable" doesn't mean "equal". Someone who is behind should not feel as though he has the same exact chance to win as someone who is ahead: it invalidates all the play that has happened so far.

If your catchup feature is so powerful that it gives all players an equal chance to win throughout the game, then it violates Howell's theorem and makes the play irrelevant. On the other hand, if it is not so powerful that it violates Howell's Theorem, then it can't sufficiently help players catch up; it can only make them lose by less. This makes players feel better, perhaps, but it's not a catchup feature.

So how do you design a game without a catchup feature? Stop allocating resources through biased randomness, and then players won't arbitrarily fall behind. Players should fall back through bad decisions, not bad rolls. (Again, this can be mitigated by extremely short play times, or frequent hard resets.)

Bonus Feature: "Kill the Elf" (A digression into probability)

Let's consider a simple dice rolling game. Two players are competing to "kill the elf." One of them is a warrior, who does one point of damage every turn. The other is a Wizard, who does 6 points of damage, but only on a roll of 6.

On average, both Wizard and Warrior do 1 point of damage on each turn, so they both should kill the elf on turn 6. However, despite the fact that the Wizard's average kill time is 6 turns, he will actually kill the elf by turn 5 nearly 60% of the time, and he is still 1/6 to tie on turn 6 if he does not win outright. The odds of rolling 5 dice without rolling a 6 are only 40.1% ((5/6)^3). All in all, despite apparent equivalence, the Wizard has 63.2% of this match.

How would you balance this? One way to do this is to give the elf 4 hit points. In this race, the Wizard wins outright 42.1% of the time (odds of rolling at least one 6 on 3 dice), and ties 1/6 of the time on turn 4, for a total slice of 46.9% of the game (the Warrior has the other 53.1%). That's pretty close.

The game would be even closer if you gave the Elf 10 hit points, in which case the Wizard has to hit the elf twice in 10 rolls (approximately 51.5%). Still not a perfect match, but getting there.

In fact, if you gave the Elf a thousand hit points, the game would get very close to fair. Now the Wizard's average damage of 1 point per turn is relevant again, because his weapon isn't an instant kill as it is in the first example.

This is an interesting case that seems to support the Gambler's Fallacy, which is that expected results normalize over a larger series of random events. I'm just beginning to analyze proto-games like this, and it's producing some interesting results. When I've done more, I'll write up another set of notes.