Intrinsic Rewards

···IN GAMES AND LEARNING···

KEVIN MIKLASZ, PHD

Intrinsic Rewards in Games and Learning

Intrinsic Rewards in Games and Learning

KEVIN MIKLASZ, PHD

CARNEGIE MELLON UNIVERSITY: ETC PRESS
PITTSBURGH, PA

ISBN: 978-1-71651-070-0 (Print)
ISBN: 978-1-716-18624-0 (Digital)

Contents

This book was inspired by a frustration. I have had my feet dipped in the world of education for over a decade now. I've also had my feet in the world of games, as a game designer for several years and as a player almost my entire life. As I began my career in education, I noticed myself learning when I played games. What I am learning can be up for debate—whether it's a bunch of fantasy world knowledge that has no application to real world work outside the game, or whether it's the exact kind of problem solving skills that are most needed in today's world. It doesn't matter where you fall on that spectrum, you have to concede that those playing a game are learning something.

At the same time that games were a place where I *always* learned something, school was a place where I *sometimes* learned something. And that was the source of my frustration. Why was school, a place specifically designed for learning, less effective for learning than a game, something specifically designed for entertainment?

At some point I discovered the wonderful works of James Gee, which played a large part in alleviating my frustration. He so clearly explains how and why games teach, and offers a set of principles that schools can emulate. Most importantly, he argues that it is not that more games should be used in schools, but that schools should take more effort to mimic the principles that are found in games.

Then educational badges came along. This movement reached the peak of its hype curve somewhere around 2013 and 2014, but is still going strong in some places today. I noticed that badges seemed clearly inspired by achievement systems in games, but differed from them in several key ways. And, most efforts at badging seemed to fall short of being effective while at the same time, achievement systems in games were making great strides to further the learning someone was getting from those games. I felt a disconnect, and noticed others identifying this discrepancy too.

At some point in this time period, I also discovered "Punished by Rewards" by Alfie Kohn, which is a fantastic book. Two big takeaways I had from this book were: 1) rewards have existed in education for a long time and 2) such rewards are mainly harmful for learning (that's a gross simplification of his book, and I highly encourage you to read it if interested). Kohn's recommendation is to eliminate rewards from

education as much as possible. Although I couldn't fault his logic, the conclusion he reached bothered me: games have rewards, and games foster all kinds of learning, so maybe it's not that rewards themselves are always harmful for learning, but it's something about how they are used that's important.

Maybe if we just had a way of understanding what a good reward system is, then this could all come together. We could understand why the rewards in games worked so well. We could understand why the rewards that Kohn rails against worked so poorly. We could understand why some educational badge systems worked well, and others worked poorly. Maybe we just needed a framework to understand the problem. And since I have a borderline compulsive obsession to try to put things into categories and dissect the subcomponents of systems, I ended up with the book you have before you.

The first half of the book is really a dissection of rewards. What are different kinds of rewards, what dimensions should we care about, how have these dimensions been used throughout history, and throughout the game industry? The second half of the book is much more about application. How can we use rewards effectively, how do we design them meaningfully, and how can we apply ideas from the first half of the book to education?

Whether you come to this book as an educator, school administrator, designer, developer, researcher, or just someone curious about these issues, I hope that you will find something useful to guide your daily work and to help us all make school a better learning experience for our students.

1. Do Intrinsic Rewards Exist?

Alfie Kohn notes the simple structure of a reward: "Do this to get that."[1] In this phrase, "this" refers to the action and "that" refers to the reward. We'll call this structure an extrinsic reward—by definition when you do the action, you are getting a reward that is extrinsic, or unrelated, to the action itself.

So what is an intrinsic reward? Intrinsic rewards have the structure "Do this, to do this more." In this case, the reward for the action is being able to do more of the action. The reward is intrinsic to the action—the reward further values the action itself. This is, by definition, always different from an extrinsic reward.

This probably seems a little bizarre. Who would ever do an action just so they can do the action more? What kind of reward is that? Well, I'd argue this is the structure that governs a lot of what we do in our everyday lives.

Let's say you have a hobby, like knitting. Why do you knit? Sure, knitting has a useful purpose—you can make socks, sweaters and scarves. You produce objects by knitting, and perhaps your motivation for knitting is simply to acquire more of those objects, in this case, socks. But then, time might be better spent earning money at a job just to buy tons of socks (which not many people do). No, it's something about the process of making a sock that makes knitting an engaging hobby.

Consider this. People knit socks to get better at knitting socks, which allows them to knit socks even faster and learn new stitches, which allows them to get even better at knitting socks. The process of learning how to knit compels someone to keep knitting, since developing that skill is actually an intrinsic reward. By practicing a hobby, you get better at the hobby, which allows you to practice the hobby even more, which allows you to get even better at the hobby. By this logic many hobbyists are driven by intrinsic rewards.

How do we more generally determine if a reward is intrinsic or extrinsic? I like to recommend the "nagging three-year-old test." Take

any given action you do in your day, and start by asking yourself "Why am I doing this activity?" And then every time you give an answer, ask "Why?" If over the course of asking why you find yourself repeating an answer and going in a loop then you have an intrinsic reward. If on the other hand you find yourself only able to answer the why with "just because," then you've probably got yourself an extrinsic reward.

Let's take another example, from gaming. Let's go to one of my classic games on this topic, Final Fantasy,[2] which has a structure typical of many video games. In Final Fantasy, you spend a lot of time fighting monsters. So, let's apply the nagging three-year-old test.

- **Why fight monsters?** To get experience points.
- **Why get experience points?** To go up levels.
- **Why go up levels?** To get stronger.
- **Why get stronger?** To fight harder monsters.
- **Why fight harder monsters?** To get more experience points.
- Etc.

We very quickly started looping back to our original action. So the reason you fight monsters is to fight more, and more difficult, monsters. In this case, fighting monsters is an action with intrinsic rewards.

Let's take a non-game example. We can ask the question, why do I take a picture of my family?

- **Why take a picture of my family?** So that I can have the picture on my work desk.
- **Why have it on my desk?** So I can be reminded of them while I work.
- **Why be reminded?** Because it makes me happy.
- **Why be happy?** Just because.

We very quickly reached a "just because" statement, making it clear that in this case, the act of taking pictures had an extrinsic reward.

But it also didn't necessarily need to be an extrinsic reward. If photography was a hobby for me, and I wanted to specialize in taking pictures of people, then taking a picture of my family could have been a way to get better at taking pictures, which would have helped me take

better pictures of my family. In other words, I might have been taking a picture of my family due to intrinsic rewards. How I perceived the action in my mind, or what that action means to me, affected whether that action had an intrinsic or extrinsic reward to me in that situation.

Visualizing intrinsic reward structures

Based on this discussion, we can now offer the basic visualization of an intrinsic reward system. This basic system has four steps. First, doing an action **produces** an immediate reward. The immediate rewards **build into** a cumulative reward. The cumulative reward **results** in some kind of effect. And that effect **modifies** your ability to do later actions (generally, it makes you better at them, allowing you to practice and accomplish the action at a more difficult level).

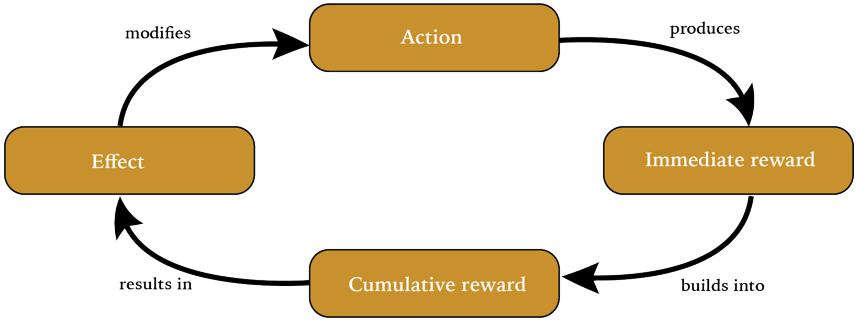

Image 1.1: The basic intrinsic reward loop.

This basic structure can be modified in many different ways, and much of this book is dedicated to going through the various forms and varieties that intrinsic rewards can take. But one element must hold true to constitute an intrinsic reward structure—the action has to loop back into itself. Steps can be removed or added, but the circular nature of the diagram has to be preserved. The reward for doing the action has to result in doing the action more.

We can also slightly extend this diagram to include extrinsic rewards.

In this case, there are a myriad of potential avenues to value a reward for extrinsic reasons, but I only show a couple here. You can think of each arrow as a different way of answering the nagging three-year-old's "why" question. Which diagram applies to a given individual in a given situation is dependent on the individual and the situation.

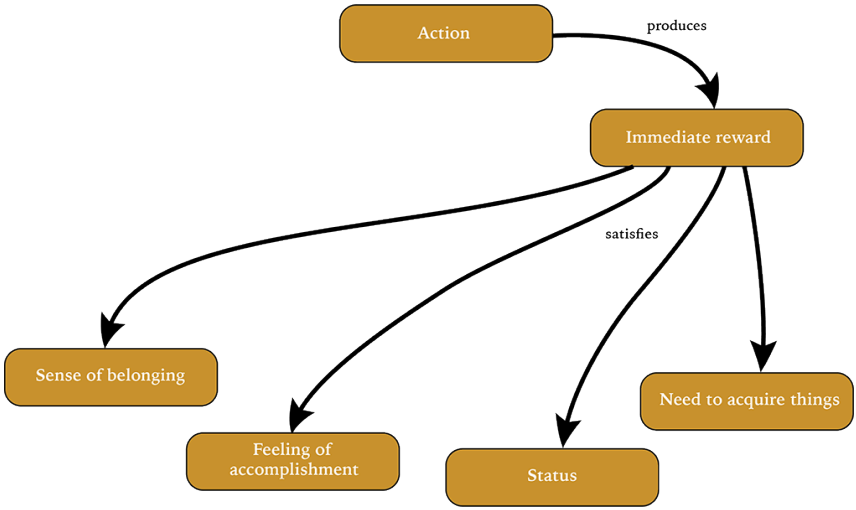

Image 1.2: The basic extrinsic reward structure.

The value of rewards

I want to be abundantly clear that intrinsic rewards are not always good and extrinsic rewards are not always bad. Whether a reward is intrinsic or extrinsic refers to its structure, not to its value. For example, taking pictures of my family to put them on my desk seemed like a very valuable kind of activity and one we wouldn't lightly dismiss, despite being a clear example of an extrinsic reward.

In fact, readers with some background in the social sciences might note a parallel between intrinsic rewards and Self-Determination Theory, or the psychology theory that describes how we are intrinsically motivated, or self-determined, to do certain actions. Although there are clear parallels, intrinsic rewards are not defined as rewards that foster

intrinsic motivation, they are defined by whether or not they loop back to the original action. This is a definition based on my background as a designer—the structure or design of the reward system defines the term "intrinsic reward". Though of course on some level if you are repeatedly doing an action just for the sake of improving your performance of the action, you are probably intrinsically motivated to do that action to some degree, so these concepts are likely strongly related.

I will use the three basic psychological needs of Self-Determination Theory to define whether rewards (intrinsic or extrinsic) are meaningful or not. But that deserves a much fuller discussion. I go into this in depth in Chapter 7. For now, I do want to note one thing—the way we think about meaningfulness in rewards should depend on whether the reward is intrinsic or extrinsic. For extrinsic rewards, we have to consider whether the reward itself is meaningful. Whereas for intrinsic rewards, the action itself is the ultimate reward, so we instead need to consider whether the action is meaningful.

One last thought for those of you coming to this book wondering why badge and achievement systems seem to work so well in games and so poorly in education. I have a proposal. The difference is that in games, the rewards are often well-designed, meaningful, intrinsic rewards whereas in education the rewards are often poorly-designed, meaningless, extrinsic rewards. The rest of this book is a slow climb towards proving that statement true, and discussing how we can make educational reward systems more meaningful and effective.

2. A History of Rewards

To start our exploration, we'll begin with the historical underpinnings of rewards. Rewards have gone by many different names in different places throughout history. Some of the different permutations are badges, points, stars, levels, ranks, merit badges, achievements, +1's, likes, upvotes, medals, grades, and microcredentials. It's really not important to understand the nuances between every specific kind of reward, especially because some rewards are especially good at blending into other kinds of rewards and it's unclear where the true distinctions between the different names lie. What is important is to understand that all of these things fit into the same kind of bucket, and that it is a big bucket with many dimensions. In this way, I use the term "reward" in an extremely inclusive manner to indicate any kind of goal-setting, tracking or advancement layer added to an experience or activity.

We often think of rewards in their most prominent contemporary uses: in the badges movement, in social media, and in games. I will spend much time talking about rewards in those modern systems, but it's worth noting that all three are relatively recent realizations of rewards, going back to the 1980's. Rewards writ large go back much further in human history. I want to explore in depth a few select examples of historical reward systems, while examining if their structure is intrinsic or extrinsic.

Rewards in the Military

The military is arguably the area of society which has been using rewards and reward structures the longest. Rewards typically take two forms in the military—military ranks and awards.

Military ranks have a long history, with evidence of use by both the ancient Greeks and ancient Romans. They are primarily used as a way to institute order and chain of command in the military, and have a highly functional purpose. But of course, there must be some criteria by

which individuals are awarded higher ranks, and so the award of a rank is typically also a marking of military skill. As such, military rank can be cited as a signifier of authority.

The system of military ranks greatly differs both throughout history, and in different countries. However, these differences are not significant for the purposes of this book. Broadly speaking, as you go up ranks, you gain command of larger and larger numbers of units. For example, in the British military, a military section might be composed of 8 soldiers or so. A platoon is made up of several sections. A company is composed of several platoons. A battalion is formed of several companies. A brigade is formed of several battalions. And so on. Each hierarchical grouping is headed by a military leader, starting with corporal, sergeant, lieutenant, major, etc. In this way, there is a clear progression by which one can advance up ranks, with clearly designated title changes, to control larger and larger military units.

Ranks could potentially be either an intrinsic reward or extrinsic reward, depending on the soldier's motivation. Some may want to go up ranks to be able to engage more fully in military operations, and thus challenge their own ability to be a better and better commander. For others, higher ranking may simply be an external signifier of their own skill or a desire to gain greater authority and prestige, and thus be a clear signifier of extrinsic reward.

In addition to military ranks, there are also a number of military awards given out to individuals for exceptional service. These are clear examples of extrinsic rewards. The earliest recorded use of such awards dates to Egypt in 2000 BC.[1] These awards can be quite varied in their use and criteria. They also differ in their frequency—some are quite rare, some more commonly given out. For our purposes, I've focused on one specific military award to provide a concrete example.

The US Army describes the Purple Heart Medal as the oldest military badge that is still given out in the United States. It's earlier form went under the name "Badge of Military Merit" and was given out by President George Washington after the revolutionary war. It was not used again until after World War 1, when it was revived under the name Purple Heart. The first Purple Heart was given to General MacArthur in 1932. It was originally awarded for both meritorious service, and for receiving wounds from an enemy. In 1942, a "Legion of Merit" award was introduced for meritorious service, and the Purple Heart was

amended to only apply to receiving wounds. After many additional amendments, the award now has its present form, in which it can be given to any military member in active duty who receives a wound. It can be given out in the field, and can be given posthumously.

Rewards in the Scouts

The Boy Scout (and the Girl Scout) badges are one of the more famous implementations of rewards that still persists to today. The Boy Scouts of America (BSA) started in 1910,[2] and it's clear that even from the start rewards were a key component of the program, with the first 57 merit badges and 6 badges of rank appearing in the 1911 Boy Scouts Manual.[3] There are at least three kinds of rewards acknowledged in the Scouts BSA program: awards, merit badges, and badges of ranks.[4] The awards themselves are the least interesting of the three for our purposes, as they most resemble the types of awards given in other contexts—mainly recognizing positive attitudes and attributes exhibited by a person or group over a long period of time with a clear extrinsic reward structure. But both the merit badges and the badges of rank are interesting and intertwined systems of rewards that are worth investigating more deeply. I focus specifically on the Scout BSA requirements from Scout through Eagle rank, though it's worth noting that interesting variations on this structure have been introduced for younger students (Cub Scout), older students (Varsity Scouts) and specialized content areas (Sea Scouts).

134 merit badges exist as of 2020, each of which is structured around a specific skill.[5] Each merit badge has a list of requirements that must be met for the badge to be earned. These requirements need to be demonstrated in front of a Merit Badge Counselor (who has to meet certain requirements to be a counselor). Once a Scout can demonstrate they have met the requirements, the counselor will sign off on their blue card application form, and they will have earned a badge. The process itself seems relatively straightforward, similar to applying for a driver's license, for instance.

Yet there is clearly more to the process than simply being recognized

for a skill. The opening section of the BSA guide to merit badges explains:

> "There is more to merit badges than simply providing opportunities to learn skills. There is more to them than an introduction to lifetime hobbies, or the inspiration to pursue a career—though these invaluable results occur regularly. It all begins with a Scout's initial interest and effort in a merit badge subject, followed by a discussion with the unit leader or designated assistant, continues through meetings with a counselor, and culminates in advancement and recognition. It is an uncomplicated process that gives a Scout the confidence achieved through overcoming obstacles. Social skills improve. Self-reliance develops. Examples are set and followed. And fields of study and interest are explored beyond the limits of the school classroom."[6]

This passage makes clear references to the ideas of extrinsic and intrinsic rewards, if not using those words directly. It notes that clear extrinsic "results" will occur, such as introductions to hobbies or career paths. Yet it also indicates that something more happens, and that it is really about the process (or the action) rather than the result (or the rewards). This shifts the emphasis of earning a badge towards an intrinsic structure, where the value is placed on the activity (earning the badge), not the resulting reward itself. And the various benefits listed in the second half of the paragraph result from the activity, rather than the reward.

The rank system is even more interesting. Boy Scouts advance through ranks by gaining certain combinations of merit badges. The requirements for each rank consist of a certain number of "required" and "elective" merit badges the scout must attain. Additionally, once the scout believes they have met the requirements of a rank, they meet with a "board of review" to undergo a final pass. Once again, the simple mechanics describe something that sounds remarkably like a school structure, with a "curriculum" of required and elective courses laid out.

Yet the description that leads off the "Mechanics of Advancement: In Boy Scouting and Varsity Scouting" section tells a much more nuanced

story.[7] For example, see this description about the purpose of advancement:

> "It is important to remember that in the end, a badge recognizes the Scout has gone through an experience of learning something he did not previously know. As a result, through increased confidence, he discovers or realizes he is able to learn similar skills or disciplines. Advancement is thus about what a young man is now able to learn and to do, and how he has grown."

This again shows a strong emphasis on the action itself, and the process rather than the result. The section goes on to make this point even more explicitly:

> "Advancement, thus, is not so much a reward for what has been done. It is, instead, more about the journey: As a Scout advances, he is measured and he grows in confidence and self-reliance, and he builds upon his skills and abilities."

The journey is the process, and the process here seems much more of an intrinsic reward loop of learning skills by doing.

With their ranks and merit badges, BSA is aiming to create and recognize intrinsic reward loops without having a specific name or terminology for that structure. BSA sets up an environment where skills are practiced, tested and refined, resulting in increased confidence in one's own abilities. An intrinsic reward structure clearly has to be at play if this language can be used to describe what is going on. It seems as though merit badges aren't meant to be an extrinsic reward in the form of physical patches, but rather a way to recognize that a scout has cycled through an intrinsic reward loop a sufficient number of times to have developed a skill. How interesting that this iconic structure that is often pointed to as the basis for many of today's reward systems has had a semblance of an intrinsic reward loop built in from the very start.

Rewards in Education

Formalized educational systems go far back in human history, potentially as far back as some of the first public schools founded in ancient Greece. The earliest origin of what became our modern schools can be traced back potentially as far as the 5th century, with universities emerging in the 12th century. These earliest universities created the foundation of the modern diploma system. But the idea of grades and the minute tracking that form the majority of how rewards are used in today's school can be traced to sometime in the nineteenth century.

Many reward structures in education focus on the microscopic elements of education—what happens in a classroom. Letter grades themselves act as a sort of reward, whose purpose is to allow a student to pass their grade level in school. In a sense, there could be a bit of an intrinsic reward loop in this system—get good letter grades, to pass your grade, to get harder assignments. Unfortunately the loop breaks down here, as the system doesn't work this way in practice. Many students will pass a grade level regardless of the grade they get, and will therefore get harder assignments the next year as mandated by the curriculum. In many schools, the grades students receive don't determine the difficulty of the assignments they receive after that work. This then leaves the rewards (grades) with an external value of judgment on how smart you are as a student.

Diplomas themselves are also a reward that becomes formalized into a progression of levels. For example, one gains a BA or BS in college, which can then be followed by one of several other degrees including MS, MBA, Ph.D., etc. These degrees act as a reward that unlocks further degrees that allow more learning to occur, in a sort of loop. Of course, the loop does end with your final degree, the purpose of which is often "to get a job." So, ultimately, education seems to have at its root an extrinsic structure in its reward system of diplomas. It's worth noting that intrinsic and extrinsic rewards are not fully binary—the number of loops you go through affects how intrinsic or extrinsic the reward is.

Rewards in the Food Industry

We've also implemented rewards in the food industry, most notably as Michelin stars. Although the first Michelin Guide was published in 1900, it focused exclusively on tires. Michelin Guides ranking restaurants date to sometime after World War 1, eventually developing a system of inspectors that visited restaurants and awarded stars on a one-to-three star scale.[8] Those stars can be removed or added to a restaurant every year based on a new visit. The Michelin committee has a fixed number of stars in circulation in every country, in effect limiting supply of this reward, which is an interesting choice as it implicitly puts every restaurant in competition with every other restaurant. This seems a clear extrinsic reward.

The food grading system is a more recent form of rewards in restaurant work. The little signs in front of restaurants in most of the U.S. record the safety grade of the restaurant, often going on a traditional school scale from A-F. Unlike the Michelin star, this reward is not limited in quantity—there are strict criteria, but if everyone passes the criteria (and they should do so), then they get an "A." This seems a clear extrinsic reward.

There are also more local forms of the Michelin system. Often local governments or magazines will declare a local restaurant the best restaurant of the year in some food category. Sometimes only one restaurant can win each year, sometimes several can win. Unlike the Michelin system, there is little consistency in how this system acts in each place. It seems more a recognition of which restaurants have a particular value to a certain community.

Finally, there are consumer-based rewards, which have mainly become prominent since the age of the internet. Of course word-of-mouth recommendations for restaurants have existed for quite some time, but new technology-based solutions have formalized those recommendations into something that very much feels like a reward system. Take Yelp, where each restaurant has between a 1-5 star rating, based on the average score given by sometimes thousands of reviewers. Along with the stars, or a quantitative reward, often comes a written review, or qualitative reward. This system of quantitative rating from 1-5 with a qualitative review is a reward structure that has

proliferated throughout our lives in many different industries, for example in movie reviews. This in general seems relatively extrinsic, but there is a potential intrinsic element. Because the main action of a restaurant is to serve customers, and customers are the ones giving the reward, it can be argued that a restaurant serves customers to get good ratings, to be seen by more people, to generate more customers, to get more good ratings, and so on. And since it is more difficult to manage a busier restaurant, there is an inherent element of difficulty built into the reward loop too.

Rewards in Games

Although we'll focus on rewards in the context of modern digital games throughout the book, I want to note ways even some of the earliest games, both digital and analog, have incorporated rewards.

Go is one of the earliest board games in human history, dating back to at least 500 B.C. in China, and having spread to Korea and Japan by the 5th and 7th century AD. From a rewards standpoint, Go is particularly interesting. Modern Go has a player ranking with over 40 levels. Players start at a rank of 30 kyu when they are a beginner, and as they improve, their rank goes down in number until they reach 1 kyu. Once they advance past the 1 kyu rank, they become a 1 dan, all the way up to 7 dan. Beyond 7 dan, players are considered professional Go players and have a rank that goes from 1p to 9p. The best players in the world are ranked 9p. This system was formalized in Japan in the early 17th century, and is the basis behind many martial arts ranking systems. A 1 dan rank is considered equivalent to achieving a black belt. The dan and p system is formally regulated and can be an acknowledged rank that is achieved through tournament play. The origins of this system date to at least the 2nd century AD in China, in which a 9 rank system, called the 9 Pin Zhi, was described as being used to rank players.[9]

What is most interesting about the classic Go ranking system (beyond its extremely old age) is that it also has meaning within the game. Traditionally, to promote fair play, lower ranked players should be given a handicap which is in proportion to how many ranks they

are below another player. So for example if a 2 dan player is playing a 5 dan player, there is a 3 rank difference between them so the 2 dan player should begin the game with a 3 stone handicap. The system is designed so that with the handicap, players should be playing an even game, or each player should have a 50% chance of winning. Additionally, ranks are considered equivalent to about a 10-point difference in final score. So, if our 2 dan and 5 dan players were to play without a handicap, we would expect the 5 dan player to win by about 30 points. The maximum allowed handicap is 9 stones, so players who are more than 9 ranks apart are not able to play a fair game against each other. This also illustrates how the ranking system is subjectively based upon the rank of other players—you don't know your rank until you play someone and see how much you win or lose by. The ranking system also affects the culture of the game. Since every game can easily be made into a fair match by using a handicap, it is expected that, in casual play at least, every game should be a fair match that challenges each player to the best of their abilities.

All of this points towards a clear intrinsic reward structure—play Go to improve your rank to play more challenging players and play Go better. Even amateur tournaments employ handicaps in games, and are considered opportunities to best determine your rank relative to others instead of elimination style competitions with a winner. Of course, there are also plenty of tournaments, especially at the professional level, that have a clear winner, such as the Meijin tournament in Japan which has roots in the 17th century.

In contrast to Go, many modern games use some form of the Elo points system, named after the Hungarian mathematician who developed the system for Chess in the early 1900s. In this system, players are given a rating that is based on whether they win or lose against opponents. Without getting into the math, the general logic is quite simple: if you consistently defeat players who have a higher ranking than you, then your rank goes up. The extent to which your rating changes depends on your wins, losses and draws, and the Elo ratings of your opponents. Like the Go ranking system, it is implicitly balanced relative to other players, but unlike Go it results in a very quantifiable number which has little effect on how the game is played. This disconnect with the game mechanics can cause the system to feel much more like an extrinsic reward. It's worth noting that the

United States Chess Federation has broken down the Elo rating into various classes of performance,[10] which resemble the system for Go in several ways. The Elo system has also been incorporated as is or with modifications in the competitive scene of many games like Scrabble,[11] and has also been used in several professional sports rating systems by organizations such as the FIFA (International Federation of Association Football) and the Intercollegiate Tennis Association. Elo-based systems—or derivatives of Elo, like the Glicko system—form the basis of many digital game matchmaking and competitive systems, such as Counter Strike: Global Offensive[12] or Overwatch.[13]

Basic point and leveling systems were available from the earliest digital games, like Pacman.[14] Launched in 1980, Pacman had two notable reward systems. First, points were calculated based on the quality of a player's play: riskier maneuvers (i.e. going for multiple ghosts after consuming a pellet) could result in higher scores per level. Additionally, getting further in the game also produced higher scores. The points were an extrinsic reward. Second, the levels themselves also acted as a reward structure, and many players would remember which level they got to, or what "fruit" they last achieved, more so than their final point total. The levels seem much more of an intrinsic reward—defeat levels, to play even harder levels. As this game has gained popularity, additional fan "achievements" have been recognized, such as getting a perfect score of 3, 333, 360 by getting all possible points on all levels up to level 255, or reaching the "kill screen" or level at which the game breaks at level 256.

The history of rewards is rich and varied, even before the recent explosion of reward design in the video game industry. Now that we've explored the history of rewards, let's attempt to categorize the design of rewards.

3. A Taxonomy of Rewards

At this point, we've noted that rewards can be separated by their intrinsic or extrinsic properties, but that is far from the only dimension we can use to categorize reward structures. In this section, I'll walk through 9 additional dimensions by which rewards can be categorized. These are not mutually exclusive categories, but can offer a useful view into the diversity of reward structures that exist. This is not by any means an exhaustive list, but I chose to limit this to dimensions which seem most useful for comparing video games to educational systems, and to comparing intrinsic and extrinsic reward structures.

If the problem with educational rewards is that they are poorly designed extrinsic rewards, then the obvious solution is to add intrinsic rewards to education. But even if we want intrinsic rewards in education, we still don't know quite what to do: there are many ways to implement intrinsic rewards. And also, there's still a place for well-designed extrinsic rewards in education, though again there are many variations on what a good design can be. This taxonomy describes those variations, and paves the way for making recommendations for the types of intrinsic rewards that should be used in different educational settings.

The chart below gives a summary of the dimensions. Many of these are borrowed from work Lucas Blair conducted in his dissertation on achievement systems in games.[1] In particular, the following dimensions were primarily inspired by Blair's work: completion vs. measurement, expected vs. unexpected, and temporary vs. permanent. Though it's worth mentioning that even the dimensions that were not directly taken from Blair's dissertation were still strongly inspired or influenced by him—for instance, his discussion of rewards as currency greatly influenced the internal rewards dimension. The dimensions are arranged from top to bottom in rough order of importance for consideration in educational design work, and the dimensions will be discussed in that order.

Image 3.1: A full map of all 10 dimensions, with a short description of each one. Arrows indicate dimensions that are a spectrum, and rewards can fall anywhere on the spectrum. Dots indicate dimensions that have discrete, mutually exclusive categories into which rewards can fall.

Discrete vs. leveled

This is one of the most fundamental distinctions between different kinds of rewards. Discrete rewards are typically a one-off, binary kind of reward. You either have this reward or you don't. The classic Xbox Live achievements are a prime example of discrete badges. Your stereotypical badge, like the Boy Scouts merit badge, is a discrete reward. The Purple Heart Medal given by the military for valor is also a discrete reward. You can often accumulate multiple discrete rewards,

as in the experience points systems in games or the “Likes” system on Facebook.

As the discrete rewards decrease in value, the amount of rewards you gain often becomes more important than the fact that you have gained a reward. At some point, this can then become formalized into a leveling system. Leveled rewards are not binary—they are rewards that can be gained to various degrees, or levels. Many RPGs do this—characters may gain individual experience points for defeating monsters, but acquiring certain amounts of experience points will increase your level. In this sense the points are still a discrete reward, but experience levels are a leveled reward.

There are also leveled rewards that don’t depend as directly on accumulated discrete rewards. For example, military ranks are a leveled reward that has no basis in discrete awards. Same with the Go leveling system. The Boy Scouts has a rank system too, which depends on collecting enough different kinds of merit badges. In this way, a Boy Scout rank is a leveled reward that is defined by a collection of discrete awards, but the variety of the collection matters rather than the pure amount in the collection.

These instances point to an additional sub-category between discrete and leveled. Discrete rewards can be single rewards, or they can be accumulated rewards. The idea of accumulated discrete rewards is sort of an intermediate between discrete and leveled rewards, from a design perspective. A player can naturally create “levels” of effort in their mind when the accumulated rewards pass certain values (i.e. passing 1,000 points, 2,000 points, 3,000 points, etc.), but it’s not quite the same as when a designer of a system explicitly acknowledges levels as a reward.

Measurement vs. completion

This dimension essentially describes whether the reward tells someone how well they did at a task (measurement-based rewards), or whether or not they simply completed a task (completion-based rewards). Although a simple distinction, the effect it has on the reward is

profound, and this distinction affects and interacts with most other reward dimensions.

Measurement rewards often contain some kind of numerical scale that accompanies the reward—a measurement of the player's ability. Two classic examples of a measurement reward structure are grades given to students on homework assignments and players getting between 1-3 stars for completing a level of Angry Birds.[2] Lucas Blair notes that achievements in games usually measure players against one of three criteria: "performance can be measured against another player's performance, their own performance, or some standard set by game designers."[3] In assessment terminology, there's a parallel idea of norm-referenced assessments (assessments are measured against other students such as "your score is in the 60th percentile") and criterion assessments (assessment scores are measured against a learning standard set by education experts such as "your score is proficient for your grade level"). Each of these measurements can affect the reward's value in a game or learning experience. A ranking on an online leaderboard is a leveled, measurement-based reward set by measuring one's performance against others. A grade on an essay in a class is a discrete reward measured against standards set by a teacher (the equivalent of a game designer in educational settings).

Completion-based rewards simply note that a task was completed or level of expertise was gained. For example, you can get a reward for the task of completing a dungeon in World of Warcraft,[4] or you can get the reward of unlocking the highest difficulty level of a game by beating the game on the lower difficulty (e.g. the "Expert" difficulty for each song on the rhythm game osu!stream[5] is unlocked by getting a score of "A" on that song on the "Normal" difficulty). Completion-based rewards can also be given for simply showing up, like giving a student a star for being in attendance. In some cases, completion-based rewards are given an additional distinguisher of being "performance-contingent" and "non-performance contingent", basically indicating whether or not you have to demonstrate performance to get the reward.

The two extremes of this dimension often blend into each other in practice. In the osu!stream game mentioned above, a measurement-based reward (a performance rating that goes from A to F) is given out every time that a song is completed, regardless of performance, and that reward contains information about performance in the form

of a numerical score. On the other hand, once an "A" rating is gained on the normal difficulty for a song, the expert mode for that song is permanently unlocked. This performance-contingent completion-based reward (an "A" rating on Normal that unlocks Expert mode) is not given out after every song attempt, but is only given out once, when a certain level of performance is reached. In this way, performance-contingent completion metrics will often "measure" performance against a threshold. By this logic, both forms of rewards (measurement-based rewards and completion-based rewards) are making a measurement: the difference is merely a matter of frequency and permanence of the measurement. Do you want a detailed, frequent and temporary measurement of performance (measurement-based reward), or a rough, infrequent, but permanent measurement of performance (performance-contingent completion-based reward).

Let's say osu!stream has a reward for playing songs 50 times (this game doesn't, but many games do have this kind of non-performance contingent reward). In this case, the reward contains no information about performance, it simply values effort. This would be an example of a completion-based reward that doesn't even have a rough measure of performance.

We can think of this as a spectrum. On one end sits measurement-based rewards measured against others' performances, and which feels the most judgmental and subjective but gives a lot of info about how you rank relative to others. Next come measurement-based rewards which are measured against either your own performance or standards set by a designer, which still contain a fair bit of information and judgment. Next are completion-based rewards which are performance contingent, which now have less info, but still contain some info because a benchmark of performance was passed, even if the performance itself was not directly measured. Finally, at the other end of the spectrum sits the non-performance contingent rewards, which contain little info about performance and no judgment.

In my opinion, none of these options is unambiguously the "best" reward to choose in any given situation. Completion and Measurement based rewards have different strengths, and you should carefully choose an option that best matches your objectives for a game or learning experience.

Internal and External Rewards

This distinction refers to whether rewards are completely external to the experience, or whether the reward itself has some internal value in the experience. Lucas Blair refers to a similar concept in whether achievements have "in-game currency," which is one way to describe a realization of an internal reward. External rewards sometimes have "out-of-game currency," but more often have no tangible value.

Internal rewards are most classically represented in an experience point and leveling system. The experience points are a reward for playing the game well, which in turn cause you to go up levels in the game and affect your character's skill and performance. Another, more recent example, is Jetpack Joyride,[6] where certain actions in the game allow you to gain achievements, which give you stars, which you can use to buy powerups. In a similar style, League of Legends[7] grants you experience points after every game, which directly add to your in game currency and let you buy new heroes and powerups.

But internal rewards can exist without an explicit currency system. Powerups themselves are a sort of reward in many games that directly increase your character's power in the game and therefore have a clear internal value—and these powerups can exist without having to invoke a currency system. A classic example is Megaman 2,[8] in which defeating a boss gives you a new attack that mimics that boss's power and makes it easier to defeat successive bosses.

External rewards can be seen in the classic achievements on Xbox Live, or the points you gain in a classic arcade game. They are an indicator of how or how well you are playing the game, but do not affect the game itself. They may have values in the social structures that surround the play (i.e. you may want to brag about them to friends) but they do not impact the gameplay itself.

A grey area occurs when it is unclear where the boundary of the game stops, like in pervasive games or games with a strong social community. For example, there might be a reward entirely external to the gameplay itself, that gives you no direct benefit when you play the game. But gaining that reward might increase social status in the games community, which might cause certain game guilds to accept you, which may in turn affect how you experience the game

as a part of that guild. Another grey area is microcredentials. Although gaining the credential at the end of a MOOC is not something that affects your ability to learn in the MOOC, it can have a sort of academic currency and give you credit at your university. Whether this is an internal value, because you are at the university to learn and this is a learning experience, or whether it is an external value, because taking the MOOC is something separate from the learning at the university, is a bit unclear.

There are positive and negative aspects to both kinds of rewards. For external rewards, the benefit is that the reward can act as a record of your performance in the game. Reflection on past performance can be aided when there is a visible manifestation of that performance, especially as a permanent, measurement-based reward. The potential detriment of external rewards is that they have no direct tie in to the activity itself. Especially when these rewards have an out-of-game currency value, these rewards can destroy intrinsic interest or motivation in the game or learning experience and act as an extrinsic motivator.

For internal rewards, the benefit is the direct tie-in and relation to gameplay. Rewards can reinforce positive actions in the game and enable useful goal-setting when those rewards are expected. Additionally, Blair notes that these rewards can create greater player agency in a game, by allowing players to choose where and how to spend their purchases and level up their characters. We'll come back to agency in spending rewards in Chapter 5. The main negative point is the same for external currency system—when the internal currency system is very strong, as Blair notes "players will end up caring about the reward system more than the game itself."[9] This reduces intrinsic motivation and creativity and can create a single-minded focus on the task at hand and the internal currency it results in.

Internal and External rewards have a strong parallel to Intrinsic and Extrinsic rewards, but they are different concepts. It's often much easier to create a reward loop when that reward has an in-game value or in-game currency, and in many ways an intrinsic reward is almost guaranteed to have some kind of internal value. But, intrinsic reward loops can also have external value that's worth recognizing and discussing. Also extrinsic rewards can have either internal value (once you get a one-off achievement, you gain in-game currency), or external

value. As a designer, it's helpful to consider whether you want your reward to have internal value to the experience or external value separate from the experience, as a separate design consideration from whether or not the reward forms a looping structure. Thus, this dimension is worth breaking out separate from intrinsic and extrinsic rewards, despite the clear parallel and correlations between the two.

Hard vs. soft unlocking structure

One main use of internal rewards is to tie them into an unlocking system. Players can "cash in" internal rewards in order to unlock advanced features in the experience. These kinds of unlocking systems have two general varieties: hard-unlocking and soft-unlocking. Hard-unlocking structures typically have discrete levels, and once a level is gained, a discrete element in the experience is unlocked. There may be an area of the map you just can't access until you reach level five, or there's a difficult activity you can't do until you complete three other activities. In an education game there could be a tutorial that you are not allowed to access until you complete a previous tutorial.

In contrast, in soft-unlocking nothing actually restricts you from accessing higher content except your own skill. You may be free to access a certain part of the map, but you will probably lose if you go there until you reach about level five. Or you can do a difficult activity at any time, but you are unlikely to be successful at it until you gain experience with three other easier activities first. Or you can watch a tutorial at any time, but are unlikely to understand it until you have watched previous tutorials.

Hard unlocking structures tend to feel more forced and unnatural, but can also direct focus and create well-ordered problem solving. Typically, soft-unlocking structures are considered more elegant, but they are also more difficult to design. Soft-unlocks allow users to continually challenge themselves, more readily leading to the optimal experience of flow.[10] Soft unlocks also allow more choice and can be less frustrating for an advanced player, who is free to skip ahead to later content if wanted. I also think soft unlocks give a greater feeling of accomplishment—being able to do something at any point in time

means that the only thing stopping you from achieving that reward is your own ability. Once you gain it, you can really feel your own progress. In a hard unlock, it is less clear that your own ability led to your success, since you weren't able to access the challenge initially.

It's worth noting that not every reward has an unlocking structure, and that there is a third category of "no unlocking structure present." Typically, extrinsic rewards do not have an unlocking structure, whereas intrinsic rewards often, but not always, do. Since the main point of an unlocking structure is to enable further actions and rewards, this structure makes the most sense to apply when rewards loop into further rewards.

Qualitative vs. quantitative evaluation

There must be some kind of criteria that determines when a reward is given. In today's era of constant gamification, we almost take it for granted that rewards are pretty much always quantitative, based on some score or measured criteria. But this isn't always the case, especially as we consider rewards in education. Grades are a great example of a reward system that is sometimes based on quantitative measures (how many math problems you got right) and sometimes based on qualitative measures (how well you argued a point in an essay). Many badges for learning, or micro-credentials, operate on a similar criteria—they are awarded by an expert, after they evaluate some evidence that you have uploaded. In many cases, this can mostly be broken down into whether a machine or algorithm determines your reward (quantitative) or whether the subjective judgment of a person determines your reward (qualitative).

Outside of education, qualitative rewards are also common in social media. Likes are based on a friend subjectively deciding that whatever you posted is worthy of their notice and recognition, in the form of a like or a "+1".

This dimension interacts with measurement vs. completion rewards. Completion rewards, whether quantitative or qualitative, always feel less judgmental than a measurement based reward, which evaluates how much or how well something was done. Qualitative measurement

rewards are potentially the most harmful to the learner or player, as in that case a person is "judging" the worth of your contribution. Grades in school are a good exemplar of a qualitative, measurement-based reward, and Alfie Kohn describes the problem with this kind of reward very clearly. In fact, he particularly recommends that if you want to use grades at all, you should "limit the number of gradation" or even better "reduce the number of possible grades to 2: A and Incomplete."[11] Or to put it another way, if you are going to use qualitative grades in a classroom, you should make it a completion-based system.

A quantitative measurement-based reward can still potentially run into problems, and is not always non-judgmental. The fact that the calculation of the measurement is automatic does help though—it can give a clear and transparent criteria that you need to strive to meet. It can then help focus the feedback and attention on how close you came to achieving the criteria (i.e. your task) rather than your innate ability. In fact, the language around the task is important,[12] as well as the extent to which the reward is leveled vs. discrete and permanent vs. temporary. Discrete temporary measurement-based rewards tend to focus on a particular instance, whereas leveled permanent measurement-based rewards tend to focus on multiple instances summed into an evaluation of your ability.

For an example of a non-judgmental quantitative reward, let's take the game Dance Dance Revolution.[13] In this game, after each song you get a grade (from F to AAA, a discrete measurement reward), which is determined by your score (an accumulated discrete reward), which is determined by how well you hit certain notes. It is quite clear from the context that this is a temporary evaluation of your performance on this particular song at this particular time, and not an evaluation of your general "dance, dance, revolving" ability. The verbal feedback from the game reinforces that with phrases like "Oh, you didn't make it!" implying that maybe you will next time. The reward also feels relatively discrete in nature—about this particular song, on this particular playthrough. It doesn't necessarily feel permanent—on your next playthrough, you will get a different reward.

For an example of a judgmental quantitative reward, let's take the ranking system behind League of Legends competitive ladder (which is generally similar to the competitive scene of many competitor multiplayer video games). In this system, you play multiple matches of

the game. Whether you win or lose causes you to gain or lose LP, or League Points (an accumulated discrete award). When you get certain amounts of LP, you can go up to higher divisions (a leveled reward), from Iron to Challenger, which is the professional rank of esports players. The division you belong to is a quantitative measurement-based leveled reward that attempts to sum up many past performances to make an evaluation of your worth as a player. And this evaluation can be measured on an online leaderboard against other players, giving the reward suddenly a bit of a qualitative nature too. In other words, these rewards very much are intended to judge how good of a player you are (in this case on both an absolute scale and relative to other players). Players in the game know this, and will mention their rank to note how good they are, or will make derisive comments towards players that have lower ranks, further reinforcing the judgmental nature of this reward.

Human vs. Machine Evaluators

In general, there are two different ways that the criteria for awarding a reward can be evaluated: either by another person, or by an automated system. Although this reflects a simple fundamental point that seems obvious, it has some important implications.

This category has some strong parallels with the last category. In general, machines are really only capable of giving quantitative rewards, or rewards based on some objectively defined criteria. This can of course get metaphysical quite quickly, as one could argue that AI has the ability to apply subjective judgments, especially as we move towards machine-grading of essay writing, which is a bit controversial. But then another might argue that AI is nothing more than a more complicated set of rules that are followed, and that even though they are not obvious or determinable by a person, they are still a quantifiable set of rules. At which point one may counter-argue that is no different to how the human brain works, and that subjective judgment is nothing more than a not-obvious set of rules by which human brains operate, etc. I'm going to just sidestep this argument and say that we don't live in the world of the uninterpretable AI (yet), and so any

machine-defined set of criteria for the most part can be called quantitative.

In contrast, humans are capable of either qualitative or quantitative systems of evaluation. We can use subjective or objective criteria to evaluate whether something is achieved. As an example, take Olympic figure skating—the "technical score" is a quantitative score based on (mainly) objective criteria that are evaluated by humans, whereas the "performance score" is more explicitly described as a subjective criteria left up to the judgment of humans.[14] One way to think of this issue is: either a human or a machine can measure how far someone jumped, but only a human can measure how beautiful the jump was.

The main tradeoff is that machine evaluations can feel more fair but less social and connected to others. Human evaluations can offer a greater social connection between the evaluator and receiver of the reward, and can go into greater detail than machines in some settings. But humans are less efficient and effective than machines and their evaluations have a greater potential to feel judgmental or unfair.

Expected vs. Unexpected

Lucas Blair introduces this dimension with a simple example—"When a player earns an achievement, the notification they receive can come as a total surprise or as the finish line they were striving for."[15] He further notes that expected rewards enable goal setting by the player, whereas unexpected rewards can encourage creativity and experimentation. When using expected rewards, it is important to make sure the reward is clearly presented to the player, and its criteria is clearly defined.

The default case for most games and reward systems is to have "expected rewards"—unexpected rewards are often the rare case. Blair suggests this balance is ideal. It is interesting to note however, some games can push in the opposite direction quite effectively. Alchemy[16] is an example of a gameplay in which every reward is "unexpected" in the sense that the game does not tell you what rewards (or elements) that you can create. But it's also "expected" in the sense that elements are found through combining combinations of other elements that fit our intuition—we may not know that water and earth will create mud

when combined, but we are not surprised when they do so. It's also interesting to note that in Alchemy, the unexpected rewards are also a key component of the gameplay itself, acting as "game currency" or an internal reward—it might be that tying unexpected rewards into an expected component of the game experience allows them to be used with greater frequency in the game.

A closer look shows us there is much grey area between these extremes in games too. For example, in Achievement Unlocked,[17] a list of 300 achievements that must be unlocked sits on the right hand side of the screen. In this sense the rewards are expected, but the rewards are not given a clear description. The name of the rewards is almost a "hint" to help you find a reward that you know must exist, and is often found through exploration and experimentation. Many adventure games employ this tactic, by advertising achievements that can be achieved, but purposefully not clearly specifying how they must be achieved. Or as another example, take the case of energy tanks in the game Metroid Prime.[18] On the one hand, when you discover a hidden passageway in a tunnel and find an energy tank at the other end, it is a surprising, unexpected reward. On the other hand, if you have played other games in the Metroid series, you know that there will be energy tanks that increase your life force hidden somewhere in the game, so you can expect to find some energy tanks over the course of playing if you creatively explore your landscape.

Temporary vs. Permanent

For the most part, we tend to think of rewards in the permanent category—why reward someone if you don't let them keep the reward? And in general this statement is true, the majority of rewards are permanent and for good reason. As Lucas Blair notes, these rewards allow post-play reflection, which can increase gameplay retention and learning. But temporary rewards also have value in certain contexts.

The typical temporary reward is the instantaneous feedback that occurs in many games as players execute good maneuvers. They may get an instantaneous message like "Nice shot!" or "Unstopppable" or "Big Combo" which fades after a time and has no permanent record.

One way to think about these messages is as rewards that have low permanence. A player is given a reward for a certain state they achieved at a certain time, but once they leave that state (e.g. once they lose the combo), the reward disappears. A good example is the "dominating" reward (a single discrete, completion-based award) in Team Fortress 2[19]—you can dominate an opposing player by killing them 3 times in a single play session without them killing you once. That reward remains on the board until the dominated opponent finds a way to kill you (thus breaking your streak), or the game session ends. Another example is the combo marker (an accumulated discrete, measurement-based award) that appears on the screen when you hit 5 or more notes in a row in Dance Dance Revolution: it indicates you are currently hitting every consecutive note in the song and it disappears when you miss a note. It's a reward that acknowledges a state that is in effect right now and might be taken away at any moment—it's an extremely direct and instantaneous form of feedback.

Some of the competitive ladder systems in the gaming world implement a variant of temporary rewards—you may have a rank in the system based on how much you have won. But if you don't play often, that rank will automatically fall over time. This system encourages the idea that being the best in the game at one time does not mean you are the best always—other players can improve their skill in the game, and to prove you are in fact better than them, you have to continually defeat them and prove your skill. Thus being in a high rank only means you are best at that moment in time. There is a parallel here with professional sports—winning the World Series acknowledges that a certain Major League Baseball team is the best team right now, or that year. But the next season, their record is reset to 0-0, and even though they may have the exact same starting lineup, they aren't the best team that year until they win the World Series again. Winning the World Series is a temporary reward, though with a longer permanence compared to the examples above.

Temporary rewards only make sense for a certain category of rewards that acknowledge that a specific state is present. But, these rewards can offer a particularly effective, immediate, and non-disruptive feedback that such a state is currently in effect.

Abundant vs. Scarce

Rewards can also differ in their scarcity. A simple way of thinking about this dimension is, is it possible for everyone to receive the reward? If the answer is yes, the reward is abundant. If the answer is no, the reward is scarce. The fewer people that can achieve the reward, the more scarce it is.

Abundant rewards are often the norm in game-based achievement systems. In many digital games, typically anyone who plays is capable of getting every achievement in that game. Even if only a few people actually gain a particular achievement in practice, maybe because that achievement is especially difficult, it is still theoretically possible for every person to gain that achievement, if they put in enough time and effort. This distinguisher is not talking about rewards which are scarcely achieved because they are difficult—it's referring to rewards that are scarce by design.

It's also worth noting that abundant rewards are also common in education. For most assignments, it is possible, and desirable for the teacher, for every student to get every question on an exam right, so that every student will gain the maximum grade.

There are a variety of ways that a reward can be scarce. One common way is by multiplayer competition. In most competitive multiplayer game systems that implement an ELO type point system, the rewards are intentionally scarce. The winner gains points, and the loser loses points. It is, by design, not possible for all players to gain points when the game concludes. Although the last example describes how accumulated discrete rewards can be scarce, discrete and leveled awards can also be scarce. Another possibility is to make a reward scarce just by limiting the number of people who receive it. For example, the Most Valuable Player trophy in many professional sports leagues is scarce because only one person is allowed to achieve it each year. Same with the Nobel Prize. Or with a class valedictorian. Or really, any elimination-style tournament that has only one "first place" reward to give out.

But there are ways to sit between "one and only one winner", and "every single person can get every reward." For example, the Michelin star system is a clear example of a reward with scarcity, but still some

abundance. Although it is true that not every restaurant in the world can get a Michelin star, it's also true that at any given time more than one restaurant will have Michelin stars. Masters and PhD degrees are another example. If we consider gaining a post-graduate diploma as a reward, it is true that anyone in the program can gain that reward, but only a limited number of people are accepted to the program in the first place every year.

Another way scarcity is achieved is by giving rewards by a relative ranking to peers (or a subjective, quantitative reward). In some colleges, grades are given out on a bell curve, meaning your grade on an assignment is determined by how well you scored compared to your peers. In this case, it is impossible for everyone to gain the highest value of the reward, because some people will always be on the lower end of the bell curve.

One last example of scarcity comes from game shows. In some cases, it may be possible for anyone to gain a reward, but the reward is only actually given out to the person that answers first (for example, Jeopardy). This also creates scarcity, although this is usually embedded in a format where there are multiple attempts to gain points by answering first—so although any particular reward has a high degree of scarcity, the system as a whole allows for rewards to go to multiple individuals and has a more moderate level of scarcity.

Other Distinguishers

These dimensions describe the main elements of the taxonomy. Below, I'd like to offer a few other taxonomy elements that have been offered by others as a way to think about rewards. These elements do not feel distinct enough to be included as their own element of the taxonomy, but still feel interesting and important enough to be discussed in some detail.

Local vs. Global Rewards

In an online reflection on badges, Barry Joseph[20] describes the

difference between badges that have local relevance for youth in their local peer networks vs. badges that have global relevance for recognizing competencies across institutions. He also inserts this subtle line that's loaded with significance: "Badges can be designed to offer both kinds of value—value within an organization and value to those outside it—but the required features and network are different." The implication is that providing local vs. global value can have wide ranging consequences on the type of badge design required, a point which I fully agree with.

Although this is a useful distinction, it feels like an outcome of other dimensions described above, rather than a dimension that is unique unto itself, and feels a little too specific to educational badges rather than a broad principle of rewards in general. It parallels most closely the concept of internal vs. external value, which I think is a more general framing of the core idea. "Local" reward value exists in the nearby community whereas "global" rewards allow you to carry that value between institutions. Internal rewards were meant to have value "within the game experience," but in a learning context that would seem to mean your local classroom/peer network. And in this analogy, external rewards would clearly mean moving the rewards between institutions (like from high school to college).

So why is Joseph's dimension worth discussing here, rather than just being a subpoint under internal vs. external rewards? Well, first, Joseph's points are very poignant and worth reading. But second, there's something to this dimension that might be useful to think about specifically in educational contexts, that is explicitly related to intrinsic vs. extrinsic rewards. If you also read Erin Knight's response post[21] and Carla Casilli's related post,[22] it starts to become clear that local, or "lightweight" rewards should allow learners to develop agency and identity through choosing what to focus on, maybe by piling up with other similar small badges in a leveling structure, or through unlocking similar opportunities in a hard or soft unlocking structure. In several ways, building rewards for local value will often lead one towards building an intrinsic reward structure, since that is something that easily conveys value back to the learner themself and others inside the local learning environment. It's the intrinsic reward structure that values the learner's experience and creates engagement for them. On the other hand, global value is almost synonymous with extrinsic

rewards that are discrete, simple and easily transferable amongst disparate systems. The complexity that is needed to make a local, intrinsic reward interesting to a learner needs to disappear to make a simple, global reward easily transferable throughout a broader ecosystem, creating a natural, and potentially unresolvable, conflict in reward design (this conflict is discussed more fully in Chapter 9).

So to go back to the quote from Joseph, I'd actually extend it and argue that the "required features" are so different between local and global rewards that they need to be different rewards with drastically different designs. This doesn't mean one badge system can't do both, but it means that one system has to have multiple different rewards coexisting within it (you can read about the GameStarMechanic example in Chapter 8 for an example of how one system accommodated both local and global badges). Because this distinction feels so important in educational reward settings, and because badge ecosystems have been discussed in depth in recent years, it felt useful to call out Joseph's distinction in this section and how it related to the other components of the taxonomy. And it's a great lead-in to the idea of hybrid reward structures.

Hybrids

It's worth noting that reward systems rarely if ever fit into only one category. Many times you end up with chimeric or hybrid rewards systems, in which one type of reward is layered underneath and feeds into a different kind of reward. I already shared some of these examples in the discussions of each dimension, but it feels worthwhile to specifically call out a few of them to show how these hybrid rewards structures can work in practice. This is not so much a unique dimension unto itself, but a meta-layer that can be invoked in reward systems to add depth and complexity.

The classic example is an experience points and leveling system in a classic RPG. In this example, you have a discrete reward system (experience points) that when certain threshold levels are reached produces a leveled reward (character level).

The Boy Scout system offers another similar example. A certain

number and variety of discrete merit badges add up to form a new Boy Scout rank (a leveled reward).

For a digital example very close to the Boy Scout system, take the technology tree in Civilization III.[23] Players gain technologies, which unlock further technologies, in a discrete badges system. When a key technology is reached, the player goes up to the next "age" which is a form of a leveling system. Unlike the boy scouts, all rewards are quantitatively given by a machine.

All of the above examples refer to discrete rewards producing a leveled reward system. For an opposite example, take the materia system in Final Fantasy 7.[24] In this system, you gain discrete rewards in the form of materia, each which gives your character a different ability. However, each materia can gain experience points and go up levels to become stronger, such that there is a leveling system within, or at a lower organization level of, each discrete reward.

With our taxonomy now fully described, let's see how it maps to the various rewards that we have been using throughout human history.

4. Using the Taxonomy

As was hopefully obvious from this previous chapter, there is not one and only one kind of reward. Although certain rewards can be characterized as "bad" and others as "good", those characterizations are very context dependent. What kind of reward might work best for a particular game or learning experience depends both on the kind of game/experience itself, as well as your objectives as a game/learning designer. Context was what helped us determine that a measurement based quantitative reward worked well at providing motivating feedback in the game Dance Dance Revolution, but poorly in the SAT math section. This of course was for a particular objective—for other objectives such "weeding out" poor performers from continuing in STEM fields in college programs (which depending on your perspective may or may not be a worthy goal), something like the SAT is actually an excellent reward system for achieving that objective. The point I want to make is that certain kinds of rewards are not inherently good or bad, but rather their context and objective is necessary in evaluating their worth. Most of the time when someone goes around saying "badges are bad," what they really mean is that "badges when used in certain contexts are bad for certain learning objectives." In order to have more meaningful conversations about rewards, we need to be more nuanced in our conversations, and avoid blanket statements about the worth of rewards.

This book began by making a distinction between intrinsic and extrinsic rewards. And the fact is that certain dimensions are much more intrinsic-reward friendly, some are much more extrinsic-reward friendly, whereas some are neutral to the intrinsic/extrinsic distinction and can be used equally well in either category.

Intrinsic rewards often feature some kind of leveled reward that increases in value the more that you play. These rewards often have internal values, sometimes directly relating to an in-game currency like money or experience points but very often relating to a hard or soft unlocking structure. They often involve quantitative, expected rewards that are permanent in nature. They can be based on either completion

or measurement metrics, depending on the kind of context that they are used in.

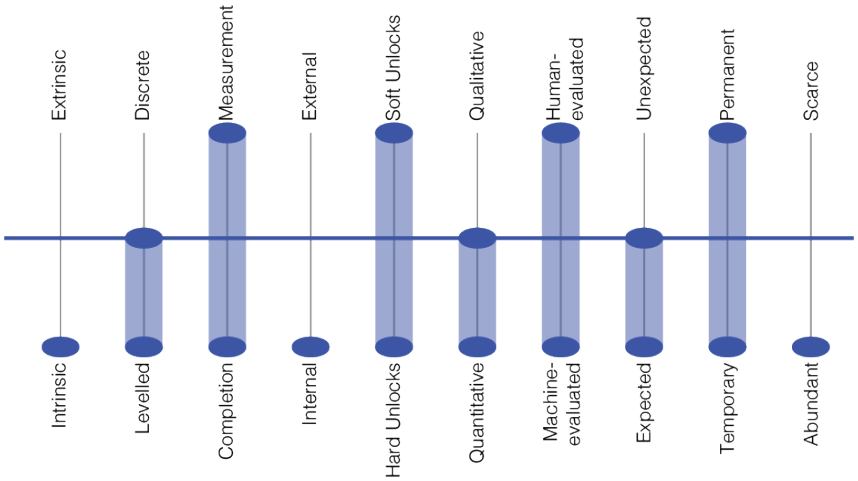

Image 4.1: A visualization of what dimensions of the taxonomy tend to be most used with intrinsic rewards.

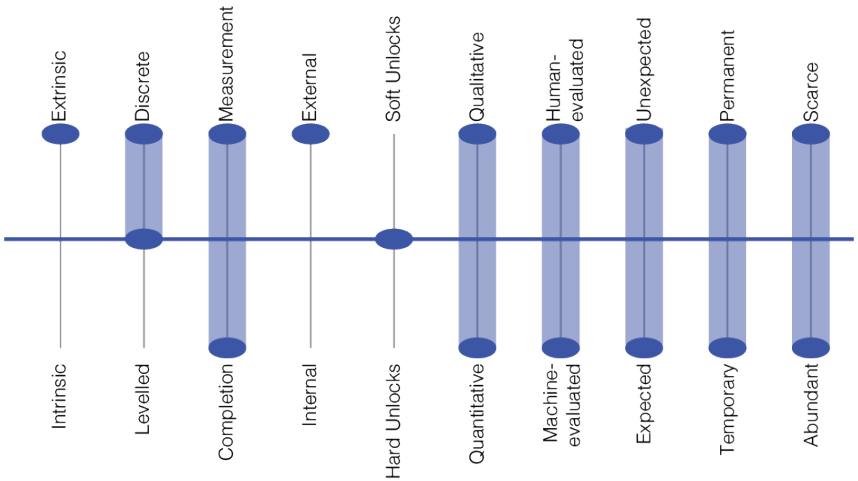

Image 4.2: A visualization of what dimensions of the taxonomy tend to be most used with extrinsic rewards.

Extrinsic rewards are often discrete, external rewards. They are not usually related to any kind of unlocking structure. On virtually every other measure, they can go either way.

It's worth also going back and considering where our historical examples fall on the spectrum given above. The variety of rewards throughout history is quite interesting.

Military

There are two rewards systems in the military: military rank and military awards. In general, both categories of military rewards are defined by being very completion-based, human-evaluated, and permanent in nature. Let's go over differences between the two types separately.

Military rank is a leveled, completion-based award. It has a very distinct internal value, in that higher ranks give you more power in the military organization, although it can also be cited in places outside military service, granting it external value too. The ranks are also permanent when given, and can be cited long after leaving the military as an indication of experience and knowledge. It's worth noting that there is a temporary form of these ranks that can be invoked in certain situations (i.e. your leader was killed amid battle), which allow a higher rank to apply to someone for the duration of a battle or period of time, but are designed in these cases with clear temporary status. The ranks are also given based on the authority of superiors, and so can also be called human-evaluated. They are also scarce, in that there are a limited number of higher level ranks that can be filled.

Military awards are a classic example of a discrete, completion-based award. These awards are meant to be given based on some merit or event achieved as evaluated by others in the military, and so are clear qualitative, human-evaluated awards. For example, a military superior in the field can evaluate and award the Purple Heart medal if they feel the criteria for the award is met. It is also quite permanent in nature, and except for exceptional circumstances is not taken back. The value of this award is entirely external in nature.

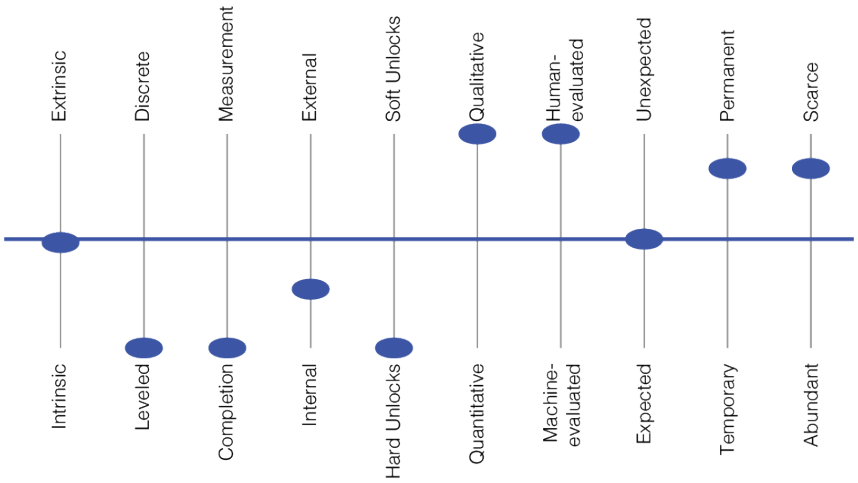

Image 4.3: A visualization of where military ranks fall on the taxonomy.

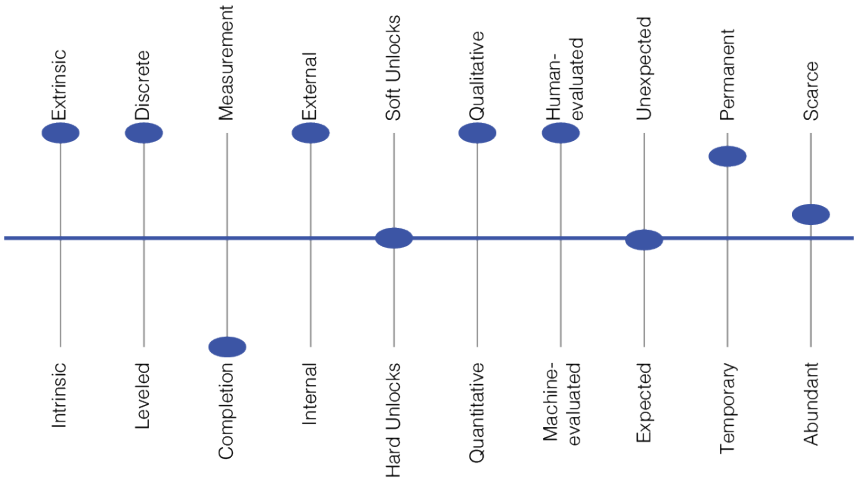

Image 4.4: A visualization of where military awards fall on the taxonomy.

Scouts

There are two interesting forms of badges in the Boy Scouts system: merit badges, and badges of rank. Both badges are very completion

based and form an interesting hybrid discrete plus leveled badge structure.

The merit badges are the most classic example of discrete completion-based rewards given by experts. The badges do undergo a qualitative review process, but the criteria of the badge are often clearly laid out, and some badges, although approved by an expert, would feel more like the expert is stamping the fact that a quantitative set of criteria are completed rather than making a qualitative judgment. The rewards are also very explicitly expected: a scout will plan out their intent to complete and progress towards a badge with an expert. And they are also quite permanent and abundant.

The badges of rank are built upon the merit badges, in one of the most classic examples of a hybrid reward structure given in the last chapter. The badges of rank are very much still completion oriented, but are leveled. First, they are achieved by a set number of merit badges, so the merit badges themselves act as points that level up into a badge of rank. Second, there are multiple badges of rank that can be achieved and are leveled with respect to each other. The ranks also act as strong hard unlocks—upon achieving a new badge of rank, a new set of more difficult merit badges is opened up to you. They are also very abundant and expected, often planned by the scout in advance.

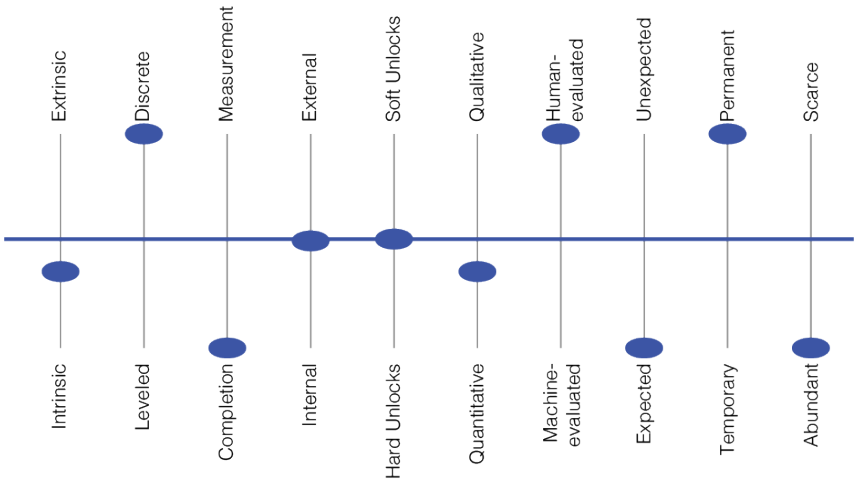

Image 4.5: A visualization of where merit badges fall on the taxonomy.

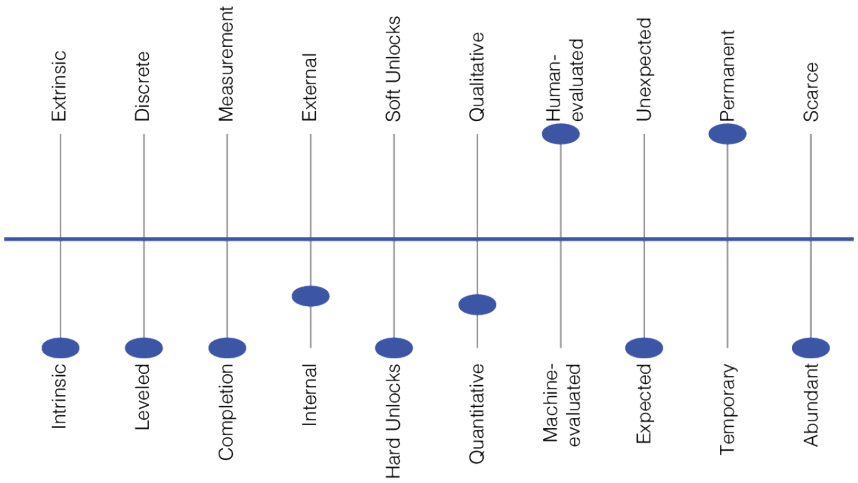

Image 4.6: A visualization of where Boy Scout ranks fall on the taxonomy.

I cannot emphasize the extent to which the Boy Scout has an interestingly complex and well-designed reward structure. It's really one of the few examples of a non-digital reward structure that contains many of the complexities that we now find commonplace in our digital games and apps. It's with great honor that the Boy Scouts system can be pointed to as a potential source of inspiration for many of the digital rewards structures found in modern times.

It's also worth noting how different the Boy Scout system is from so many of the other historical reward systems—most notably with its emphasis on completion-based merit badge rewards that tie into a very hard unlocking rank system. The difference to a school environment could not be more stark.

Education

We discussed two main kinds of awards in education: grades and diplomas. Both education awards offer an element of qualitative judgment by human-evaluation, and have external value—beyond that, grades and diplomas are quite different in structure.

Grades are a classic form of measurement-based awards given based

on feedback from a human expert, namely the teacher. It's worth noting that the grades do vary between being based on qualitative or quantitative criteria, often based on the subject: math for instance tends to have much more quantitative grading with a clear indicator of whether an individual answer is right or wrong, whereas english tends to involve more qualitative grading. One of the main advantages of using a quantitative award is to allow instantaneous, objective machine-gradable feedback. Depending on the class, grades can be more or less permanent, based on whether or how often a teacher allows students to resubmit work for higher grades. Grades are mainly extrinsic in nature—the school year continues moving on whether or not you get good grades. Of course, if you do fail most of your classes, you will not advance to the next year in school which does give grades a sort of longform intrinsic reward loop. But the timing on this loop is rather infrequent compared to the frequency with which grades are given through the school year, such that for the most part, grades feel extrinsic. Individual grades on assignments are also very discrete, but can be averaged together into a final grade on a course. This is a unique phenomena that can be achieved by discrete, measurement-based rewards—these kinds of rewards often are "averaged" rather than "accumulated." Another example of this will be given in the Food Industry section. Grades can also vary in their abundancy and scarcity depending on the class—if everyone can get an "A" then it's fully abundant. But often grades occur on a curve, which effectively limits the number of students that can get an "A".

Diplomas are very different in structure. They are completion-based awards—you have a diploma or you don't, although diplomas can be given "with honors" which is a modest form of leveling. Their other defining features involve permanence, and having both internal and external value. Diplomas are permanent, and are only rarely if ever taken away from someone. The diploma has internal value to the education system as a whole as it allows you to advance from one stage to another in a hard unlocking system (e.g. high school diplomas allow advancement from high school to college). On the other hand, diplomas also most notably have currency in the workforce, allowing one to apply to certain jobs, or to get higher pay. The higher pay especially is a literal example of cashing in the diploma's value for external currency, an exceptionally clear example of the meaning of

external reward. In this sense, diplomas have both an intrinsic reward loop (in unlocking further learning opportunities) and an extrinsic reward structure (in enabling better employment opportunities). Like our photography example in the introduction, it mainly depends on someone's framing in determining how they perceive a diploma.

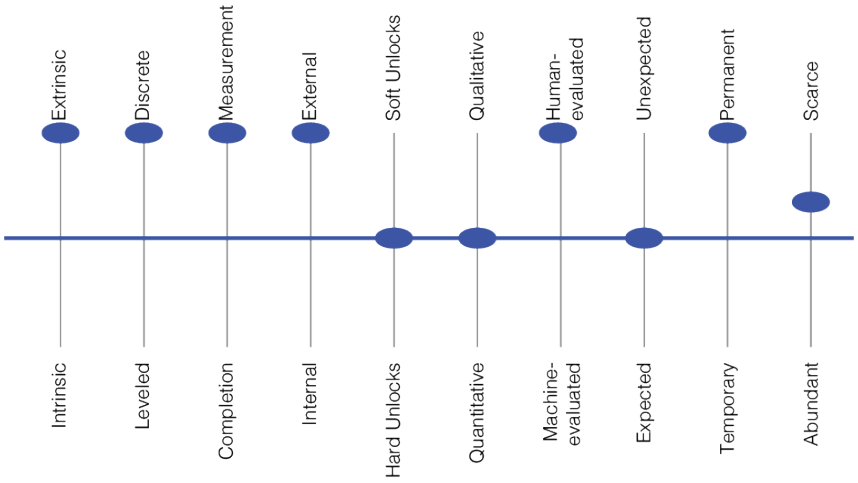

Image 4.7: A visualization of where traditional school grades fall on the taxonomy.

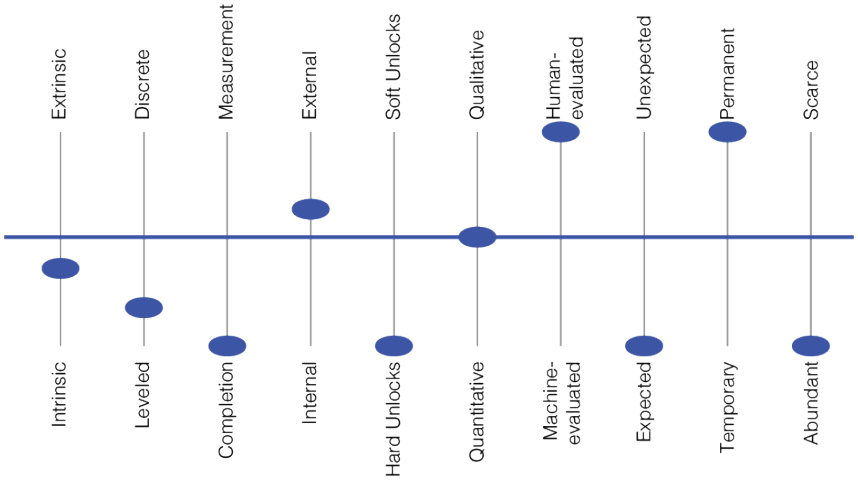

Image 4.8: A visualization of where diplomas fall on the taxonomy.

Food industry

We discussed three main kinds of rewards in the food industry: Michelin stars, Food safety grades, and Yelp reviews. For the most part, all of these are extrinsic in nature, qualitative, human-evaluated awards, and temporary. Nothing about the reward allows the restaurants to practice being a better restaurant (except potentially the Yelp reviews). But, there are some differences too.

The Michelin stars are a discrete, scarce award in that not every restaurant has them or can get them, and there is a cachet in being a Michelin restaurant. In some sense, even gaining the status of a Michelin restaurant is a sort of completion-based reward, but the reward itself is explicitly measurement-based in the form of having one to three stars. The reward's number of stars is determined by qualitative feedback from experts. Most interesting is that the reward is both temporary and unexpected. It's temporary in that you can lose the award from year to year, and unexpected in that you don't know when you might be evaluated by a Michelin critic, who visits the restaurant in disguise.

The food grade system is similar to the Michelin stars in some respects—it is also a qualitative, temporary, measurement-based feedback given by an expert. The main difference is in the expected nature of this reward—there's far more certainty and timing in when restaurants are reviewed for food safety ratings. And it is abundant; every restaurant can (and probably should) get the highest rating.

The reviews on a platform like Yelp are very similar in many respects—they are a reward given based on qualitative-measurement-based feedback. But there are some differences. In this case the reward is human-graded by peers or amateurs, rather than experts. Additionally, the reward is permanent, not temporary. Yelp reviews don't really ever go away, unless the author deletes their review. Although your overall rating can change and fluctuate as more reviews come in, every review is treated as an individual, discrete reward. It's interesting to note that although there are many discrete rewards, they do not "accumulate", they average together, a unique feature that can be accomplished by a series of discrete, measurement-based rewards (also found in grades in the Education section).

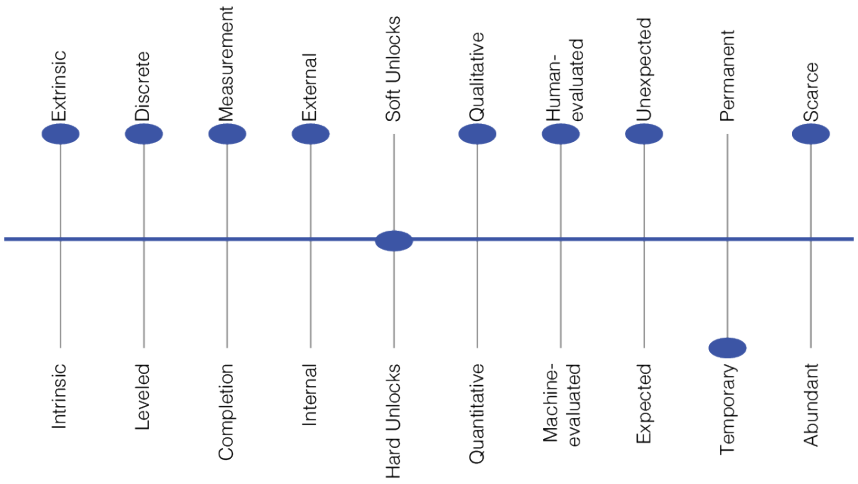

Image 4.9: A visualization of where Michelin stars fall on the taxonomy.

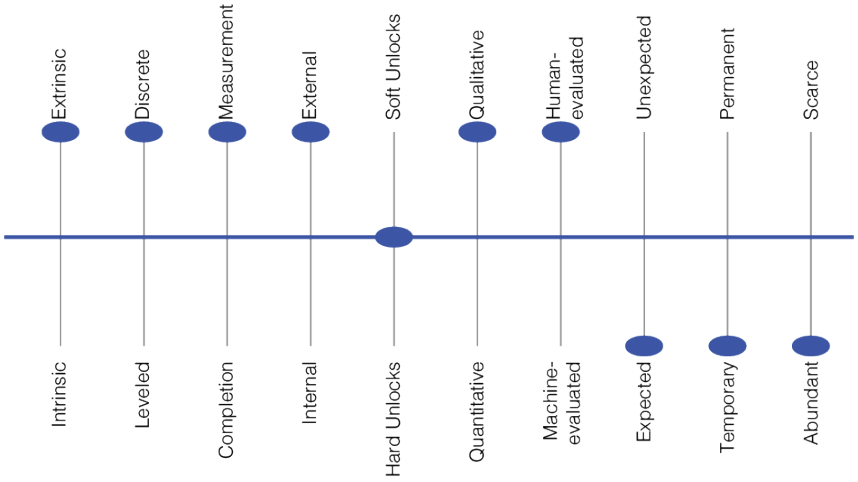

Image 4.10: A visualization of where the food grade ratings fall on the taxonomy.

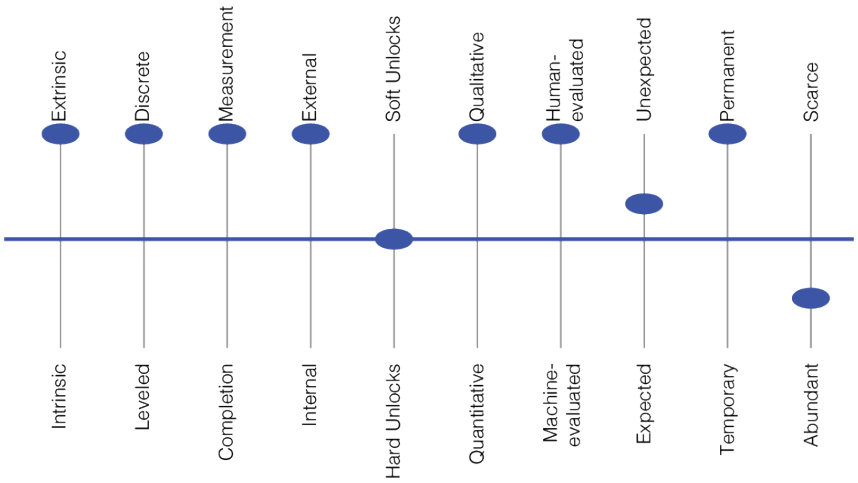

Image 4.11: A visualization of where Yelp reviews fall on the taxonomy.

Early Games

There were three examples of early game reward structures: the ranking system of Go, the Elo ratings system, and the rewards found in Pacman. All are similar in being quantitative, and their ability to foster an intrinsic reward loop in some situations—other than that, they are quite different in most respects. And even in those two respects, they are not all similar.

The ranking system in Go is a relatively unique example of reward structures. It showcases leveled design and has very strong intrinsic value, mainly because the rank is used to inform future play so strongly. This is unique because many systems with strong internal value usually feature some sort of currency, which is usually expressed in many discrete awards, rather than one leveled reward. The system is also a good example of a soft unlock system—there's nothing preventing you from successfully playing at a higher rank except your own skill. These rewards are also a hybrid mixture of quantitative and qualitative rewards—they are qualitative in that they only have meaning with respect to other players, but are quantitative in that they have a very specific meaning in that system and are defined on an exact system relative to how you play (every 10-point difference in your final score

translates to one rank). Because the reward is defined relative to others, it also creates a degree of scarcity. It is also a measurement-based reward, as your rank is a measure of your overall skill.

The Elo (or Glicko) system is similar to the Go ranks in several respects—it is also a hybrid quantitative/qualitative structure. But it most notably differs in lacking internal value, and having a large degree of scarcity. In some digital games with a matchmaking system, the Elo system can be argued to have internal value with a soft unlocking structure as higher points correspond to having more difficult matches, but not every system uses the Elo for such a purpose. Similarly, although the USCF was noted to create ranks out of Elo scores, the use of leveling in Elo is not widespread. Although there are clearly many ways to repurpose the score, it is most notably a measurement-based reward of your skill (though it is not "averaged" like grades or Yelp reviews, but evolves over time according to a formula). The higher granularity of the reward relative to the Go system reinforces its value as a measurement reward.

The two reward systems in Pacman also straddle the opposing ends of the spectrums described above. The points system is an accumulated discrete award, the levels are a leveled reward. These two systems are almost the quintessential archetypes of a discrete and leveled reward. They also have other differences, with the levels being a completion-based award with internal value that hard unlocks later levels, and the points being a measurement-based award with external value (bragging to friends, or being on the high score leaderboard at the arcade). Besides these differences, they have a major similarity of being entirely quantitative, which is true of almost all reward systems in single player digital games. They are also both expected rewards (you know exactly what sequence of actions will produce the awards), are abundant (every player who plays the game could get the highest scores) and are temporary—they only exist for that playthrough and are reset when you lose all your lives and restart the game.

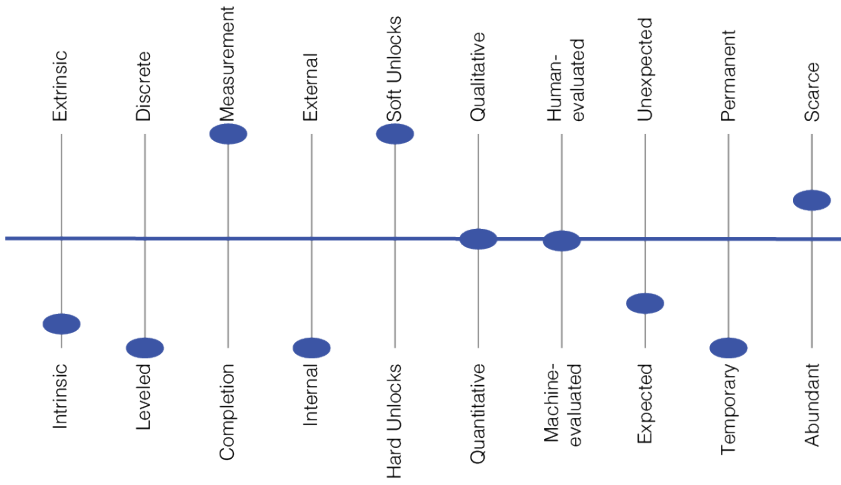

Image 4.12: A visualization of where Go ranks fall on the taxonomy.

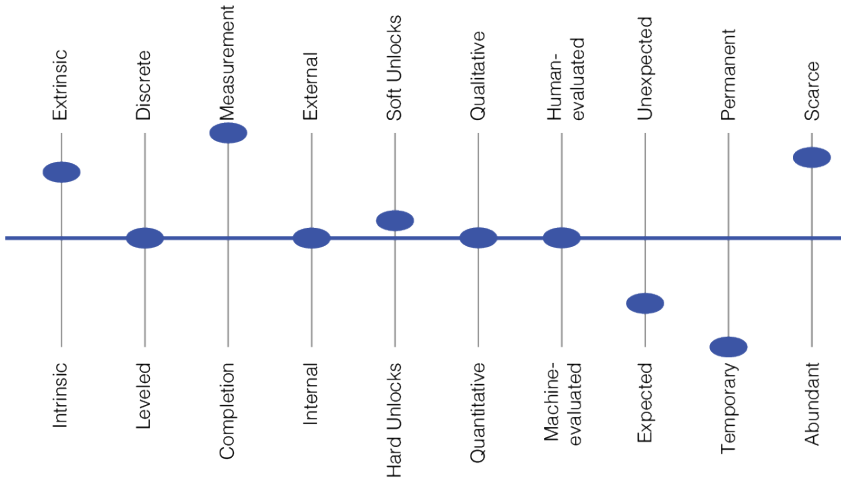

Image 4.13: A visualization of where the ELO type rating system falls on the taxonomy.

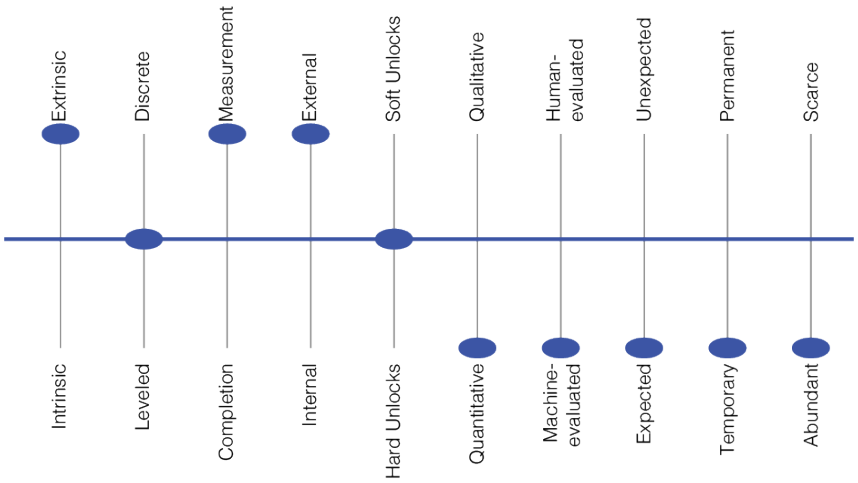

Image 4.14: A visualization of where Pacman points fall on the taxonomy.

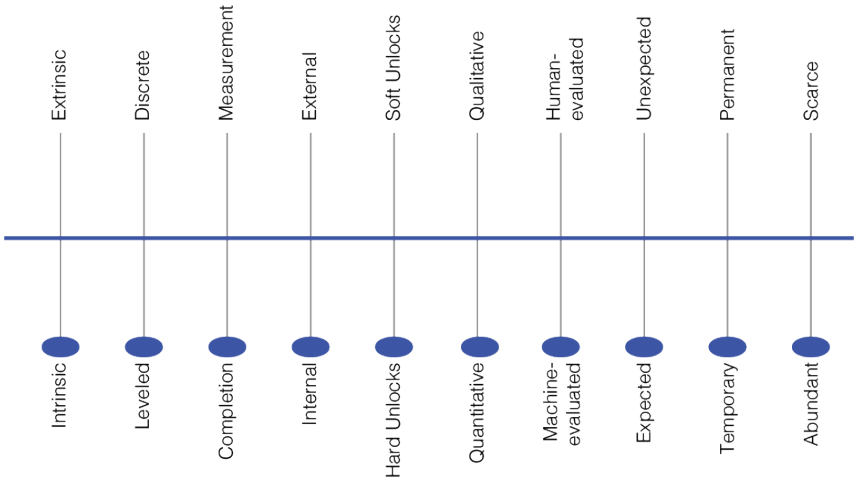

Image 4.15: A visualization of where Pacman levels fall on the taxonomy.

Looking forward for a second into further digital games that evolved from Pacman, many of these same elements still ring true: accumulated points and levels with internal value, quantitative and expected. One main difference is that the temporary nature of the

rewards was an interesting function of early arcade games that did not have save states—as later games evolved the ability to save progress, games no longer had to be "reset" but could be "continued." This difference had a far-reaching consequence in allowing the rewards to be far more permanent in nature.

Research on rewards and intrinsic motivation

As mentioned before, the research is less complete than one might hope on offering recommendations for this taxonomy. But, there are a few key studies that can offer some insights.

The earliest study was published in 1999 by Deci, Koestner, and Ryan[1] and their ideas were continued with a followup report in 2001.[2] They looked at which aspects of rewards might undermine intrinsic motivation for a task. This work served as an early meta-analysis on what factors had been shown to increase and decrease intrinsic motivation. Of course, this work was done before rewards existed in high prevalence in digital games and platforms, so their conclusion—which were developed for physical rewards that are mainly given by human systems—might have limited applicability to digital systems that are mainly awarded by machines. Given that qualification, the findings are still very insightful.

Their high level summary is that rewards which give information about performance tend to increase intrinsic motivation while rewards that act to control or coerce the learner tend to decrease it. These are not quite design principles so much as statements on the outcomes that should be achieved by certain designs. I have tried where appropriate to accommodate this thinking in the taxonomy in recommending how to use given dimensions.

They offer several specific design recommendations on how to implement these principles. They talk about how expected and "tangible" (which I think in a digital age could mean "permanent" or "external") rewards are the worst for intrinsic motivation. Verbal rewards are generally described as ok, mainly because they lack the permanence of tangible rewards, and these rewards can be filled with information. They also talk about the "contingency" of the rewards,

what we might call now the criteria needed for achieving the reward. They found that rewards that were non-task contingent, like participation rewards, tended to have no effect on intrinsic motivation. Engagement-contingent rewards, or rewards given for starting a task, tend to have a significant negative effect on intrinsic motivation, as they control action. Completion-contingent rewards (rewards for completing an activity) also have a significant negative effect—this tends to be quite controlling but potentially with some information, depending on what completion involves. The simple act of being able to complete something can be informative in its own way. Finally, performance-contingent rewards do have a negative effect, but it's much smaller than the previous two effects. Basically, this reward can be controlling, but also contains a fair bit of information about performance, and so the reasoning is that the two effects balance each other out somewhat. Overall, this work is an inspiration behind the completion vs. measurement dimension, but the idea of control vs. information can be applied to concepts well beyond that one dimension. For example, leveled rewards contain more info than discrete rewards, and human-evaluated rewards can feel a bit more controlling than machine evaluated rewards.

For the next work, we need to jump forward to 2013, to a study by Abromovich, Schunn and Higashi.[3] These researchers implemented a few different kinds of badges in a math tutorial system. They compared participation rewards to skill rewards given for demonstrating ability. This harkens back to Deci, Koestner and Ryan's categories. The rewards themselves are described in the paper; the "participation badges" seem to be a mix of engagement-contingent and completion-contingent rewards. On the other hand, the "skill badges" are very clearly performance-contingent rewards. They also measured the concept of goal-expectation, in particular in performance vs. mastery goal orientation. These concepts end up being very analogous to the concepts of fixed and growth mindsets, discussed more fully in Chapter 9. I'll avoid going into a description of goal orientation or mindset right now. Instead, in describing the results, I'll simply note if there was an increase or decrease in the type of goal orientation that is good for learning.

The results were very interesting as they indicated that the rewards had different effects on different students. Low-achieving students

seemed relatively unaffected by skill badges, but strongly affected by participation badges in a mostly negative way (it decreased the type of goal orientation that is good for learning). High-achieving students seemed unaffected by participation badges but were affected by skill badges in a mostly positive way (it increased their expectations for success and therefore confidence in the subject). What is interesting is that the direction of these effects generally correspond with what was suggested by Deci, Koestner and Ryan, but that the effect did not apply to all students equally. Or put another way, context matters: how a learner approaches a reward matters as much as how the reward is designed. Or, engagement-contingent rewards do seem bad for learning, but only for some students, whereas performance-contingent rewards conversely seem good for learning (maybe the positive effects of feedback more than outweigh the negative effects of being controlling), but not for all students.

Next, let's look at a study by Filsecker and Hickey in 2014.[4] In this study, the researchers looked at students using a science learning game called Quest Atlantis,[5] and created two conditions—one in which students received rewards after completing certain assessment activities, and one in which they did not receive the rewards. The rewards were technically given by humans, but since the researchers were behind the scenes it probably felt more to students as if they were given by the machine. Interestingly, although the rewards were extrinsic (they were not embedded in a reward loop) the researchers also gave the rewards clear internal value—once you got the reward, it could be pinned on your avatar and displayed in game. The rewards could also be displayed in your classroom on a leaderboard, giving them an additional internal value (if you consider the classroom to be a part of the full game/learning experience).

The reward would qualify as a clear completion-contingent reward which means it should have a negative effect on students based on the previous work discussed, but Filsecker and Hickey adopted a different perspective on the issue. They first noted that the issue of competition can be harmful to rewards, but only in "impoverished learning environments," or learning environments that tend to both lack inherently interesting tasks for students, and to be a little controlling in how they enforce student learning. In well-structured learning environments, of which Quest Atlantis seems a good example, rewards

that otherwise might be harmful could be beneficial because there are more interesting and student-agency-respecting opportunities for learning.

Additionally, the authors take a socio-cultural perspective, which basically means rewards need to be considered in the context of a student's interactions and social environment. This influenced their decision to give the reward internal value, both in the game and in the classroom environment. One way of considering this perspective is that when rewards have internal value, the reward might feel less controlling because it allows students to more deeply engage in their socio-cultural setting rather than simply manipulating them towards a certain course of action.

Thus, one way of looking at this study is that it starts to truly mimic some of the more powerful uses of rewards in digital games and media, rather than simply being digital versions of the same physical rewards that Deci, Koestner and Ryan studied. By incorporating more modern and innovative designs, the rules on what works and what doesn't could be rewritten.

In fact, this does seem to be true. Filsecker and Hickey generally found no significant increase or decrease in student motivation for those that use badges, but they did find significant increased learning gains for those groups. It's worth noting that the badge group did have higher levels of motivation than the non-badge group, even though it did not reach significance. This all in general seems to counter the main results of Deci, Koestner and Ryan, by noting that factors other than just the contingency of the reward (namely, other dimensions offered in this book) can affect and possibly override the effect of reward contingency.

In the same year, another study emerged from a group tackling the problem from a different angle. In 2014, O'Rourke et. al[6] wanted to know if the type of feedback affected a students mindset (discussed more fully in Chapter 9). They designed two versions of a game and tested each version with thousands of students through an online educational gaming platform. One version contained more vanilla badges—these badges were given out as a reward for completing a level, and had neutral messaging. These could be called "completion-contingent" rewards or discrete, completion rewards by our taxonomy. The rewards could also be "cashed in" to advance to the next set of

levels and advance through the game, creating a rough semblance of an intrinsic reward structure, or at the least giving the rewards internal value.

A second version of the game represented the intervention condition, and contained something called "Brain Points." These rewards were given out at the end of the level, but were only rewarded if 2 of 4 criteria were met in the level. The criteria were based around actions students performed when solving levels that indicated the presence of a growth mindset: testing new hypotheses, clearing the board to start fresh, making math progress, and making multiple moves. This then forms a sort of performance-based reward, though it is not strictly measuring performance, or ability. In our terminology, this is a discrete, measurement-based award. The rewards also contained growth-mindset-inducing feedback throughout the game, whether or not a reward was given out. It's not possible to see if the alterations in reward design vs. the feedback language itself caused the changes seen in this study, and to an extent it's a bit of a moot point anyways—the language used would be disingenuous to students if used with the control condition's badge structure. To truly acknowledge growth mindset activity in a meaningful way, something more than discrete, completion-based awards was needed. These rewards also awarded points which could be cashed in to advance to new worlds, giving an intrinsic reward structure with internal value.

The results are quite encouraging. They found that the Brain Points intervention increased persistence in the game and promoted growth mindset play patterns (both effects are small but significant). The intervention also helped struggling children persist more (which was also a small, but significant, effect too). The intervention condition may help students react to tough challenges in a more positive way. In other words, the Brain Points system was able to change a student's mindset in playing the game simply through changing the rewards system design and related feedback messaging. This is again further evidence that rewards that more directly integrate best practices from the game industry are able to show relatively promising effects.

The above examples are single studies of particular systems, and of course the larger pattern in the literature could show a different pattern than any single study finds. One way to study large scale patterns is to look at meta-analyses, or research that looks for patterns amongst

other published studies. There are, at this point, several meta-analyses into game-based learning and game-based rewards. However, the majority of such studies focus on the primary question of "are games-based systems effective for learning/engagement?", or they have focused on what dimensions of learning are most fostered by games,[7] or what subject areas have benefited the most from games.[8] Some of the above studies have gotten into design, but mainly from a media perspective (e.g. graphical realism), and not from a game design perspective. Very few have focused on what design elements of game-based systems have led to the most learning or engagement, which is the primary subject of this book. There are two meta-analyses that do discuss this issue, and we'll dedicate a little time to each one.

First is the study by Clark et. al in 2016.[9] This study looked at the design elements of games used in learning settings, but many of the design elements chosen are about the reward structures. This study found that game conditions produce more learning than non-game conditions, but that the type of game mattered a lot when looking for a positive effect. Single player without a competitive or cooperative component has a strong significant effect, cooperative games have a potentially also significant effect, but competitive games tend to not have a significant effect on learning (in fact it may have a detrimental effect). "Simple gamification" (which they define as just points and badges) was found to be effective, but so were more complex gamification elements that went beyond points and badges. It's unclear what "going beyond points and badges" means, as the taxonomy we just established noted the complexity of design that even points and badges can have. The study does generally confirm both simple and complex gamification as effective. The authors have found that there are also positive effects associated with integrating game elements into the learning design, and that enhanced scaffolding (which can also be described as feedback) is more beneficial than low or no scaffolding. Finally, there is a very interesting analysis on the issue of media design, that goes into elements like camera viewpoint, realism in graphics, anthropomorphism, and amount of narrative. But the key finding is that the effect of different game design elements is at least as large as the effect of what form of media is chosen, which furthers the argument to pay more attention to game design elements like the ones described in this book. The final sentence of the article

really drives home the point: "We should thus shift emphasis from proof-of-concept studies ("Can games support learning?") and media comparison analyses ("Are games better or worse than other media for learning?") to cognitive-consequences and value-added studies exploring how **theoretically driven design decisions influence situated learning outcomes** for the broad diversity of learners within and beyond our classrooms." (bold emphasis is my own)

Second, a study by Dichev and Dicheva in 2017[10] focused specifically on "gamification" systems, or reward structures. This study primarily reported a lack of consistency in the type of gamification systems used, and correspondingly a lack of conclusive patterns in the literature. Although more positive effects were found than negative effects when gamification systems are used, the "inconclusive results" category was by far larger than the positive and negative results categories. They additionally found that there was little consistency or rigor in what game elements were chosen for inclusion in studies, often no justification offered for why that combination of game elements was chosen, and reward structures were often used without a theoretical framework guiding the implementation. As they put it: "the process of integrating game design principles within varying educational experiences appears challenging and there are currently no practical guidelines for how to do so in a coherent and efficient manner." As such, they concluded that it is almost impossible to identify what specific elements or principles might be leading to the success of gamification systems, if any. The primary conclusions of the study are to argue for a more nuanced approach in future studies, noting that their results "indicate a need of a systematic program of experimental studies mapping game elements to the learning and motivational specifics of individual (groups of) learners," making a very similar point to that of the previous meta-analysis.

The taxonomy of rewards offered in this book is hopefully seen as a way to characterize rewards throughout our history, and as a way to make sense of some of the seemingly contradictory research findings. At this point, it's worth taking a dive back into our original idea of intrinsic rewards, and exploring how this particular dimension has manifested itself in the game industry.

5. A Close Analysis of Intrinsic Rewards in Games

With our basic definitions and a taxonomy of rewards defined, it's now time to take a deep dive into what intrinsic rewards look like in games. I will start with one of the classic models of intrinsic rewards—grinding in an old-school RPG. I will then move through different kinds of games, and see how the basic reward loop is modified by different games and game genres, while keeping the basic idea of a reward loop intact. To avoid an overly lengthy chapter, I will avoid describing the exact details of each game's mechanics and play style, and so this chapter is naturally geared more towards those readers who are avid gamers and familiar with many of the games mentioned, or other games like them. For those readers who are not avid gamers, I have tried to describe just enough of the game for you to follow the argument without getting lost (but I'd encourage you to play the games themselves if interested, as there's no amount of description of gameplay that substitutes for direct experience). I have also shared info or videos of gameplay in the notes to allow you to experience the basics of the game in some form, especially when the game itself is old enough to be relatively unavailable to play in its original form.

Let's start by taking a look at games that feature "grinding", as that's often indicative of the presence of a reward loop. Grinding is a gaming term that refers to the process of repeatedly performing an action to go up levels. In the classic RPG game, grinding refers to repeatedly battling monsters to gain experience points and money, but grinding actually takes many forms in today's diversity of games. When players grind in a game, I would argue they are undergoing multiple cycles of an intrinsic reward loop.

To me, the classic intrinsic reward game is Final Fantasy I[1] (which I'll refer to as FF1), because it is one of the first games in which I experienced grinding, but it is also one of the classic RPG games that

helped define the genre. In this game, you explore a world, finding towns, dungeons, and monsters. There is a story to the game that you progress through by adventuring, but that adventuring mainly occurs through repeated encounters with enemies big and small that you need to battle. Typically, the further you travel from your starting location, the more difficult enemies get, such that you need to become stronger to fully explore the world and complete the story. Tangentially, this is a description of a classic game-based soft unlocking structure—in this example the only thing preventing you from venturing further from your home base is your own strength. In the Final Fantasy 1 game itself, this basic soft unlocking structure is intermixed with several hard unlocks, to create an overall game progression and story.

There is a very basic loop created in this game. Namely, you 1) fight monsters, to 2) get gold/exp, to 3) go up levels and buy better stuff, to 4) get stronger, to fight better monsters and repeat the cycle. The key components of the cycle are steps 1) and 3), as these are the steps where a player performs actions and has agency. In this game, the fighting of monsters is somewhat strategic and takes some player skill that is independent from how powerful their characters are (you need to know what attacks to use against what enemies in what order). Buying stuff also has some strategic value, mainly because you never have enough money to buy everything that you want, even if it's always clear what you should buy for which character if you have the money available. But in general, having more money and levels makes you better at fighting monsters. So in general, you end up with a pretty basic reward loop.

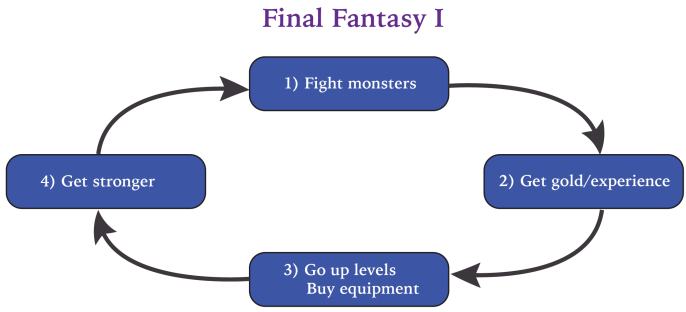

Image 5.1: The reward loop for Final Fantasy I.

At this point, let's jump to another game, The World Ends with You[2] (TWEWY for short). In this action-RPG, you progress through a story which is divided into days that you play through. In each day, there are certain quests or objectives that you must complete to advance past the day (offering a classic hard unlocking system), and those objectives often require battling enemies called noise. Unlike FF1, battles are not randomly generated as you wander—you choose when you fight battles in TWEWY. Additionally, the battle system is extremely dynamic, and involves executing complex series of actions by tapping, swiping, and pressing on the Nintendo DS touch screen. What actions you can do depends on what "pins" and items you equip, each of which has different powers and actions that it enables. Both the player themselves and their pins can grow stronger over the course of the game.

In this game the formula really isn't that different from FFI. 1) fight noise, 2) get XP and Yen, 3) raise the levels of your character, pins, and buy better items, 4) get stronger, and fight more difficult battles. What's really different here are a few minor but important details. I'd argue that step 1) is now an even more meaningful interaction than FFI, as every fight is basically a complex puzzle that requires different kinds of actions to be successful. Additionally, finding the best pin/item equipment is also a puzzle in itself, making step 3) a more meaningful interaction. But the biggest difference is that a significant amount of player agency has been added to determining difficulty of battles. Players can constantly decide how difficult they want the game to be, by lowering their level or increasing the game difficulty. Doing either increases the proportion of valuable "drops" (or the rewards given out after a successful battle), which allows the player to get stronger faster. In other words, players now also have control over the pace of grinding, or the speed at which they cycle through the intrinsic reward loop. This control over the difficulty level creates an even stronger form of a soft unlocking system, as players can really feel their ability to play the game better than by setting themselves to a lower level and still being able to complete one of the more challenging battles or days.

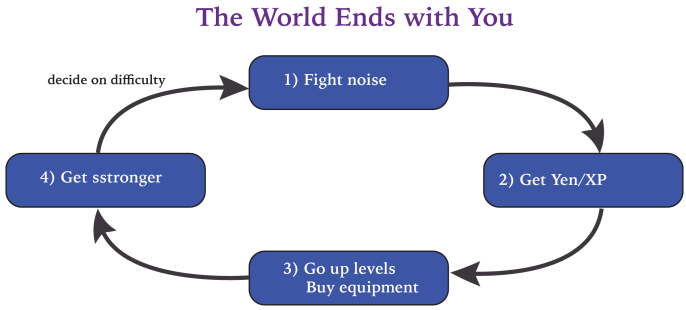

Image 5.2: The reward loop for The World Ends With You.

Jumping one step further in this direction are puzzle games, like Osmos.[3] Osmos is an intriguing kind of puzzle game, in which you play a sort of bubble that can absorb other bubbles to grow bigger—but can only absorb bubbles that are smaller than you. The act of acquiring smaller bubbles repeatedly over time is a kind of grind within a level, but also most importantly for our purposes you are grinding your personal skill through repeated play attempts, without any real notion of a reward. This is somewhat similar to TWEWY in that you did have to grind your ability to play battles well, but in that game you also had stats that grew stronger over time through intrinsic rewards built into the games (namely, leveling up), so your own personal skill was not the only thing causing your performance to increase. In puzzle games, the game rewards are taken out, so the loop looks like this: 1) beat puzzles to 2) increase your puzzle solving skill to beat more difficult puzzles. There's lots of parallels to other puzzles like crossword puzzle or sudoku. Note, that this is similar to how hobbies and other real life activities work, and is an even stronger soft unlocking structure. It's worth noting that Osmos does have a series of levels, and that you need to beat earlier levels to unlock later levels that increase in difficulty. This is both an explicit reward structure, and a hard unlocking structure, which seems counter to the point above. The challenge is that to unlock higher levels, the main issue preventing that from happening is your own play ability, which can only be improved by repeated play attempts (aka grinding) and doesn't have any explicit reference in the game itself.

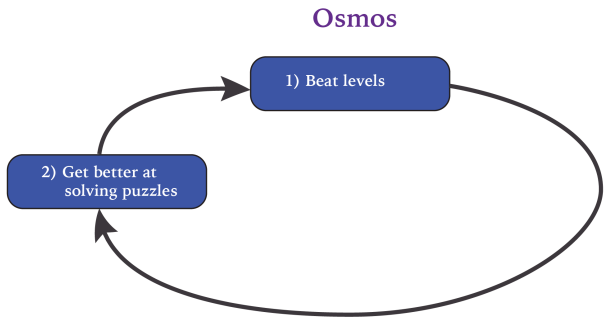

Image 5.3: The reward loop for Osmos.

As another example, take Clash Royale,[4] the popular multiplayer mobile game. This game features a simplified version of the basic tower defense game format, combining real-time strategy with something that also feels like a deck-builder card game. This game also contains a clear intrinsic reward loop, in that you play battles to unlock reward chests, to collect cards, to level up your creatures. Although the loop is interesting, it also contains elements that decrease the value of the loop. First, you progress towards unlocking chests whether or not you win, although you progress further when you win. Second, you cannot choose your reward, the reward is given out completely randomly when you open the chest. Compare this to FF1, in which the character that you are using is the one that gains experience and levels up—in Clash Royale, a random character is leveled up with a win. There is also a limit on the number of rewards you can gain each day, which can only be circumvented by spending money. The game though is interesting in that there are multiple ways to gain rewards, though each time the reward is random. For example, you can join a clan, and your clan can play in "clan wars." Your clan as a whole can then gain levels, and at the end of each battle period your clan will be awarded a reward based on the clan's level. This is again a multiplayer-based intrinsic reward loop, in that the clan fights clan battles to go up levels to get better rewards to have a stronger total clan. The clan can also trade unwanted rewards with each other too, allowing members to specialize in the cards that align with their strategy using cards that are otherwise useless to other

clan members. Thus, random rewards in a way enable a more strategic conversation around sharing rewards in a clan, creating some element of strategic choice in the rewards stage. If considered with clan reward elements, then an intrinsic reward loop seems clear.

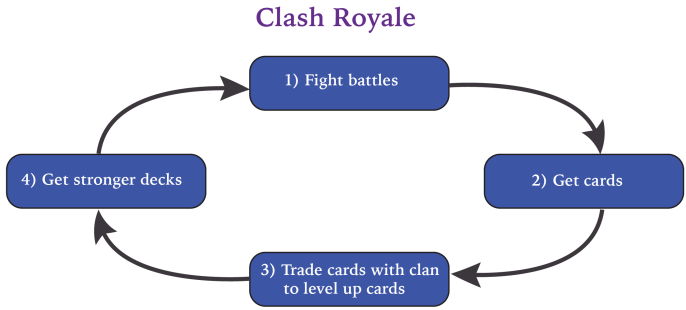

Image 5.4: The reward loop for Clash Royale.

To take one more example, let's look at the gun mettle campaign in Team Fortress 2.[5] Team Fortress 2 is one of the classic first person shooter games, structured around several different game modes like capture the flag or king of the hill. This game genre itself doesn't have a strong grinding structure built into it, but that changed when an update released the Gun Mettle campaign, a series of quests that were basically achieved by completing grinding-like actions in their multiplayer game. The actions are performed on multiplayer games, and so in some sense aren't like an internal game grinding structure. The quest rewards are also extrinsic rewards and social in nature (which makes sense, given that the game play experience is social). So the structure is 1) Play multiplayer games to 2) get gun mettle XP points to 3) get better at performing the action that gets you those XP points. Action 4) is to repeat 1-3 until the quest is achieved and you unlock a unique-looking gun that makes you look extra cool in future multiplayer matches. It's unclear whether 4) links back into 1) and creates an intrinsic social reward loop, or simply sits on its own as an extrinsic social reward. But in either case, there is something that resembles an intrinsic reward loop for a certain period of time, and certainly involves grinding (in a game that otherwise does not involve

grinding). What makes TF2's system especially meaningful is how carefully they've crafted their quest point-gaining system. You gain points most quickly by performing specific play strategies with certain classes, meaning it implicitly encourages you to perfect new play strategies that you may be unfamiliar with and therefore improve your skill at the game.

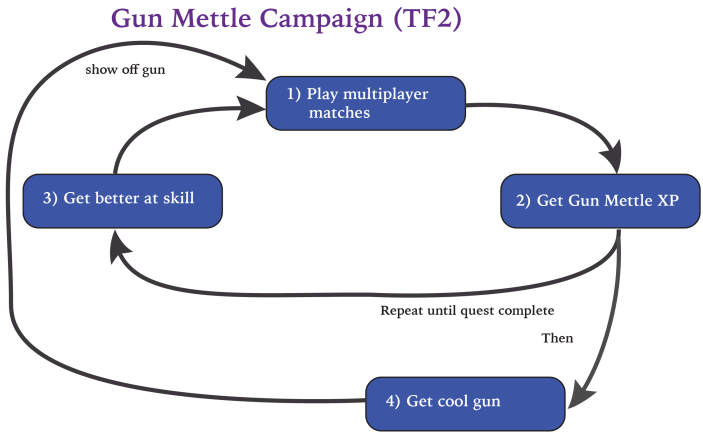

Image 5.5: The reward loop for the Gun Mettle Campaign in Team Fortress 2.

At this point we can jump back to puzzle games with quest structures, most notably Candy Crush.[6] Candy Crush is a "match 3" game mechanic, where you have to switch two candy shaped pieces each turn to match 3 in a row and cause them to disappear, allowing new pieces to appear. Some of the unique features of candy crush within this genre are that you play through levels with a unique shape and unique obstacles, and that you have a limited number of moves to achieve some kind of goal. To compare Candy Crush to Bejeweled,[7] Candy Crush basically adds a questing and grinding structure to what was otherwise an infinite-play puzzle game.

So as the classic example of puzzle-game-with-quests, Candy Crush calls for a grinding-like reward loop. But, it generally falls short of being meaningful. Let's look at the structure. 1) Play puzzles to 2) unlock more levels, and repeat 1) and 2) until eventually you 3) play more novel puzzles with new mechanics. Some differences—notice that 2)

and 3) don't always lead to more difficult puzzles, and I think this is important in making the whole loop itself less meaningful. The puzzles don't structure themselves in order of higher difficulty, they structure themselves in ways that induce frustration and promote pleasure at the discretion of the game designer, rather than the player. A few easy levels, with a tough level, followed by a few easier levels, induces players to feel accomplishment for a while until the inevitable frustration at getting stuck at another hard level. Playing more is also not encouraged, as failure causes one to wait for health to regenerate before being able to keep playing. The only way to get through difficult levels quickly is through paying money, meaning there is an extrinsic input to the system to avoid the frustration of the poorly crafted level structure (poor from a game design standpoint, excellent from a money-making standpoint). In other words, this game does induce grinding, but the grinding does not link into an intrinsic reward loop, but rather can best be relieved by an extrinsic monetary input. Nothing about playing the game allows you to play the game more or better. Although the fundamental mechanics of the game (or the actions taken by players) are meaningful, the game structure is not meaningful.

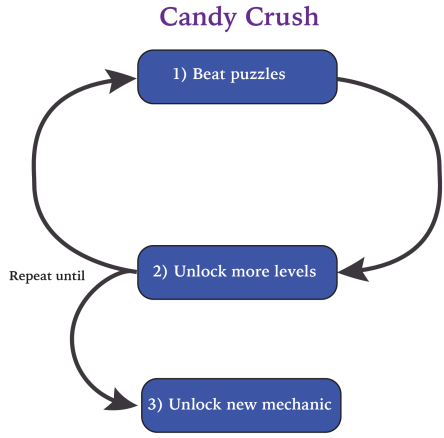

Image 5.6: The reward loop for Candy Crush.

Compare Candy Crush to Tetris[8]—in particular the unlimited play mode that formed the basis of the NES Tetris game. In Tetris, you try to optimally arrange 4 pieces of various shapes into a structure that contains the fewest gaps as they fall down a well. Tetris has incorporated grinding in a nice intrinsic reward loop. 1) Play levels to 2) advance to harder levels. There's also a minor additional loop of 1a) play initial levels better to 1b) start harder levels in a better place to 2) be better at advancing to even harder levels. This minor loop encourages more skilled play, making 1) an even more meaningful interaction. Although I'd argue that based purely on game mechanics Candy Crush is as meaningful an experience as Tetris (or Bejeweled for that matter), I'd also argue that the structure of Candy Crush creates a situation where grinding is used to degrade the play experience whereas in Tetris the game structure causes grinding to enhance the play experience.

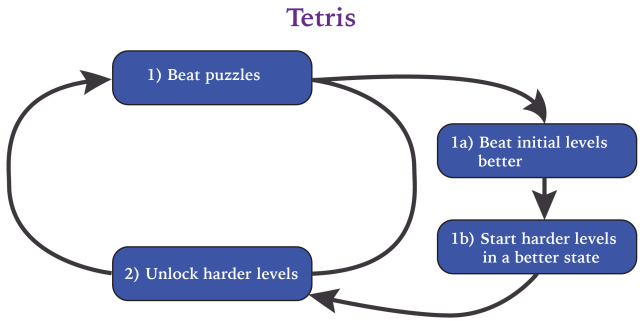

Image 5.7: The reward loop for Tetris.

There's another direction we can take here too, as we explore another genre of games without explicit rewards, but with a strong looping structure. In Tetris or Osmos, there are explicit levels that you progress through as you defeat earlier levels. The same idea holds true in the classic action-adventure games developed by Capcom, for say Mega Man 2.[9] In this game you play through platformer levels to defeat bosses, with the key point that once you defeat a boss, you gain their power. This power makes you stronger, and unlocks a new ability in

your arsenal, which is essentially a kind of reward. This can allow you to fight new bosses in a different way, or to use your powers to access hidden parts of other levels. Tangentially, it's worth noting that the bosses can be played in any order and all are capable of being defeated without any powerups, such that there's an interesting kind of soft-unlocking structure in play. If you struggle with a boss, you can either improve your skill at the game so that you can beat the boss without powerups, or you can defeat other bosses first to gain powerups to make your character powerful enough to defeat the boss.

In the Mega Man 2 reward loop, we really lose step 2 for the same reason puzzle games do—there's no explicit reward that builds up over time, there's no experience points system. There's just discrete rewards gained at key junctures in the game, that unlock later levels, and new abilities that affect how those levels can be played.

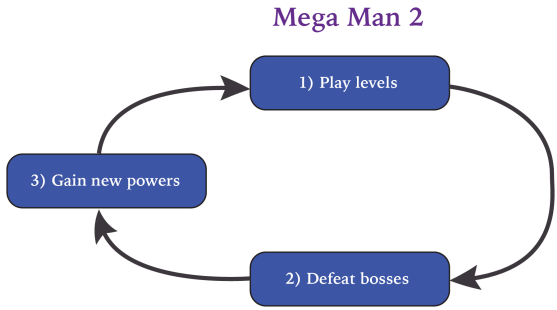

Image 5.8: The reward loop for Mega Man 2.

At this point we can go one step further into the genre of pure adventure games, with Metroid Prime.[10] In this game you are in a giant world that you can explore by moving around. The core mechanic of the game is really to explore the world. Over the course of exploring, you can find new powers or abilities, or you can find keys or other similar items. Whether it is a new power or a key, the main purpose of these abilities is to unlock access to new parts of the world, so that you can explore further, so that you can find even new abilities or keys, so that you can unlock even more parts of the world, etc. And so you enter into a kind of world exploration loop. The funny thing is that the

reward is even less explicit in the pure adventure genre than in the action-adventure genre. You do usually have a map, and the map does grow bigger as you unlock new areas, which is kind of a recognition of a reward in some kind of visible form. And also the ability to enter a new part of the map feels very rewarding when playing the game, even if it's not explicitly called out in a kind of reward.

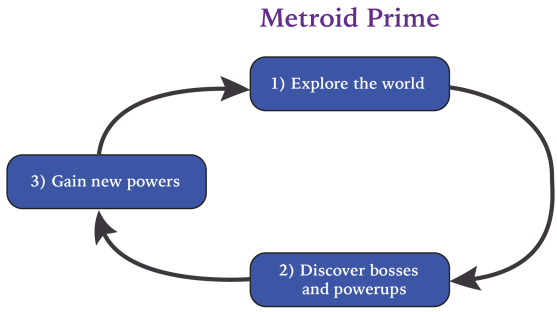

Image 5.9: The reward loop for Metroid Prime.

Now, let's jump back to FF1 and go down a different path. Real-time strategy games, like Warcraft[11] for instance, also feature an interesting reward loop. In Warcraft the goal is to build up an army that can defeat an opponent (human or AI). To get an army, you need resources, which in the original Warcraft was gold and wood. And to get resources, you need a functioning economy—workers that harvest resources and buildings that could process resources. This forms a nice loop—you have workers that you assign to gather resources, to spend resources on more workers and buildings, to be able to gain even more resources (the 1a to 4a loop in the figure). But as you are harvesting your own landscape, there's another larger loop to also traverse, in using resources to build an army to expand to new territories to gain access to even more resources (the loop that starts with 5 and goes through 1b-3b), which in turn allows you to repeat the original worker-economy loop enough times to build an even larger army. Although real-time strategy games are often advertised as games where you lead an army against an opponent, in practice you spend the majority of the time figuring out how to most efficiently manage your economy. Generally,

whoever manages a better economy tends to have a bigger army and usually wins. In this game, you are still directing the actions of individuals to gather or gain access to resources, but most of the meaningful strategy is embedded in the choices of how to spend those resources (buying workers vs. buildings vs. armies), or the reward phase. An intrinsic reward loop still exists and there's still some meaningful choice in the action phase, but much of the meaningful choice has now been shifted to the reward phase.

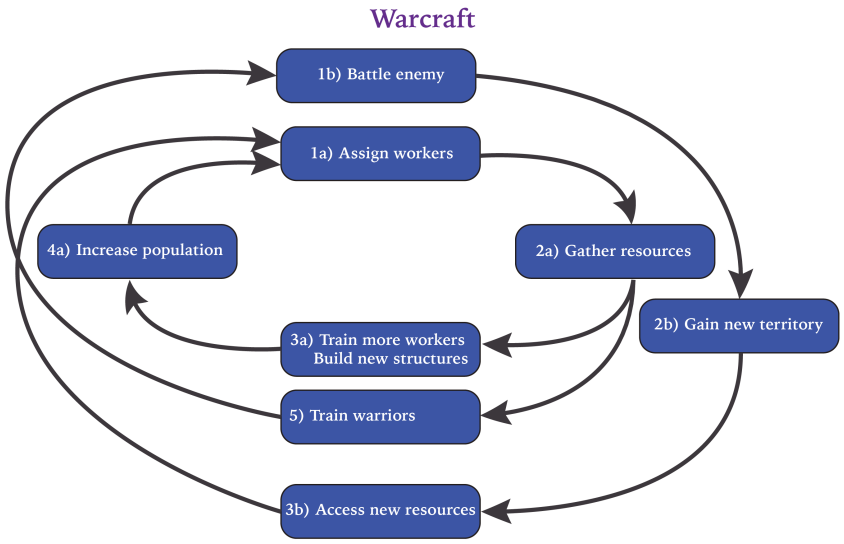

Image 5.10: The reward loop for Warcraft.

If we go further in this direction, we can see the loop diverted even more towards the reward phase. Tower defense games are a genre of games in which you build and manage defenses (usually towers) from advancing enemies. Although the roots of tower defense games go back to the 1980's, they really became popularized as mods (or player-created modifications) to several real-time strategy games in the mid 2000s. In many ways, tower defense games feel like a subgenre of real-time strategy games that have simply automated many of the elements of that genre. And that automation has affected the structure of the intrinsic reward loop.

Rather than dig up one of the original real-time strategy mods, we can analyze the loop for one of the more popular tower defense games developed for mobile devices and still readily available, Kingdom Rush.[12] In a sample level in this game, hordes of enemies advance in waves along preset paths, and you have preset locations where you can place towers that will attack enemies as they go by. As you defeat enemies, you gain gold which can be used to buy new towers and to upgrade your existing towers. This creates a clear intrinsic reward loop (that is fundamental to every tower defense game)—defeat enemies, to gain gold, to buy and upgrade towers, to defeat the even more difficult enemies that will appear in the next wave. In this core loop to the genre, you have no choice over the action phase—the way that your towers shoot at enemies is entirely automated. Your only choice is how to set up the towers, upgrade your towers, and create your overall defensive strategy. It's worth noting that many games (like Kingdom Rush) find ways to additionally insert action choice by embellishing the reward loop. For example in Kingdom Rush, you also control a hero that you can advance around the path to help combat enemies as needed, and who also levels up by defeating enemies in a separate intrinsic reward loop, which does add some agency to the action phase, but this is still a more minor element of the gameplay compared to the strategy of placing and upgrading towers.

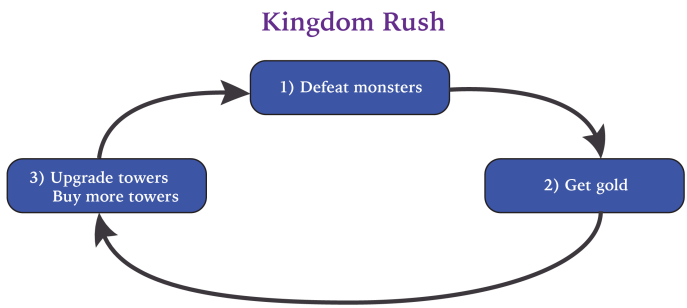

Image 5.11: The reward loop for Kingdom Rush.

We can see this same trend if we go back to the evolution of the RPG genre. The Final Fantasy series has been changing with time and

continuing to embrace its narrative theme with each new title, while simultaneously phasing out grinding. In other words, although you do go up levels in later games in the series, the process of going up levels never really seems to advance or impede your play (thus the designers have effectively removed the soft unlocking structure that most defined the earlier games in the series). They have either balanced the story so well, or have enemy strength respond to your current ability, that you are never stalled from proceeding further in the story because the enemies are too strong. The omnipresent save has also eliminated the idea that there just might be too many consecutive battles for you to survive a dungeon. This then entirely focuses on moving the game through a compelling narrative, while using the battle as a relationship-building exercise between the game characters and the player. The interesting thing is that the reward loop still looks exactly the same as FF1's reward loop, it's just that the action of fighting monsters becomes less meaningful. To compensate, the game designers have created more interesting ways of building up character's abilities, like the Final Fantasy X's[13] sphere grid. In Final Fantasy X in particular, the meaningful interaction of fighting monsters has almost entirely disappeared in favor of more meaningful leveling up and character customization. So 1) becomes less meaningful as 3) becomes more meaningful.

Image 5.12: The reward loop for Final Fantasy X.

This raises an interesting point, in that now we can see a connection between RPGs and god-view simulation games. If we take a game like SimCity 2000,[14] we can also break it down in this way. So here's the loop in SimCity: 1) city grows, to 2) get income, to 3) improve your city infrastructure (zoning, government buildings) to further grow your city. Notice here that there is essentially no fighting-monsters equivalent—step 1) is an entirely passive experience that involves no agency on the part of the player. "Fighting" has essentially been replaced with "waiting." To grind is to wait. But, the use of the currency to level your character has now been greatly expanded and made more meaningful, and by itself creates a fantastic experience. So the gameplay experience is made meaningful with a meaningless grinding action, but a meaningful use of those grinding rewards. The system also recognizes this, and gives you a fast forward button to move the grinding faster when necessary and avoid meaningless gameplay. This is in essence the opposite of the original FF1 recipe, in which the grinding action is the focus, and the choice of rewards was something simply done in service of improving the grinding action.

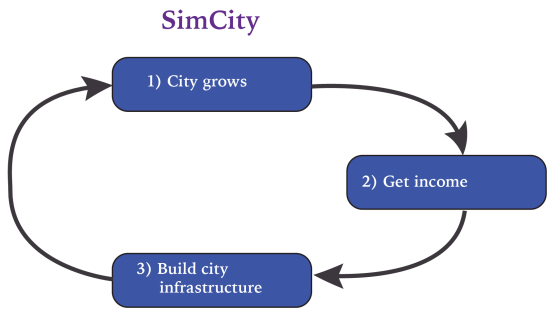

Image 5.13: The reward loop for SimCity.

Of course the issue with such a situation is that we can take it even further, into the realm of the meaningless. Enter Farmville[15] and its genre. Here's the reward loop: 1) crops grow to 2) reap crops to get money to 3) plant more crops, to cycle back through to 1). This basically took the god-sim formula for grinding actions and the FF1 formula for using grinding rewards and smashed them together. Or, they took the

least meaningful element of both recipes and combined them into a new game. By giving you no interesting actions to do with your rewards, they put the emphasis back on the grinding action, but that action is... waiting. Wait, to get stuff, to wait more. Meaningless grinding at its extreme. It still offers you a reward that is intrinsic, and there is a very well-defined intrinsic loop here, but it's a meaningless intrinsic reward because it helps you do a meaningless action better. Of course such a meaningless action can get tiresome, so they offer an extrinsic input (i.e. pay money) to bypass the meaningless activity, but that only gets you rewards that help you do the meaningless activity better.

The issue of meaningfulness will be described further in Chapter 7, but it's worth noting that Farmvillle is a successful game, and that although the reward is meaningless towards the game loop itself, going through the loop does unlock aesthetic customizations and allow one to build out what can be quite a meaningful farm, which is a reward. In the language of our taxonomy, this meaning comes from a sort of external value that this extrinsic reward carries, towards aesthetic customization and being able to express oneself to others through aesthetics that are unlocked through the reward loop. The aesthetic rewards, also a result of going through the loop, do not affect or determine how one goes through the loop, so are effectively an extrinsic reward pinned to an intrinsic reward loop. What's most interesting about Farmvillle is how there is a clear presence of an intrinsic reward loop, and there is a strong element of meaningfulness that has led to the game's success, but that these two items don't really seem connected to each other in an appreciable way.

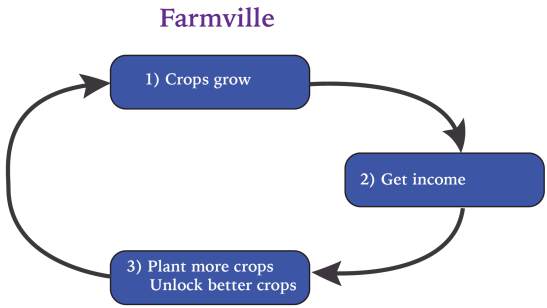

Image 5.14: The reward loop for Farmville.

Farmville reminds me of another reward engine: slot machines. See reward loop: 3) pull lever, to 1) watch wheels spin to 2) sometimes get money to 3) pull the slot machine lever more. The reward for doing an action (should you get the reward) is you get to do more of an action, but that action is not meaningful, and is never harder or more difficult to do. It's almost directly analogous to Farmville.

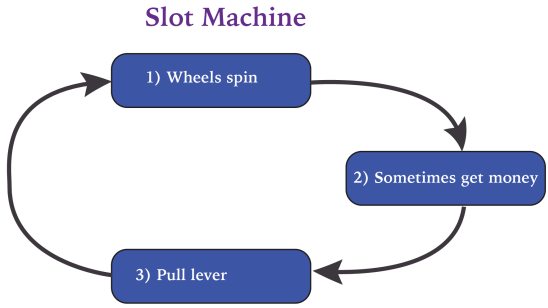

Image 5.15: The "reward loop" for a slot machine.

Which finally brings me to my final example, Fallout Shelter.[16] I was able to neatly categorize each game into a clear grinding structure and level of meaning above, so I wanted to end with a game that defies an easy categorization (as game designers are quite good at defying any attempt to neatly categorize or define games). After playing this game, I'm not quite sure where it falls. On a personal level, I really like grinding heavy RPGs and puzzle games, and also will happily get immersed in a good god-view sim. But I despise the meaninglessness of Candy Crush, Farmville, and slot machine. Part of me is enthralled by Fallout Shelter, but part of me is disgusted at the game and myself for being enthralled.

So let's try to break down Fallout Shelter. This seems to be this basic intrinsic reward loop: 1) let your shelter grow to 2) gain caps to 3) buy more room and upgrade rooms to help your shelter grow faster. There's a relatively complex system by which these steps occur, as growing shelters can mean getting more dwellers or leveling up existing dwellers, and there's a "gain resources to keep people alive to get more resources" sub-loop in the shelter growth model. Plus a questing

system that gives even more caps for practicing certain strategies. But at its basic level, it seems to follow the Farmville/Simcity model, with the core grinding action being "waiting" and most of the meaningful stuff in the upgrading phase. There is a system that needs to be kept in balance and is rather hard to maintain, and your actions help you maintain the system in balance better—in this way it is very similar to SimCity.

So the real question—is the grind in Fallout Shelter meaningful? There are some differences from SimCity that point to Fallout Shelter being less meaningful. Most important, you can't fast-forward in time, but are forced to actually act out your waiting grind. Of course you can rush rooms to fast forward in a small way, but only so often and with increasingly negative consequences the more you do so. So fast forwarding is a mechanic, rather than something free, indicating that it is valuing waiting as a grinding mechanic rather than allowing you to easily bypass it. You can also offer extrinsic inputs to bypass the normal reward system, which certainly points toward a meaningless mechanic. The system is also grossly oversimplified compared to SimCity. Understanding the complexity is part of what makes interacting with SimCity rewards so interesting, and so the lack of complexity would seem to cause Fallout Shelter gameplay to lose meaning.

Fallout Shelter also doesn't have the meaningful elements that are present in other games. Like Final Fantasy X, Fallout Shelter naturally increases difficulty as your shelter becomes bigger, meaning you don't really see your increase in power realized in a significant way, and there's no soft unlocking structure present. And it gives you very little control over how you encounter challenges, whereas that control was part of what gave The World Ends with You gameplay additional meaning.

But Fallout Shelter is also more meaningful than Farmville by comparison. First, the system is more complex, and your choices of how to use rewards most certainly have consequences and need to be chosen well. Most prominently, there is a "loss" condition, and it takes strategic action to avoid that state, which is notably different from Farmville and slot machines.

What's clear is that Fallout Shelter is both more meaningful than Farmville, but less meaningful than SimCity and other games with meaningful intrinsic reward structures. But the real question is: where

is the line between Farmville and SimCity, and what side does Fallout Shelter fall on? When does a game have a meaningful intrinsic reward system and grinding structure? When does a game devolve into the addictive-but-meaningless zone so clearly dominated by Farmville and Candy Crush? In reality, Fallout Shelter is only the first in a series of games intent on blurring this line further, and making it less clear what is a meaningful game and what is a slot-machine-esqe money-maker.

In practice, there probably isn't really a line at all, just a continuous spectrum of meaningfulness. For our purposes here, it is useful to wonder where and how meaning is lost in the process of evolving reward structures to generate new gameplay mechanics. Since we are interested in rewards that exist outside the context of games, it's more interesting for our purposes to know the principles that cause game reward structures to become less meaningful, so that we can extend those principles to other contexts that aren't games. From the analysis above, two things have become clear—meaningfulness is more of a spectrum than a discrete point, and player agency is a key component to when actions feel meaningful. The idea of meaningfulness in reward structures is vitally important, and will be addressed further in the next two chapters.

The chart below lays all of the games described above, showing them on a map that indicates how they relate to each other.

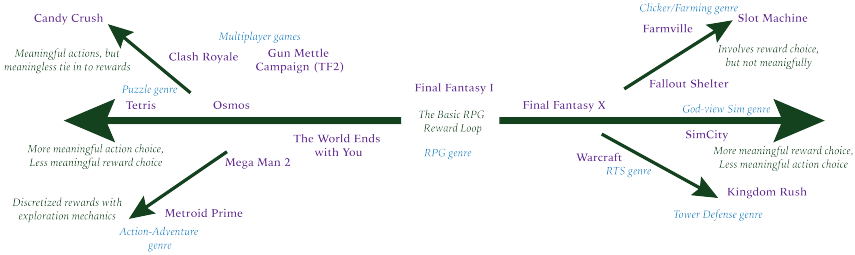

Image 5.16: A visualization of how the intrinsic reward loops of all the games described in this chapter relate to each other. A fundamental spectrum runs along the middle, that relates a tradeoff between meaning action choice and meaningful reward choice. Some side branches are shown, representing interesting diversions from the main spectrum. Games are listed in purple, genres are listed in blue.

6. Gamification

Before going more fully into meaningfulness directly, it's worth taking some time to discuss another example from outside the game industry: gamification. This term generally refers to the application of game-like systems to non-game situations, which feels like it can apply to most everything in this book. In practice, this is mainly used by business and marketing teams to form the basis behind game-like systems used to boost sales of consumer products. The most prevalent example is the "reward points" systems that have become relatively omnipresent in today's consumer market.

So what are these reward points systems? If you have any kind of credit card or membership card to a major retailer, you are probably familiar with this structure, and have probably used it yourself. These are systems that typically reward you points for making purchases, often only with a certain brand, and then give you tangible benefits or rewards once you've gained enough points.

For example, take the airline mileage rewards systems offered by most major airlines. When you fly with airlines, you earn points or "miles." These build up the more you fly, until eventually you can cash in your miles to gain a free flight. Many other systems work essentially the same way, but substitute buying airplane tickets with buying something else.

Since we have taken to concerning ourselves with rewards systems in all places where they occur, it's worth asking, what kind of reward system is this? In particular, this system is inspired by the game industry, which is most known for designing some of the most well-designed intrinsic reward systems. Is gamification potentially one of our most prominent examples of intrinsic rewards that exist outside the game industry?

The Structure of Gamification

So, let's examine the structure of a typical gamification system. We'll

take the airline rewards system as our archetype first, and then talk about a few variations after we've dissected the archetype.

On the surface, this looks very much like an intrinsic reward structure, and clearly harkens back to the game structures it is based on. You perform an action (buy airplane tickets) to gain an immediate reward (airplane miles) which build into a cumulative reward (a free airplane ticket). But, it's a little unclear how the next step would work—it doesn't appear that the cumulative reward results in any kind of effect. And finally, it seems very difficult to argue that any kind of loop exist—it's unclear how the system of airplane miles causes one to practice the action of buying airplane tickets more, or more difficultly, or more completely.

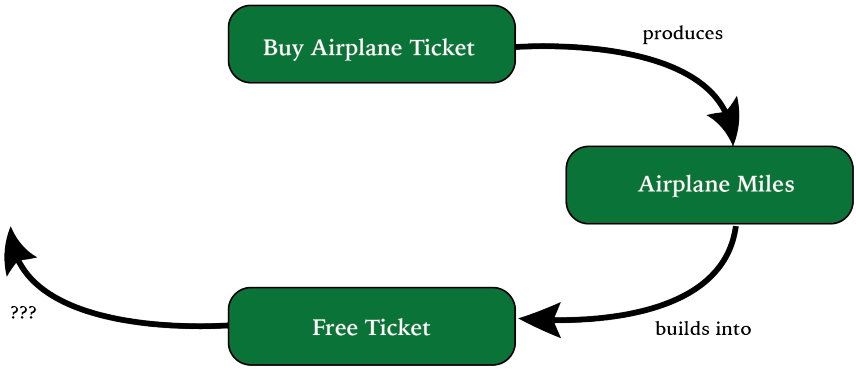

Image 6.1: The attempted loop for an airplane miles system. It's unclear how to connect the loop.

What this means is that superficially this carries the structure of an intrinsic reward loop, but in actuality it lacks the key elements that from a functional standpoint allow it to truly mimic the rewards seen in games. It contains the same initial steps as an intrinsic reward, but then it fails to actually loop.

But one may offer a counter-argument. The reward is intended to create a psychological effect of brand loyalty. Creating an investment in a system can count as buy-in to that system, causing one to become loyal to that brand over others. In this case, the "effect" is brand loyalty,

which does in fact cause one to purchase more, and the loop is completed.

This argument falls apart relatively quickly though. In all of our previous examples, we always discussed the nagging 3 year-old test from the perspective of the player or the user, or in this case the consumer. We never discussed it from the perspective of the designer, or in this case the marketer. For games there is often no need to distinguish these perspectives, as the game designer and player have relatively aligned goals—both are aiming to increase the fun and enjoyment of the player. For a marketer vs. consumer perspective, there are very different and conflicting goals, such that perspective matters. A consumer wants to get the best deal, and a marketer wants you to buy their brand's product at the highest price that they can manage. These can align, but many times they do not, and gamification reward systems seem specifically designed to force alignment where it might not exist naturally.

From this perspective, it's clear that a marketer wants to have the desired effect of brand loyalty, which will cause more purchases of their brand from the customer. But what does the customer think about? Well, for our example, I could at least answer it for myself.

- Why buy airplane tickets? To get miles.
- Why get miles? To get free airplane tickets.
- Why get free tickets? To visit family and friends more often.
- Why visit family and friends? Because it makes me happy.

From a consumer standpoint, this is a clear extrinsic reward. Which is really no surprise—there's nothing intrinsically rewarding about the action of buying something. Purchases are a means to an end (although the act of shopping can be enjoyable, that doesn't mean it's enjoyable for intrinsically rewarding reasons). In some cases, one might argue that the first "why" question given above should actually go right to the extrinsic reward—I might be buying tickets in the first place to visit family and friends, not to get miles. Getting miles is just a side effect of my real purpose in buying the ticket. Or I could be buying the tickets for work-related travel, in which case I'm not choosing to buy a ticket, but I am choosing what brand to buy from. The first question might be better phrased: "Why buy airplane tickets from this airline?

To get airline points." This is then no longer even describing an action itself, but a decision point in the action, which has even less parallel with games. But, this decision point is exactly what the reward points and brand loyalty are specifically aiming to influence.

So at this point we can then draw out a reward loop for airplane miles. We'll draw out two reward loops though, for our two perspectives: the marketer perspective and the consumer perspective.[1] Notice the differences are quite large: one is not even a loop.

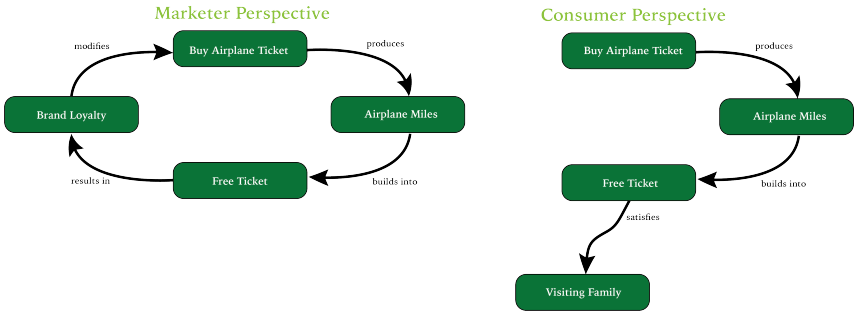

Image 6.2: Two different ways to finish the diagram for airplane miles. One is by taking the marketer perspective, and one by taking the consumer perspective.

There is one additional difference worth noting on the issue of brand loyalty. This concept is quite foreign to a game designer. In one sense, game designers do want you to play their game more, and do enable this through enjoyable game mechanics, and sometimes through their game rewards structures. Especially for games that rely on a freemium structure, increased play is especially important to their business model. But even in freemium games, what the game designers aim for is frequent engagement, not brand loyalty. I don't think any game designer expects to, or specifically aims to, make their game the only game you play for the rest of your life. This difference fundamentally makes gamification quite different from the types of rewards structures found in games.

In a sense, this leaves us with a conundrum. Gamification is described by those who use it as a system inspired by games. But it's really a source of inspiration for something that superficially mimics rewards

in the game industry, rather than something which truly reuses the core elements of game reward structures for another purpose. Game designers often turn red at the mention of gamification, and for good reason—it seems to have little to do with the kinds of systems that they design on a daily basis. Game designers for the most part aim for deep meaningful intrinsic reward structures—gamification at its core is a pretty obvious extrinsic reward that is superficially dressed in an intrinsic reward costume.

Characterizing Gamification

So gamification is an extrinsic reward: where else does it fall on our taxonomy?

Using our archetypical example, gamification is a completion-based reward with a strong external value that acts as a form of currency—you gain the free object once getting enough points. It's also usually an accumulated discrete reward, in that there are a large number of points that are accumulated. The reward structure itself is machine-graded and quantitative. Although you do buy something from a person (usually) at a register, the accumulation of points to your account from the purchase is something that happens automatically, and by a clear quantitatively-defined proportion to the amount you spend. It's also fully abundant—there's no limit to the amount that you can spend to gain more points. These are the main features that seem to hold true across different systems.

Variations in implementation are usually reflected in other elements of the taxonomy. For example, different systems have expected vs. unexpected rewards. Airplane miles tend to be relatively expected rewards—you know how much the points you've accumulated translate into dollars, and exactly what you have to buy to get points. On the other hand, many gamification systems add an unexpected element. For example, Sephora has a points system that can be cashed in for one of several free objects on display near the cash register. Each object is worth a different number of points. But also, the stores seem to rotate the objects in the case on a regular basis. So any given time you enter the store, you expect there to be some objects you can

cash in points for, but you don't know which objects specifically will be offered.

Another variation is on the idea of membership, particularly to a VIP level. For example with American Airlines, there you can gain certain Elite member status that confers additional benefits, when you gain a certain number of airline miles in a year. This effectively turns the system into a hybrid level/discrete reward structure. You use accumulated discrete rewards (points) to reach a certain level (membership), which interestingly is sometimes a temporary reward that can be lost or gained from year to year.

Lessons Learned

I'd like to end this chapter with a lesson or two. I came down rather hard on the idea of gamification, noting that it was only superficially similar to an intrinsic reward system, and that the idea of brand loyalty had little parallels to the gaming world. But I'd like to end by noting that gamification systems are highly effective at achieving the goals that they aim to achieve. Marketing teams are trying to get consumers to buy more of their product—and an extrinsic reward system that creates a sense of investment and buy-in is an effective way to do so. I noted in Chapter 3 that there is no one best reward structure, but that different structures are best for different purposes.

My major objection comes down to the term itself, "gamification," which by its phrasing implies a strong connection to how game designers use rewards in games. It's to this point that I would wholeheartedly object—gamification is not similar to the systems that game designers make. This point is important because words do matter. I know that not everyone plays games, and without playing games, or even better, designing a game, your experience with how games use rewards will be second-hand. On the other hand, because of their omnipresence, almost everyone has interacted with "gamification" systems, and there's a danger in thinking that the rewards expressed in such systems meaningfully mimic the rewards expressed in games. And that implication can create a misunderstanding about what is one of the most powerful inventions of the game industry, the intrinsic

reward loop. Although I will argue in this book that schools should adopt more intrinsic reward structures, that does not mean that schools should adopt anything that looks remotely like gamification systems.

As Ryan and Deci[2] have put it: "In fact what is important and promising about gamification is not the idea of making everything look like a game, but rather the application of principles that make games fun in the design of other activities." The issue with most gamification systems is that they fail this test.

7. Meaningful Rewards

So now that we've seen many iterations of rewards we can take a step back and ask the question—are all of the rewards we've seen meaningful? In the last few chapters, we started to use the term "meaningful" to describe different rewards in games and in gamification systems, though we didn't get into the nuances of why certain actions or rewards were more or less meaningful than others.

Why we do things

Why do we do anything that we do in our lives? What gives any action that we do meaning? One attempt to answer this question goes by the name Self-Determination Theory.[1] This theory attempts to explain the forces that drive us to do actions, or at least those drives that exist from within our own selves, or are self-determined. Of course we take many actions which are determined by others—because we want to avoid punishment, or simply because a superior instructed us to do something. Self-determined actions are not things we do for some other purpose, but are simply things that we like to do for their own sake. When we do something for it's own sake, I would argue that action has meaning. If we use our nagging 3 year-old process, at some point, for any meaningful action we will stop asking why because we reached one of the reasons underlying Self-Determination Theory.

There are three basic psychological needs that form the basis of Self-Determination Theory, or three main reasons that explain why we do actions for their own sake. These three needs are:

1 Competence

Competence is the desire to experience mastery over tasks that are important to you, or basically to experience mastery over tasks that are important to you, or feel like you are in control of the outcome of your

actions. If you have a desire to do something well, then you are able to do it well.

2 Agency

Agency is the desire to feel like you are in charge of your own actions or in control of our own behavior, or to "be causal agents of one's own life." Agency is about being able to make choices, and to feel like those choices matter.

3 Relatedness

Relatedness is the desire to "to interact, be connected to, and experience caring for others." It's about connecting with other people, and feeling like a valuable member of a community.

Relating Self-Determination Theory to meaningful rewards

With this framework, we can finally start to ask what gives intrinsic and extrinsic rewards their meaning. Let's start with extrinsic rewards, as they are a bit simpler.

The Meaningfulness of Extrinsic Rewards

If we go back to the nagging 3 year-old test, an extrinsic reward leads us to a "just because" statement as our final answer. This final answer more often than not leads to one of the three pillars of Self-Determination Theory, and what we would therefore call a meaningful extrinsic reward. Or, the final answer is about a materialistic value or something that really doesn't have a value that we can articulate, leading to what we would call a meaningless extrinsic reward. In reality, it's rarely a binary option between meaningful and meaningless, but more of a spectrum.

The classic extrinsic reward has a simple structure: Do an action to get some reward which has some value. The trick in determining meaningfulness is figuring out if the value of the reward satisfies one or more of our three Self-Determination Theory needs.

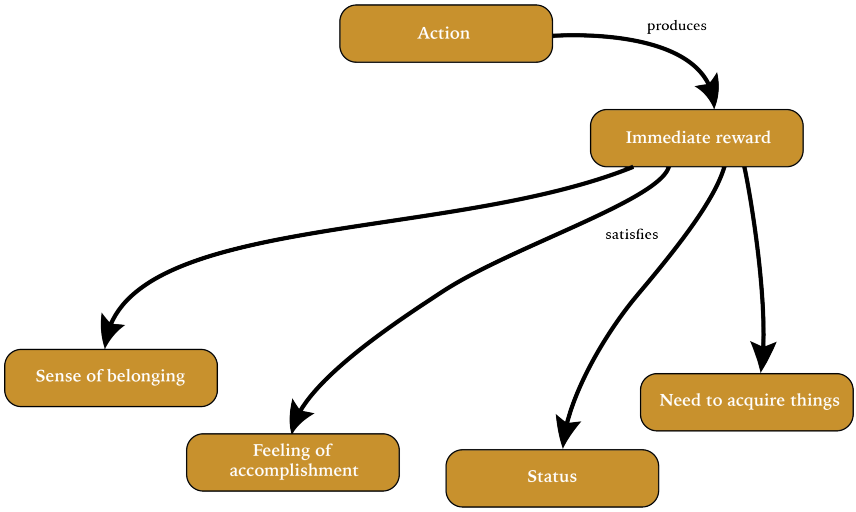

Image 7.1: For reference, the same visualization of the extrinsic reward diagram from Chapter 1.

Competence in Extrinsic Rewards

A good starting place to discuss competence is the example in Chapter 1 of photography. In our aspiring photographer analogy, perhaps this hobbyist posts photographs to Facebook (or more likely Instagram), because they want to be known as someone who is good at capturing the essence of a party, like let's say a wedding. In this case, the "likes" are a sign to them that they are in fact good at, or competent at, that goal. Rewards can be a powerful way to offer feedback that competence has been achieved.

Competence can also be achieved through many different kinds of rewards. Trophies and medals are clear examples of competence-inducing extrinsic rewards. In general, measurement-based rewards or completion rewards that are performance-contingent help satisfy

needs for competence, since acquiring the reward is dependent on demonstrating competence. In this way, any reward that by itself indicates some level of mastery, or rewards that give feedback that help someone understand how to become more competent both help satisfy needs for competence. Which covers quite a large range of extrinsic rewards, including things from grades to classic game achievements.

Agency in Extrinsic Rewards

Agency is a little trickier to pin down. Some rewards can get down to feeling outright manipulative. The more manipulative a reward feels, the more we feel as though our actions are being controlled or coerced by others, and the less agency we feel like we have. The more the reward feels like a feedback mechanism, the less manipulative it also tends to feel. There are no strict rules on when this does and does not happen, but there are certainly some elements of our taxonomy that are more prone to feel manipulative. To summarize some key points from Chapter 3 and to draw on related research from the Self-Determination Theory authors[2]:

- Measurement rewards contain feedback in their measurement itself, which is contained in the reward. This gives them a greater natural potential to act like feedback than completion-based rewards, though completion-based rewards that are performance-contingent in meaningful ways can offer feedback too. When there are high stakes attached to measurement-based rewards, which essentially means only high scores are deemed acceptable, then the measurement feels less like feedback and more like a judgmental evaluation. Other factors greatly influence whether this dimension acts as feedback.
- Hard unlocking structures are more forcibly created by the designer of the structure, and can feel as if they are guiding the player in a certain direction. Soft unlocking still exists in a designed environment, but it is an environment that allows for more choice, and feels less constrained and manipulative.
- Qualitative systems that are evaluated by people can always have

a tendency to feel manipulative and judgmental, since there is a person deciding your fate and reward.

- Unexpected rewards have no potential to limit agency, since they come as an unexpected surprise. Expected rewards have the potential to feel manipulative, depending on how strongly the endpoints are thrust in the face of players and the players are pushed in the direction of that reward.
- Temporary rewards are the type of rewards that feel most like feedback, and have little potential to feel manipulative. Permanent rewards have the potential to feel manipulative, depending on how much the permanence of the rewards is attempting to pin people into categories.
- When rewards act as in-game currency, this has the potential to either limit or improve agency. On the one hand, currency directly offers more control and agency in allowing players the ability to decide how best to use their reward. On the other hand, when currency is overemphasized in the context of the game, and there are limited choices for what the currency can be used for, the currency can feel very manipulative and be one of the biggest drivers towards meaningless rewards.

There's also a few perspectives on meaningful agency that go beyond the taxonomy. The "Action phase" rubric described in the "Agency in Intrinsic Rewards" section later in this chapter can also provide guidance on when extrinsic rewards are meaningfully incorporating agency (the "Reward phase" rubric in that section is only relevant to intrinsic rewards).

Additionally, the number of rewards available in the system for a person to choose from is potentially one of the biggest factors affecting agency. This is a different concept from scarcity, which is related to how many people can potentially get a single reward. In contrast, this is related to how many rewards a single person can potentially get. This is really a system design issue, of how many rewards exist in the overall system. From personal experience playing games with different kinds of rewards, a good rule of thumb is to have two to three times as many rewards in the system as can be achieved in a single course of normal play. Building this tends to result in a system that has a decent amount of agency. The real question is then really, what is "normal

play"? For many games, this often means playing through the game to finish one time. Thus, to get all the rewards, one would have to play the game through to finish several times, but also someone who only played the game to finish once would still have a satisfying experience with a choice-filled reward structure. For non-game environments, or games without a clearly defined end state, it may take a little more creativity to define "normal play." Some options can include what the average person might typically spend on the platform, or what might be deemed by the designer to be a minimally satisfying experience (a play off of the idea of a minimum viable product, this can mean the minimum someone would need to have engaged to have felt satisfied that they experienced something).

Finally, a discussion of agency should also discuss self-selected goals. Allowing the ability for someone to create their own goals is one of the ultimate expressions of agency. To link back to the previous paragraph, when there are two or three times as many rewards for someone to choose from than they can possibly achieve in one playthrough, they will naturally need to self-select which rewards they want to achieve. But to go a bit further, nothing builds a sense of agency more than a reward that you yourself have designed and chosen to accomplish. Of course, the design of an experience can make these self-selected goals more or less available to a person. Sometimes, self-selected goals are a scaffolded experience build right into the design itself (for instance many fitness games like Wii Fit[3] set workout or weight loss goals as a part of the play experience). But many times the goals are not explicitly built in, yet players still readily engage in self-selecting goals. These self-selected goals are almost always extrinsic rewards, but the system the goals are set upon are not always fully extrinsic. Two relatively common examples of self-selected goals in games are speedruns and personal bests.

Speedrunning is attempting to beat a game from start to finish in as little time as possible. This typically occurs in action-adventure games, or games with a clear end state. Players will try to beat a game as fast as possible, often recording and posting online their playthrough as evidence of their success.[4] It's interesting that the designed reward built into these games is just to specifically get to the end of the game, the attempt to beat the game as fast as possible is a self-chosen goal. The games themselves might or might not have intrinsic rewards

structures built into the game. The concept of speedrunning has become so engrained that some games now offer achievements that explicitly challenge players to speedrun (such as the "Speedrun 1" or "Speedrun 2" rewards in the Action Adventure game Hollow Knight[5]).

The second idea is a natural one that occurs in many arcade games with endless (or just very difficult) intrinsic reward loops. In these games, since there is no end state the goal in the game is to simply get as far as you can. It's natural as a player to set your own goals to beat personal bests. In a game like Pacman,[6] for instance, this can mean getting to a specific score threshold (thus engaging with the extrinsic reward system in the game) or getting to a specific level (thus engaging with the intrinsic reward system in the game). In either case, achieving a certain personal best, or beating a previous personal best, is a self-selected extrinsic reward with a lot of meaningful agency.

On one final point, let's briefly return to Chapter 6 and gamification. We noted that gamification systems have competing purposes—the purpose of the system for a consumer are different from the purposes intended by the designer. One way of saying this is that the designer of gamification systems is trying to control or manipulate the consumer into certain courses of action, or that the design of such a system is meant to limit the agency of the consumer. As a general rule of thumb, any reward system where the goals of the designer of the system don't match the goals of the user of the system is likely going to suffer an agency problem, and lack meaningfulness for the user.

Relatedness in Extrinsic Rewards

The examples given in Chapter 1 are a great starting place for talking about relatedness. Taking a picture of my family to have a picture on my desk seems to be a clear example of satisfying a relatedness need. Feeling like a valuable member of my family, being reminded of my membership with my family while at work, and being able to show my closeness with my family to colleagues all helps me fulfill my need for relatedness. For another example—I post a picture on online photo sharing websites so that I can get "Likes" so that I am bringing happiness and joy to friends. Again, another clear example of an extrinsic reward bringing about the value of relatedness.

In general, relatedness is best achieved with community-given rewards (a subset of human-evaluated rewards), since a reward that is conferred by a peer is a process that by itself helps create relatedness. But as shown in the photo-on-the-desk example, there are lots of objects that can form connections between people and bring them together. For relatedness to be satisfied, the reward needs to be easily viewable by other people, meaning it usually has to be permanent and easily shareable. The Farmville[7] example discussed in Chapter 5 is an interesting example here—creating an aesthetically pleasing farm that expresses who you are and can be shared with friends on Facebook is a good example of a reward that achieves relatedness.

The Meaningfulness of Intrinsic Rewards

Intrinsic rewards are a bit more complex. Since the reward itself is an action, then the action itself must be used to determine whether or not the reward is intrinsic. Actions are an exertion of agency—you can only take an action if you have choices to make and the agency to choose amongst those choices equally. Therefore the issue of meaning in intrinsic rewards is most closely tied to agency, but there are useful insights related to all three pillars.

Competence in Intrinsic Rewards

Competence is clearly embedded into the action of an intrinsic reward loop. As mentioned in Chapter 1, hobbies are one of the basic activities that involve an intrinsic reward loop. And hobbies are often developed because we like to feel like we are getting better at things we do. This feeling of increased competence in your actions through repeated intrinsic reward loops can help reinforce intrinsic motivation.

It's worth mentioning that for this to be true, the action being performed has to be one in which you can actually express competence, and one in which you care enough about to be competent at. This may feel obvious, but it needs to be stated. Many games achieve this naturally—the core mechanics of a game need to be something that you actually can be good at and that you care

to be good at; therefore you can enjoy being good at it. But as we stretch gamification, and games themselves, into new territories, we can enter situations where the action being performed is not one that has different levels of competence or is not one that a normal person would care to be better at. This was what some of the games in Chapter 5 bled into. For example—in Farmville, the action of "waiting" is not one with a high degree of competence associated with it. Sure, patience is a virtue, but I don't think anyone considers practicing their ability to "wait out time better" to be something that they care to be more competent at.

There are certain dimensions of our taxonomy that can play to this feeling of competence especially strongly. In particular, permanent leveled rewards that allow a player to progress upwards over time often work well at showing an increasing progression of competence. Or, when rewards are embedded in an unlocking structure, that can also help reinforce and communicate to the player how the repeated actions result in greater competence and therefore more unlocks. This holds most true in a soft unlocking structure, but can also be reinforced in a well-designed hard unlocking structure.

Agency in Intrinsic Rewards

Agency is the most complex concept to talk about in the context of intrinsic rewards. Agency is not simply the ability to choose—it's the ability to have what might be described as meaningful choices. Sid Meier has famously defined a game as being something which has "a series of interesting decisions."[8] Thus this idea of interesting, or meaningful, choice is tied not only to game loops or rewards, but to the entire idea of what makes a good game itself. Since an intrinsic reward has as its reward doing more of the action itself, there is an implicit connection in the reward structure to the idea of agency, and how that agency relates to the core actions, or mechanics, of a game.

As an example, let's jump outside a game environment for a counterexample of meaningless agency. Let's take a multiple choice standardized test. This isn't a context that typically contains a lot of choice or agency—you can choose what answer to give for each

question, and sometimes you can also choose what order to answer the questions. But that's about it.

So let's say we wanted to offer some more intrinsic motivation for taking this test. Maybe the test designers allow you to make an avatar before you take the test. You can choose the look of the avatar, maybe the style of hat and color of clothes that it wears. There can be a whole bunch of choices allowed in the selection of the avatar, and those choices could even be related to some kind of reward structure (do better in the task, to unlock more avatar customization options). And once the avatar is chosen, it appears on the top corner of every screen of that test.

So, there's agency here right? That might be true, but it would be hard to call this form of agency meaningful. The avatar seems totally separate from and unrelated to the core component of the test, which is answering questions. In fact, it's unclear how answering a test question should be related to the appearance of an avatar, thus it's unclear how these choices could ever be made meaningful for this kind of task (though I'm sure any game designers reading this book might immediately be thinking about creative ways to prove that statement wrong).

At which point, one might ask the question—is having choice over an avatar always a meaningless choice? There are so many games in which the rewards systems are based around the cosmetics of a character. And many players spend much time modifying their avatar's look. Is this a meaningless use of time?

Well, not necessarily. In many cases, the avatar is shared publicly online, either in the game itself, or in the "lobbies" of the game, which affects how you are viewed by other player's (potentially pointing to satisfying a relatedness need). The look of your character could affect gameplay, either directly or indirectly. For example, in Mario Kart,[9] you choose an avatar at the beginning of each round, and each avatar has different stats that affect the acceleration and turning radius of your cart, which directly affects gameplay. Or, in a multiplayer first person shooter game, you might specifically choose an outfit that better blends in with the surroundings and therefore camouflages your character from others, which indirectly affects the gameplay. Both direct and indirect effects on gameplay are indicators of a meaningful form of agency in avatar cosmetics.

Thus it is clear that our avatar in a multiple choice test is simply a poor fit of agency to core mechanics. There were really two problems here. First, this was a case in which the core of the activity did not have a lot of meaningful agency, and so layering on an avatar creation system with or without some kind of reward structure attached to the cosmetics really has limited ability to amplify or embellish that agency. Second, the use of an avatar has little to no connection to the task of answering multiple choice questions, which has no use for an avatar. This is in contrast to the Mario Kart or first person shooter examples, in which the core mechanic of the game involves controlling an avatar.

The larger point to be made from this example is that in discussing meaningfulness in agency in intrinsic rewards, it becomes unavoidable to discuss and evaluate the core mechanics themselves, and how the reward system is related to the core mechanics. So, let's get into game design theory a bit.

From our analysis of games in Chapter 5, we found two key components in the reward loop where agency can be inserted: in the Action phase (which is highly related to the core mechanic of the game) and the Cumulative Reward phase (or more specifically, the phase that translates the cumulative reward into an effect). It's possible for games to reduce agency, and therefore meaning, in one of these phases, as long as the other phase retains meaningful agency. For example, SimCity reduced all Action meaning in favor of only having Cumulative Reward meaning. Correspondingly, puzzle games like Tetris reduced all Cumulative Reward meaning in favor of Action meaning. It's only when you get to a place where meaning is lost in both places that the intrinsic reward becomes truly meaningless, such as in Farmville.

So as an example, let's ask a question—is going to dinner a situation that has agency? Well, that depends. Are you able to decide what you can wear to dinner, or is the dress code set? Can you decide what to order, or is it a preset menu? Or on a broader level, can you even decide where to go to dinner? Or even if you want to go to dinner? There are multiple levels at which agency can be allowed or limited, and these greatly constrain the extent to which the decision allows for a meaningful amount of agency.

Which brings us to a framework offered by Scot Osterweil: the Four Freedoms of Play.[10] These freedoms are meant to describe how

elements of playfulness are encouraged or facilitated by an experience—if you don't have the freedom to do these elements, then something is lost in the playfulness of the experience. These four freedoms are: 1) Freedom to Experiment, 2) Freedom to Fail, 3) Freedom to Try on identities, and 4) Freedom of Effort.

Each freedom describes the ability to have choice over your decisions in some way. These four freedoms can act as a good guide in determining what choices are meaningful. You can ask yourself for a given choice, for example our dinner party:

Does it allow the freedom to experiment with options?

- *Can I choose what to order, how to combine or substitute ingredients, in my dinner?*
- *Can I choose where to go to dinner?*

Does it allow the freedom to fail?

- *Can I order something I know that I won't like?*
- *Can I order something I can't afford?*

Does it allow the freedom to try on new identities?

- *Am I allowed to order something I don't usually order?*
- *Am I allowed to dress and act as I want?*

Does it allow the freedom of effort?

- *Can I choose to go to dinner or not to go to dinner?*

These four criteria offer an excellent guide to determining how much your intrinsic reward loop is meaningful. But, we noted that the choices can be embedded at either the Action phase, or the Cumulative reward phase. Accordingly, we can develop two sets of rubrics based on this framework, for each phase. If you can pass these rubrics, then you have meaningful agency embedded in that phase of your intrinsic reward loop. It's worth noting that these are not all or nothing propositions—you can pass three of the freedoms, but fail at one of

them. Or you can even partially pass all of the freedoms. These are spectrums more than binary distinctions.

In the Action phase:

Freedom to Experiment

- *Can you act in many different ways?*
- *Is there more than one way to find a solution to a problem you might be facing?*
- *Can you go down worthless or suboptimal paths?*

Freedom to Fail

- *Most basically, can you lose?*
- *Is it psychologically safe to fail?*
- *Can you set the difficulty of your actions?*

Freedom to Try on Identities

- *Can you perform your action from different perspectives?*

Freedom of Effort

- *Can you decide not to act?*
- *Can you decide how much you act or hold back?*
- *Can you decide when to act?*

In the Cumulative Reward phase:

Freedom to Experiment

- *Can you spend your rewards in ways that develop different skillsets?*

Freedom to Fail

- *Can you choose to spend your reward in a way that will lead to failure, or at least to more difficulty?*
- *Does gaining the reward maintain the ability to fail?*

Freedom to Try on Identities

- *Can you spend your reward in ways that better express an identity you want to try out?*

Freedom of Effort

- *Can you choose not to spend your reward, and remain as you are?*

It's quite easy to apply these frameworks to some of the game examples described in Chapter 5 and see how those games lack meaningfulness. For example, let's fill out this matrix for the Meaningfulness of Actions in two contrasting examples: SimCity[11] (which was described as having low Action agency, but high Reward agency) and Tetris[12] (which is the opposite—high Action agency, low Reward agency). First let's go through the rubric for SimCity:

ACTION (sims performing actions)

Freedom to Experiment

Can you act in many different ways?

- In general, no—the "actions" are completely automated in that the sims perform the actions that make you money. You can choose a few tangential actions regarding taxes and policies.

Is there more than one way to find a solution to a problem you might be facing?

- In general, yes—a crowded intersection, for instance, can be solved in multiple ways.

Can you go down worthless or suboptimal paths?

- Not really—the sim's algorithm always tries to choose the most optimal path for each sim.

Freedom To Fail

Most basically, can you lose?

- Not really—on an individual action or meta-level. You can go bankrupt, but you don't really lose.

Is it psychologically safe to fail?

- A bit irrelevant since it's not possible to really fail.

Can you set the difficulty of your actions?

- You can choose a more difficult initial map, so to some degree, yes. But it's an initial choice removed from the turn by turn action.

Freedom to Try on Identities

Can you perform your action from different perspectives?

- Not really—while you are waiting for action, there's little room to exert multiple identities.

Freedom of Effort

Can you decide not to act?

- Technically, yes, you can choose not to click on the screen. But the game does go on without you. You can't choose to stop the city from growing—you are forced to always take the fundamental actions in the game, because they always occur automatically, without you.

Can you decide how much you act or hold back?

- Yes, you can decide how much to support your city with an intervention or policy, or let it keep growing as is.

Can you decide when to act?

- You can choose when to jump in and make interventions—but you can't choose when to make your city grow. It grows with or without you.

REWARDS (spending money to build city infrastructure)

Freedom to Experiment

Can you spend your rewards in ways that develop different skillsets?

- Absolutely. You can grow your city however you want to.

Freedom To Fail

Can you choose to spend your reward in a way that will lead to failure, or at least to more difficulty?

- Definitely—you can mismanage your city and stunt its growth with poor practices.

Does gaining the reward maintain the ability to fail?

- Yes—you can choose how to spend the reward in a way that mismanages your city.

Freedom to Try on Identities

Can you spend your reward in ways that better express an identity you want to try out?

- Yes—you can definitely take on roles and choose what kind of mayor you'd like to be.

Freedom of Effort

Can you choose not to spend your reward, and remain as you are?

- Yes—you can stockpile as much money as you'd like as a greedy mayor!

Next let's go through the rubric for Tetris:

ACTION (choosing where to place bricks)

Freedom to Experiment

Can you act in many different ways?

- Yes—blocks always fall down, but you choose what position and where to place blocks.

Is there more than one way to find a solution to a problem you might be facing?

- Yes—there are many ways to place blocks and advance up levels.

Can you go down worthless or suboptimal paths?

- Yes—stack blocks up any weird way you want! You might lose, but you have the choice to try it.

Freedom To Fail

Most basically, can you lose?

- Yes—if the blocks get to the top of the screen, you will lose. You also have agency over when you lose.

Is it psychologically safe to fail?

- Yes—it's easy to restart, failure is expected at some point and treated as a natural part of the game experience. There is in fact no way to "win" at Tetris, really you can only fail, the game is just based around challenging how far you can get before you fail. This is an ideal model for psychological safety around failure.

Can you set the difficulty of your actions?

- Sort of—there is a difficulty setting in Tetris when you start that allows you to begin play with a more advanced level. But for the most part, difficulty is set automatically by the reward loop.

Freedom to Try on Identities

Can you perform your action from different perspectives?

- Not exactly—there isn't a strong sense of identity in this game based on geometric pieces.

Freedom of Effort

Can you decide not to act?

- Sort of—you can choose not to act and the game will go on and stack up blocks in a tower until you lose. It is a time based game, so the decision not to act is a decision to fail.

Can you decide how much you act or hold back?

- Yes—you can choose how much to move or rotate pieces. For example, you could try playing the game without rotating a single piece.

Can you decide when to act?

- No really—as a time-based game, once it starts, it is going with or without you.

REWARDS (getting 10 lines and going up a level)

Freedom to Experiment

Can you spend your rewards in ways that develop different skillsets?

- Not really—there's only one skillset to develop, and you can't really choose to advance with that skill with rewards.

Freedom To Fail

Can you choose to spend your reward in a way that will lead to failure, or at least to more difficulty?

- No—you don't have a choice in spending the reward, you always just advance up levels.

Does gaining the reward maintain the ability to fail?

- Yes—you are actually more likely to fail when you gain the reward, since the action grows more difficult.

Freedom to Try on Identities

Can you spend your reward in ways that better express an identity you want to try out?

- No—you don't have a choice in spending the reward, you always just advance up levels.

Freedom of Effort

Can you choose not to spend your reward, and remain as you are?

- Not really—once you've made 10 lines you always advance. You can choose not to make 10 lines, but that's more about Action agency, and is also a choice to lose the game rather than one not to exert effort.

It's quite clear that SimCity had most of its agency in the reward phase, whereas Tetris has most of its agency in the Action phase. Notice

that even the games with high agency in one of these steps were not able to answer yes to every question, but they were able to answer yes to at least some of them.

Relatedness in Intrinsic Rewards

Relatedness is perhaps the most difficult to discuss in the context of intrinsic rewards. Multiplayer games are an interesting place to start looking, since those games implicitly involve an interpersonal connection. Although reward systems have been attached to multiplayer games for hundreds of years (most notably in Chess and Go), those rewards have been primarily extrinsic in nature. The idea of going through intrinsic reward loops while conducting actions in competition with others is something that can be explored more in depth in the future, once such systems continue to be developed (for some early examples of promising directions, see the section on Hearthstone and Splatoon 2 in Chapter 8).

The reward loop itself is something that is very personal to the individual going through the loop, which is one reason Relatedness is so hard to think about in the context of intrinsic rewards. Reflecting on the taxonomy of rewards, there are some dimensions that can lean towards relatedness. The idea of permanence is perhaps most relevant to the idea of Relatedness—permanent rewards displayed in a public place can then allow for public expression of one's abilities in a social group, which can serve to reinforce connectedness in the same way that an extrinsic reward can. It's worth noting though that the relatedness that comes from this sharing outside the reward structure is relatively unrelated to the action that is accomplished in the reward loop itself. In other words, it's not really clear that the needs of relatedness are achieved inside the intrinsic reward loop; it's the sharing of such intrinsic rewards on other platforms that achieves relatedness, which is true for both intrinsic and extrinsic rewards.

As a related point—intrinsic reward loops do exist for individuals acting in cooperation with each other, say as a part of a game guild. If the guild itself has a series of intrinsic rewards it can loop through from well-coordinated cooperative efforts from the individuals, then such loops can certainly reinforce a feeling of relatedness and are perhaps

a primary example of how to achieve relatedness with intrinsic reward structures. In other words cooperative rather than competitive multiplayer games are probably the best venue for achieving relatedness with intrinsic rewards. An example can be the clan leveling system from Clash Royale,[13] where the reward is achieved once the clan as a whole performs well in repeated clan battles over time. In this case, the group itself is really the primary unit going through the reward loop together. This example feels like it can have high relevance for education, in which group work is already a common practice.

In comparing intrinsic and extrinsic rewards, it's interesting to note that the sharpest divergence happened on the issue of agency. In intrinsic rewards, agency was the most natural and well-built out pillar, the one that is easiest to think about promoting. On the other hand, in extrinsic rewards, agency is the pillar most likely to be damaged or destroyed by the design of the rewards. Or, to go back to one of the original seminal research papers[14] on the issue, it's important that rewards not feel like they are manipulative or constraining agency, which seems much easier to design for and think about when you are in the intrinsic rewards headspace. We'll see how some leaders in the education industry have designed in this space in the next chapter.

8. Exemplar Educational Reward Structures

At this point, let's take some time to walk through some real world examples of innovative reward structures that are used in a few different contexts. These represent some best in class examples of how to use reward structures in effective ways. I'm going to walk through three case studies from the education world, and one final case study from the gaming world. I'll use each example to describe what the educator or game designer wanted to achieve and what particular kind of reward system they chose. I'll then offer an analysis of what reward system was chosen and why it worked well for it's given purpose.

Gamestar Mechanic

Gamestar Mechanic[1] is an educational game and game creation platform that teaches students about systems thinking through the process of game design. It starts as a game where you have to fix broken levels to make them playable. But it eventually expands into a student creation platform where you can make your own original games and share them online with other players. Details of the reward system and reasoning behind that system are based on an interview with Scott Price, former Product Manager for Gamestar Mechanic.[2]

There are multiple layers to a rather extensive reward system on Gamestar Mechanic. First are the Gamestar badges, which feel very much like classic extrinsic reward achievements in games. The badges are structured to reward players as they progress through the early stages of the game and reach specific milestones. When those milestones are reached, the game gives out assets that can be used in the in-game level editor. So basically, when you master a skill in the game by beating a level, you get a discrete achievement-based reward to recognize that skill, and you also unlock elements in the level editor that are related to using that skill in the games that you can create.

This added feature, of unlocking elements that you can use in the level editor, turns what otherwise feels like a standard extrinsic game achievement into one that starts to feel more like an intrinsic reward loop. Play through levels, to show that you can master the mechanics of the game, to use the mechanics you have mastered in your own levels, to be a better game designer and player, to be able to play and make even more levels, etc.

Additionally, there is an experience points system, and 4 categories in which you can gain experience points. Those categories are based on the 4 general skills that the designers wanted students to practice: Designer, Player, Reviewer, and Citizen of the sharing community. Experience points can be gained for individual actions that contribute to one of those four categories. Gaining a Gamestar badge will reward a lot of experience points, but those badges are not the only way to gain experience points—everyday actions in the site also contribute points.

There are also mechanic ranks. Your mechanic rank is a leveling system that is publicly displayed on your name. The rank is based on the lowest of the four skills' experience points, so the only way to increase your rank is to gain experience points in a balanced way across all four skills.

Finally, a fourth system was developed after the initial release, to develop badges that could be passed to other systems and recognized outside the Gamestar Mechanic ecosystem. These were called World Badges. They were designed in parallel to all of the other systems described above, attempting to recognize elements of Gamestar Mechanic play experience that might contain value outside the system, or be based on "transferable skills." To help increase the validity of these badges, a rubric and human review element was added, as well as a place for student reflection.

So, this four part system clearly contains a wealth of rewards. Let's first describe how this reward system maps onto our taxonomy.

First, there is a giant hybrid structure that involves the first three rewards. There are the Gamestar badge achievements, that lead into an experience point system, which leads into a ranking system. This is a discreet-to-accumulated-discrete-to-leveled reward system. Each element builds upon each other, and the value of each reward can only be identified by how it ties into this larger hybrid system.

The Gamestar badges are intentionally designed in the classic

formula for achievements—as discrete, completion-based awards that are machine-graded, quantitative, abundant, and permanent. They are also relatively expected, and for the most part occur over the normal course of playing the game, but some are unexpected and first recognized as they are unlocked. These rewards have the semblance of an intrinsic reward loop—play the game to gain achievements to unlock elements to play the game better. It's in a hard unlocking structure, with the designers explicitly using the unlocks to scaffold the student learning. In a way, this gives the reward an internal value, in terms of building up a player's palette of game design actions.

The experience system is an accumulated discrete reward, like many other experience point systems. It seems to be given out with a mix of expected and unexpected rewards—some of the Gamestar badges are pointed out in advance and are very much expected. But some of the more minor elements could come semi-unexpectedly. I say semi because they were not laid out in advance for players, but also over time players would recognize and expect certain actions to produce experience points for one of the four categories. These are also all given out in a completion-based manner—if the action is completed, the full rewards are given out; there is not a case where the reward gives partial points for completing a level poorly. This system does not seem to contain any direct internal or external value, although there is an indirect internal value in how it feeds the rank system. It is machine-graded, quantitative, abundant and permanent.

The rank system is a leveled reward. It's interesting that it is a combination of the four different experience point systems, and only acknowledges the value of the lowest of the four skills. The main feature of this badge is that it has external value (unless the community in which your Gamestar profile is shared is still considered within the game). The primary point of this system seems to be to take the elements of the hierarchy below it, and translate them into a simple metric that can be shared and easily convey status in the community. The rank system is machine-graded, quantitative, permanent, abundant, and very expected.

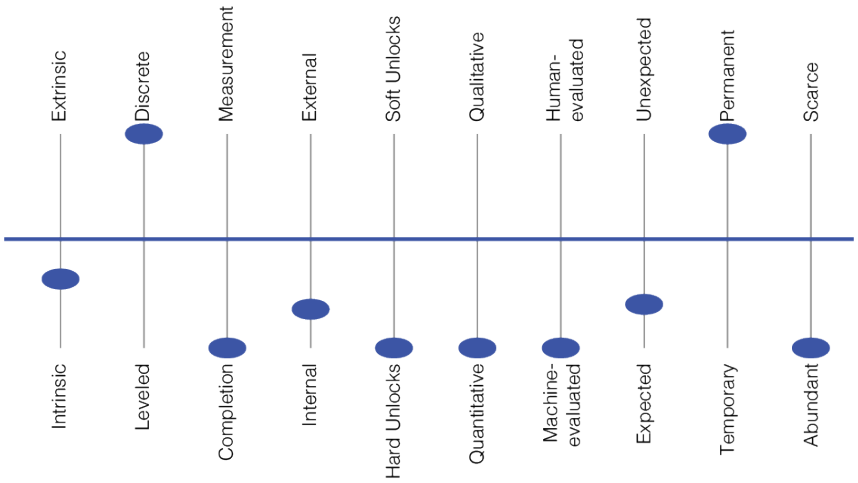

Image 8.1: A visualization of where Gamestar badges fall on the taxonomy.

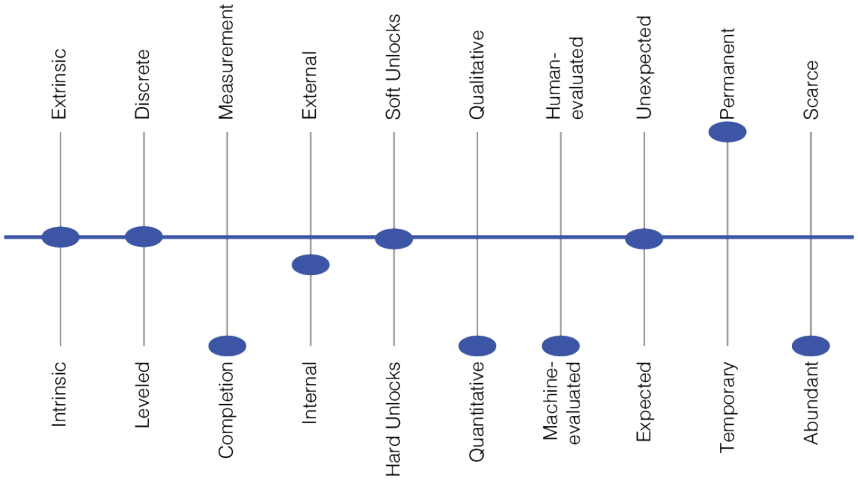

Image 8.2: A visualization of where the Gamestar experience points system falls on the taxonomy.

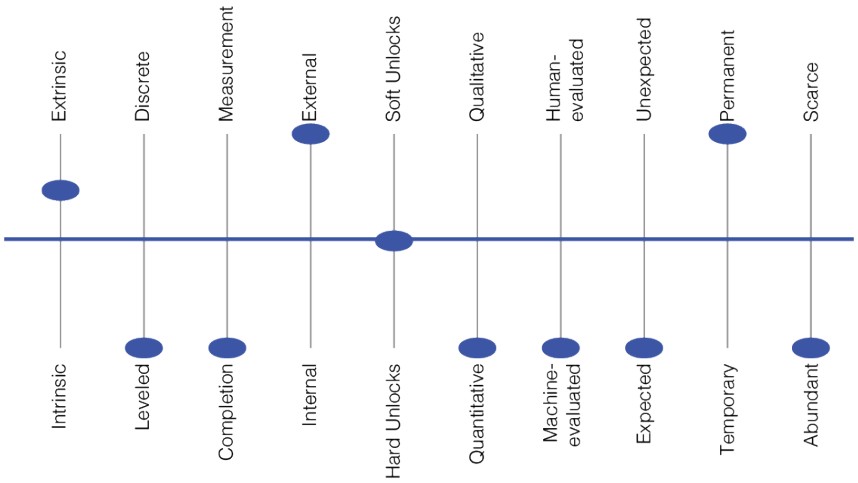

Image 8.3: A visualization of where the Gamestar ranks fall on the taxonomy.

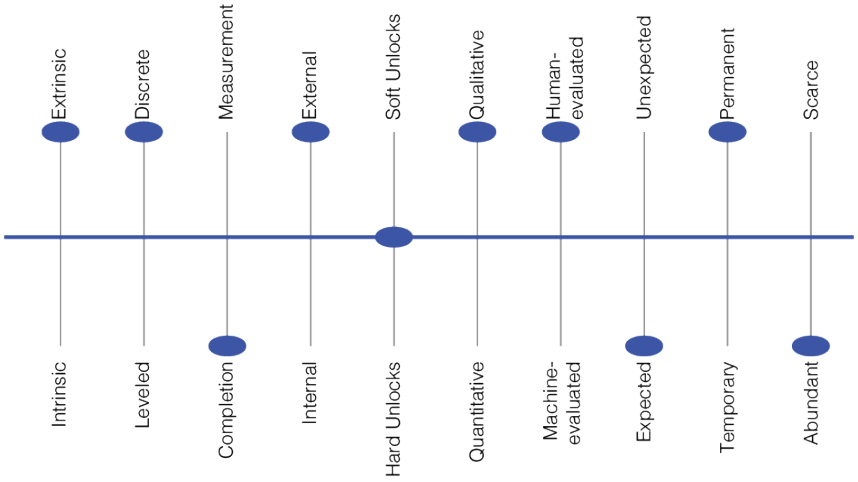

Image 8.4: A visualization of where the World Badges fall on the taxonomy.

The World Badges are quite different. These are clearly a discrete extrinsic reward that is qualitative and human evaluated (like most OBI badges). They are intended to have a clear external value, but because

an ecosystem for these badges has continually failed to emerge (i.e. no credentialers have accepted the badges for credit), that value for the most part does not exist. But these badges are clearly meant to point outside the game system itself.

Several interesting points can be taken from this structure. First, it's interesting that the first three badges are relatively internal facing, and all link into each other in a relatively complex hybrid structure. Whereas the fourth badge, which has a clear external value, sits on a completely different system. This feels like more than just a coincidence. Internal facing badges have to be built on the elements of the system itself, which are meant to encourage the right kind of play. These internal facing badges may not always point to elements that are highly transferable, and to create that kind of transfer, the game designers opted to build an entirely new reward system with a different underlying design. Scott Price explained some of the reasoning behind that decision:

> "[Designing the World Badges] made us rethink our entire system—we'd already done this crazy metacognitive nonsense around learning game design while you played a game about game history. For World Badges, we had to step back and think about what in our game was specific to the game and what was a transferable skill."

Much of this is consistent with Barry Joseph's comments[3] offered in Chapter 3—if we think of the first three badges as what Joseph calls "local" badges, and the World Badges as "global" badges, then we can see Joseph's constraints playing out in Gamestar Mechanic. The local badges, aimed at engaging students, had a hybrid intrinsic/extrinsic structure with a clear unlocking structure and internal value. The global badge, aimed at satisfying credentialers, was a clear extrinsic reward with external value, and felt different and out of place compared to the other badges. It was hard to accomplish these two goals of local and global badges in one well-integrated system.

It's also worth noting the multiple values achieved for students in the multi-leveled set of three rewards. By creating an interconnected set of hybrid rewards, some of the detriments of one reward choice can be alleviated by other choices.

For instance, the Gamestar badges are discrete, which can help with goal setting, but don't give a sense of overall progress. But by interlinking the Gamestar badges into an acculumultated experience points system, students can also get a sense of overall progress.

As another example, the experience points have a lot of detail about progress, but the system involves four different categories and so isn't easily shareable in a format that allows someone's proficiency to be quickly understood by others. By translating the 4-prong experience point system into a single rank, an easily shareable marker of progress is created- what it lacks in detail, it gains in simplicity.

There are other values balanced as well. See Price's words about how to achieve a balance between setting clear goals for students, while still fostering intrinsic motivation.

> "Some badges are listed and appear in your profile as unearned so that you can set them as a goal and work toward them. These are anticipated, contingent rewards, which are shown to decrease intrinsic motivation because players may ascribe the pleasure of success to the reward rather than their effort. However, showing the badges is necessary to direct new players and communicate the structure of the game. To offset that risk, we made other badges that are hidden until you earn them, because rewards based on natural, unrewarded behavior are shown to increase intrinsic motivation. The hidden badges are also more descriptive of your style of play or advanced skills, so that players earning them will feel that the game is recognizing their personal style. That kind of responsiveness also supports intrinsic motivation."

One last point I want to make—both the Gamestar badges and all elements that contribute to experience points are completion-based rewards. These are the main systems that determine your proficiency, and they all are binary—you have completed them, or you haven't. General proficiency is determined by completion-based accumulated discrete rewards (aka points that add up), rather than by measurement-based rewards (aka a grade). In my view, this choice worked out really well for Gamestar Mechanic, and it was the right choice for them to make—and is currently the opposite of how most

traditional micro-rewards in education (e.g. grades, badges) typically work. We'll be keeping a close eye on how this issue plays out in the following case studies.

TAKEAWAYS:

- Using a hybrid badge structure accommodated rewards with different designs that serve different purposes.
- Given that, it was still hard to build for both "local" and "global" needs in one hybrid system.

The Ward Game

Paul Darvasi is a high school English and Media Studies teacher who teaches at Royal St. George's College, in Toronto, Canada. In 2013, Darvasi decided to try something a little different- he developed a pervasive classroom gaming experience called The Ward Game. He used the narrative, setting, theme, and characters from Ken Kesey's novel to create a 30-day game where his students are immersed in the world of the asylum from One Flew Over the Cuckoo's Nest. This game was enacted by turning his all-boys school into the all male world of the asylum, where he used game mechanics such as missions ("prescriptions"), surveillance, play-acting, propaganda production, and group challenges to bring the novel's narrative to life. Details of the reward system were based on an interview with Paul Darvasi.[4]

The classroom experience that Darvasi created is fascinating from many angles, and he has presented and written about his game many times.[5] The use of pervasive gaming in such an effective classroom setting is worthy of a treatise in itself. But for our purposes, the interesting element to consider is, what did Darvasi have to do with his reward system, specifically with his grading and assignment system, to make this innovative play experience work?

The mission, or prescription system, is where most of the reward design is contained. In the first year of the game, Darvasi would send emails to students "from The Big Nurse's account" offering them tasks for rewards which they could accept or reject. As the game evolved in

subsequent years, students could request a prescription at any time from a specific category (art, medicine, creative writing, journalism, music, engineering, etc.). They would be sent a task with a catalog of rewards for completing the task, which primarily took 3 forms: points, in-game money called "c-sticks", or mystery envelopes that gave them bonuses or valuable information. Again, they had the option to accept or reject anything that was sent to them, and they could request another prescription in the same or another category.

So let's dissect each of these three rewards. Points were like standard accumulated discrete rewards. They were also related to the core mechanic of the game: Gain 100 points, and you "escape" the ward, win the game (and get an A in the class). The points were directly translated into a letter grade in the class too, on a normal 0-100 point scale. Points were generally given for completing missions. Also, winning the game early meant you gained free time to do whatever else you wanted to do during class time, though many students opted to continue playing the game even after winning.

One additional use of points was to punish students for violating the rules. For example, one rule was that students could not talk about the game, outside of class. To enforce this rule (and create a culture of paranoia similar to the one created in the novel), if a student told on another student for talking about the game outside of class, then they would gain points for ratting out a student, and the student caught would lose points. By tieing a ratting out system to the points system that became their grade, it added a bit of stakes to finding and tattling on others.

It's also worth noting that the point system has a graduated, accelerated nature. As you gained more points, Darvasi decided to make it harder to gain additional points. At 50 points, the student's disease became "Acute" and the points they gained were reduced by half. At 80 points, the students gained a "Chronic" status and points were reduced by half again, to a quarter of their original value. This essentially created an exponentially increasing system of points, found in so many games, while still keeping a clarity and simplicity in how point thresholds translated into a grade. This was an elegant solution to balancing complexity and simplicity.

The mystery envelopes were relatively simple—they would allow students to gain various bonuses and benefits in the game. Some

were directly related to the reward systems, some were related to other narrative elements of the game. They could help you with completing subsequent missions, or help you mediate elements that might cause you to lose points or C-sticks. Here are a few examples of what was contained in mystery envelopes.

- ***Prescription Extension 48****: This card buys you a 48 hour extension on any prescription. This can also be sold or traded in the Day Room. If this card is lost or stolen, it will not be replaced.*
- ***Acute Immunity****: You can use this to preserve your natural point value until you reach 65 points with no losses for being Acute. This can also be sold or traded in the Day Room. If this card is lost or stolen, it will not be replaced.*
- ***Make it Stop****: This card allows you to ask Dr. Spivey to stop playing Lawrence Welk for any session. This can also be sold or traded in the Day Room. If this card is lost or stolen, it will not be replaced.* (Notes from Darvasi: "In the novel, McMurphy is driven a bit nutty because Nurse Ratched only plays Lawrence Welk in the ward. He begs her to stop, but she won't and he eventually reacts violently. At one point in the game I started playing a Lawrence Welk song ("Misty") on repeat for all subsequent classes. This card was the only way to make it stop.")
- ***Counter Espionage****: If you are caught participating in an unauthorized discussion, you can give this card to The Big Nurse to convince her she has been misinformed and not lose any points. This can also be sold or traded in the Day Room. If this card is lost or stolen, it will not be replaced.*
- ***Film Fest****: This card allows you to complete a task using images from the Milos Forman film One Flew Over the Cuckoo's Nest. This can also be sold or traded in the Day Room. If this card is lost or stolen, it will not be replaced.*

Last were the C-sticks, which were the in game currency system. C-sticks were used to trade for various in-game elements, on an in-game auction site that Darvasi set up, called the Day Room. Most transactions were for the cards contained in mystery envelopes, but sometimes The Big Nurse would be requested to transfer funds from one account to the other for "services rendered," for reasons Darvasi

didn't always know. As an example, one resourceful student received the rare "password hack card" and there was a lot of buzz about a password protected PDF that would prove useful to many of them, so he posted a message on the Day Room that said that he would release the said password at a set time to any other players who paid him 10 C-sticks. He made a fortune, and came back from business school a year later to tell Darvasi how much he loved the game.

One additional use of C-sticks was that they also boosted your grade by being converted into points once the game ended. Every 5 C-sticks translated to 1 point. In this way, this very mutable reward also gained some of the characteristics of the points system.

There was also some trading that seemed to go outside the strict confines of the game. Much of this was never made fully clear to Darvasi. One notable example is that the students who were not formally part of the game began recording people speaking out of turn about the game. Instead of turning in the recordings to The Big Nurse, they sold them back to the perpetrators for a lower C-stick cost than what it would cost them in points if they were turned in. This was done on a black market auction site one of them was running in a secret Facebook group. Whether this is a part of the game, or "outside" the game, and thus has internal or external value, might be a bit outside of the point, as Darvasi clearly bent the boundaries of the game and had it pervade all aspects of his students' lives.

So let's go into the three different kinds of rewards used. First, the points are potentially the easiest to analyze. They resembled a standard experience points system in many ways, as an accumulated discrete reward given for completion-based criteria. The system was still strongly human-evaluated and qualitative, as most educational-based grades are. The points were also one of the more expected elements of The Ward Game's reward structures. The system has a clear internal/external value dichotomy (depending on whether winning the game and getting an A in the class is a part of or separate to the "Game"). It's relatively permanent, except for when points are lost due to the in-game punishment system. It's interesting to note that these points form a reward loop structure, but it's a weak one: even though getting more points does reduce the number of points you get on later assignments thus making advancing more difficult, there's no particular unlocking that occurs, and the assignments themselves

don't get more difficult in any way, they just count for less points. There's a sense where the goal of the loop is really just to get to the end and achieve the extrinsic reward (keep getting points to get an A in the class). The rewards are also abundant—one students ability to gain points is not limited in any way, except by their ability to complete their assignments well.

In contrast are the C-sticks. These are again accumulated discrete rewards. But their most prominent feature is to act as a currency with internal value. They are also notably very temporary, and can be spent or removed at any time, in often unexpected ways, giving them a sort of scarcity. In a sense, the rewards are still gained in a completion-based, human-evaluated, subjective, expected format, since when you get a mission you'll know how many C-sticks you'll get as a result of completing the assignment. But there are so many other facets of how this reward operates in the game, it seems unsatisfactory to stop there. This reward seems to operate mainly in a way to allow the unexpected elements of the game to have an impact on a player.

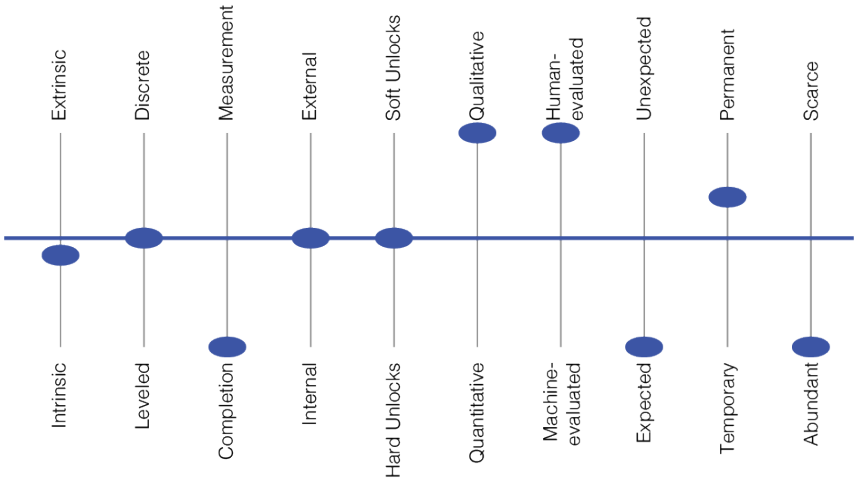

Image 8.5: A visualization of where the Ward Game's points fall on the taxonomy.

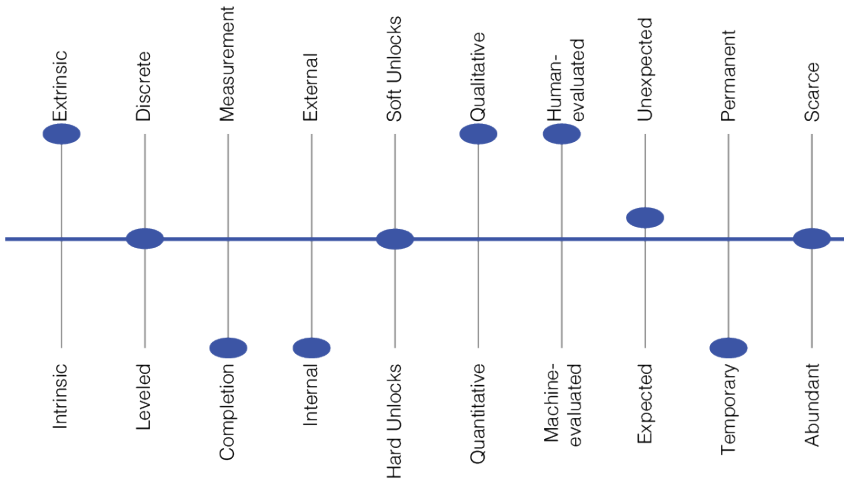

Image 8.6: A visualization of where the C-sticks fall on the taxonomy.

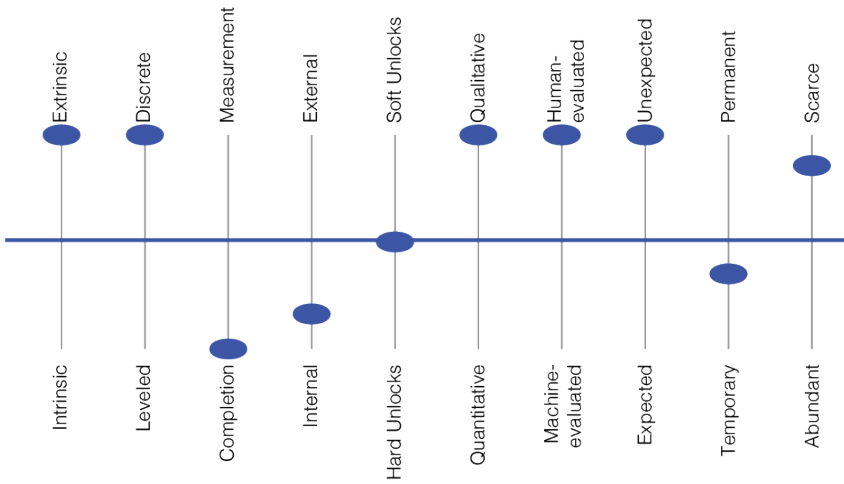

Image 8.7: A visualization of where the mystery envelopes fall on the taxonomy.

Finally are the mystery envelopes. Unlike the previous two items, these are not accumulated, but simply a discrete award. They are most prominently unexpected, and have some kind of clear internal value

for which they can be cashed in. They are also scarce, to different degrees—there is a limited quantity of each kind of envelope. Like the other rewards, these are given out as the result of successfully completing assignments and are completion-based, human-evaluated, and subjective.

There's a few interesting further points to make about student motivation. First, it's worth noting that to make this system work, Darvasi converted his assignments into an all-or-nothing type of system (or, he converted them from measurement-based rewards to completion-based rewards). Although he had given the students in this class standard, measurement-based grades in previous semesters, once the Ward Game started, all assignments became completion-based. He had a clearly defined rubric that guided when students accomplished the assignments' criteria, and he would return work back to students if it was too low quality to gain the points, and allow resubmission. Darvasi said he actually returned very few assignments back to students, and the work he received during the course was some of the higher quality overall work he received from his students that year. This is interesting, because the assignments themselves don't qualitatively differ from the assignments Darvasi gave before the Ward Game started—the major difference is not the assignments themselves, but the reward structure around the assignments (as well as the narrative game framing that they were now placed in). Those additional elements changed how students approached the assignments, and how much effort and quality they put into the assignments.

Another point that surfaces the high engagement of the game is in the win condition. Once students "win" by getting 100 points, students earn an A, and are also allowed to spend their time in the class on whatever they'd like. But, many students opted to stay in the game even after winning, to continue to participate in the missions. Additionally, there were several students not in Darvasi's class that also found ways to participate in the game through unsanctioned black markets. There was clearly a high motivation to participate in the game, which was likely at least in part caused by the rewards used.

Of course both of these motivational points could be attributed to the framing of the game environment itself, as much as it can be attributed to the reward structure. But the reward structure is relatively tied into

the game mechanics and the assignments in a way that's hard to separate, so it would be hard to attribute motivation solely to one and only one of these elements. Really, that's kind of the point—to engage in this highly innovative, and highly effective, model of instruction, the typical assignments-and-grade reward structure needs to be rethought and reintegrated with the new mode of instruction. Darvasi couldn't have changed the way that he ran his course without rethinking his grading structure. Simultaneously, he also couldn't rethink his grading structure without changing the way that he ran his course. It would be hard to research the effectiveness of these elements completely in isolation from each other.

TAKEAWAYS:

- A key shift to a game-based classroom environment is moving from measurement-based rewards/grades to completion-based rewards/grades. This change also implies other changes (like allowing students to resubmit work, and being on call as a teacher to grade work as it's submitted).
- There are a lot of interesting ways to use scarce currency in a classroom to reinforce your learning goals. Currency can have value without an explicit intrinsic reward system.

The Multiplayer Classroom

In 2009, Lee Sheldon decided to try something new with the college courses he taught. He called the experiment "The Multiplayer Classroom" and has documented the experiment in a book by that name.[6] The main goal was to use Massively Multiplayer Online games (or MMOs) like World of Warcraft[7] as a guide for running a class. This involved many changes, like having students choose avatar names, engage in a virtual world, join "guilds" in which they did group work, and of course involved changing the system of grades and assignments in the class. Lee Sheldon describes multiple iterations of his grading system in his book. I'll briefly describe some of the evolutions.

In the first version of the course, the idea was relatively light. Sheldon put students in the role of an MMO character, and for the most part just relabeled existing assignments according to the new course theme. Midterm quizzes still existed as well as regular quizzes, and became mod fights and normal fights. Students worked in groups, called guilds, on a game prototype together, for which they received a collective grade. There were also some newer inventions—a peer review system allowed players to grade each other within their guild on their contributions to the prototype. Sheldon also instituted a PvP quiz ("Player vs. Player", a game term to note areas where players can directly attack and challenge other players), in which players had to shout out answers first to questions to earn points for their guild, and points would only be given to the guild that answered first. The biggest change was that now, instead of averaging point grades on assignments and midterms to get a final course grade, the point grades were accumulated, students reached different levels as they gained points, and the final course grade was determined by what level the students reached.

In this system, the micro-rewards mostly look like those given in a normal classroom. They are measurement-based rewards, qualitatively given by an expert teacher. They are also permanent, abundant, and mostly expected. The exceptions are the peer evaluations, which change who is doing the qualitative judgment, and the PvP quiz system, which feels more like a micro-completion-based reward, that is relatively scarce.

The macro-reward in the class, the overall course grade, was what looked and felt different. It was no longer the average of a series of discrete rewards. Instead, there was a hybrid accumulated discrete reward that built into a leveling system. In Lee Sheldon's own words,

> "Games "grade" a player's performance by attrition. While you could lose XP and even a level in some early MMOs like Everquest, today a player is always gaining XP when he is victorious. This way of looking at achievement has something to teach us educators. Letter grades- the way we align them as penalties for failure- and how our educational system focuses on achievement learning can hinder student progress; the direct opposite of experience points mounting to starts."[8]

The idea expressed here is so similar to our framework, and pointing towards a truly leveled, completion-based permanent rewards system. Sheldon began his course by telling his students "You all have an F" and then challenging them to work up to an A. On a technical, mathematical level, this system can be technically similar to a normal grading system—but the perception, the way the points are presented, is very different. You are given a goal to build up to, rather than a pedestal to fall from. It's more a matter of perception than a real change to the grading/reward structure.

In the third iteration of the class, there was a major change to the preparation for the midterm exam. It became a "guild vs. guild, PvP" session. Each guild had one copy of the textbook between them, they could look up answers in the book, but had to close the book before answering and shout out their guild name to "buzz in" and answer. The questions themselves were more difficult than typical questions, often having multiple parts and requiring guilds to memorize their answer as a group and collaborate. What's interesting is that this activity in itself was a little outside the reward structure. The winners did not earn experience points. But, everyone did very well on the followup midterm exam, better than in any prior class up to that point, which earned them a lot of experience points on the exam itself.

In the fifth iteration of the class, quizzes were changed to no longer be a PvP guild opportunity, but instead were a solo task, but with a twist. Each quiz contained one bonus question. You would only get credit for the bonus question if you answered it right, and everyone else in your guild also answered it right. It encouraged collaboration in studying, and avoided a situation present in the PvP quiz where everyone in a guild got points simply because one person did well.

The sixth iteration of the class once again involves some significant changes. This class was explicitly about designing interactive characters for games. Sheldon asked students to take on the role of designing a character or avatar for themselves in the class in order to learn about character design, blurring the line between where the game starts and ends. The game no longer was restricted to the physical limits of the classroom, and moving around the room of the classroom had a parallel to exploring new geographic areas in the game world and uncovering puzzles to be solved. The syllabus accordingly laid out some "expected" rewards, but they were far below

the total number of points needed to max out the chart and receive an A—unexpected rewards were lurking in the course.

A few other minor but important changes were made. In the syllabus, there was still a leveling system, but grades were no longer specifically listed. There was a general reference that "Careful readers will notice that from Level 10 up the XP system mirrors percentages associated with letter grade."[9] But he also noted clearly that level 10 is needed to pass the course—a clear discrete, binary, completion-based benchmark. This is a movement away from measurement-based rewards towards completion-based rewards, and a defocus on the final letter grade itself.

Additionally, certain levels conferred "skillz points", a new type of reward structure. See Sheldon's note on this:

> "Again, much like characters in standard RPG games, players would achieve new skills every few levels. At first I considered actual skills that would help them in their game quest. This started to feel too elaborate. So instead I decided to cut way back on the random extra credit opportunities and give them 5 XP when they reached 4 levels to add to any grade they wanted."[10]

This seemed to point to a potential intrinsic reward direction—gain XP to gain levels and skillz points, which allow you to play the game better, which allows you to gain more XP. But, given the complexity, the skillz points were backed off, and became more of an extrinsic reward, but one with clear internal value.

There were other iterations towards intrinsic and internal rewards. Players had to design characters, for themselves and for their guilds. Many assignments were about building characters' backstories and actions, and many of those actions affected what future assignments they received. Many of these were conveyed through the narrative, but some were brought into the rewards structure itself. For instance, Sheldon told his students that on the midterm, in addition to receiving experience points for the number of answers they got right, they would also receive one drop of a lizard's blood per answer. If the entire class as a whole gained enough drops of lizard blood, then a character in the story could be healed, and then later assignments would involve

more interactions with that character. So the class's collective skill on the midterm would affect the story, and their assignments. It wasn't quite a reward loop, but it did link rewards to class assignments in an interesting form of internal value. Do well on this assignment, to get enough lizard blood, to keep a character alive, to have more assignments that involve that character. Sheldon noted that the number of blood needed to save the character translated to a C average for the class- but this class scored higher than any class ever did on the midterm, well above a C average. The lowest grade overall was a B-. The character was saved.

Overall, this system had many interesting features, from its original inception to its final product. The overall grade in the course was treated as an accumulated discrete award embedded in a leveling system rather than an averaged score from several discrete measurements. This framing itself changed a lot, even if a lot of the individual assignments remain standard human-evaluated, qualitative, measurement based rewards. The effect of this change is rather profound, and can be clearly seen in the student evaluation comments offered through Lee Sheldon's book. There were also a variety of different assignments that varied in their abundance and scarcity.

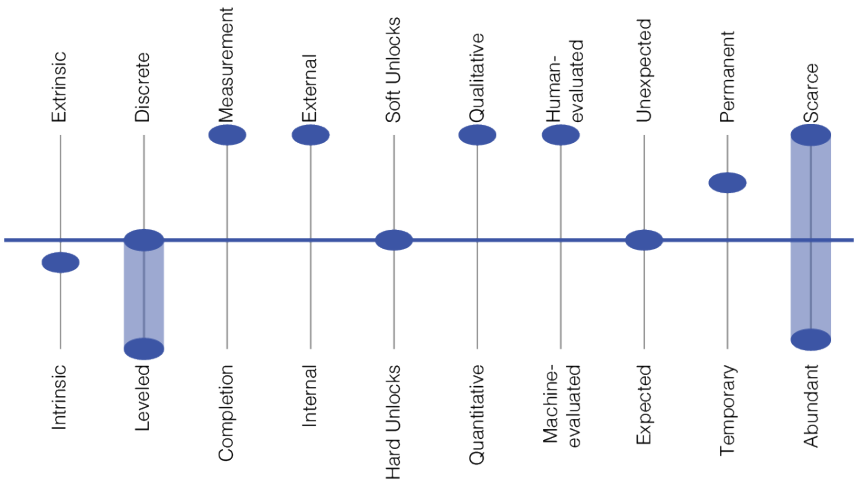

Image 8.8: A visualization of where the multiplayer classroom's grading system as a whole falls on the taxonomy.

But many of the assignments themselves changed too. Some became more completion-based (e.g. the bonus question that was only awarded if everyone got it correct). Some gained clear internal value to the course assignments (e.g. the midterm lizard's blood). Some became unexpected (e.g. the evolving assignments based on the story in the story class). Some became scarce (e.g. the PvP pop quizzes where whoever answered first got the points).

What this showed in general was the amount of change that can be achieved just by changing how a final grade is calculated and perceived. Many of the changes made by Sheldon just involved how individual assignments built up into a course grade. The assignments themselves were still often very traditional, compared to what Scott Price or Paul Darvasi created in their earlier case studies.

TAKEAWAYS:

- A feeling of building up into a final score is a powerful feeling, and uniquely enabled by accumulated discrete rewards that sit in a hybrid leveling system.
- A lot of interesting motivation can be achieved by thinking about team-based, or group rewards, vs. individual rewards.

Innovative multiplayer reward systems in games

As a last example, I wanted to discuss two interesting and innovative competitive multiplayer games that used an innovative rewards structure. Although it's not exactly clear how to directly translate this to educational environments, there are some situations where competitive multiplayer scenarios are used in schools (chess, math olympiad, FIRST Robotics, as well as any sport). On a fundamental level GPA is often used to rank students, and has a competitive feel in many schools. This feels like a situation that could have unexpected inspiration, especially as educational rewards grow and evolve in unexpected ways.

There's a basic problem presented to anyone who has a competitive

multiplayer environment. You want to rank players, so that you can match players up with an equally skilled opponent to optimize the fun. Additionally, players also want to know how good they are in a game, which implies sharing a ranking score with other players. But the obvious solution to this problem is one that we explicitly recommended avoiding (see Chapter 9). Namely, a quantitative, measurement-based, permanent reward system in which you are judged relative to peers is bound to foster a fixed mindset, a feeling of constantly being judged, which is unhealthy from a psychological viewpoint. Additionally, it also demotivates play for those not at the top, or those not motivated by hyper-competitive instincts.

So how do you design a system that provides a satisfactory play experience but avoids the more judgmental elements of the typical reward structure used in competitive multiplayer environments? Go offered one example with its ranking system. Something about reducing the level of quantification and making it less leveled and more discrete means that players move up and down less frequently. It focuses less attention on how a specific performance mapped to your stated ability, and makes the change to a new rank a more significant, noticeable, and reflective event. Which definitely seems like it is in the right direction, but two very recent games went even further, and in my opinion to good result.

HearthStone and Splatoon 2 draw from very different genres of games, but share the fact that they have relatively active competitive multiplayer scenes. And both have chosen a structurally similar response to managing rewards in that multiplayer environment. I won't get into the details of the play in each game, but both games have a structure of playing a multiplayer match in which you either win or lose the match, and then enter a post-game rewards screen. I'm going to solely focus on that rewards screen.

In Hearthstone,[11] the system is strikingly similar to Go. There are 50 normal ranks in HearthStone (almost the same as the 30kyu system in Go). You start at 50, and as you play better you can advance to rank 1. The first 25 ranks are "new player" ranks—as you gain ranks, you permanently keep them. The final 25 ranks are "normal" ranks. Each "season" in the game lasts about a month—whatever rank you have gained by the end of the season relates to a reward, with higher ranks

giving larger rewards. At the end of the season, you then drop 4 ranks (going no lower than 25).

Image 8.9: A sample rewards screen taken after a match. This player is currently at rank 14, which is called "Raid Leader." They have 2 of the 4 stars needed to advance to the next rank. Image taken from the HearthStone Wiki site: https://hearthstone.gamepedia.com/File:Ranked_win_streak.jpg

Within each rank, there are a certain number of stars that you can gain. In general, as you get to higher ranks, more stars are required for that rank. For every win, you gain a star. For every loss, you lose a star. Once you have all the stars in a level, your next win causes you to gain a rank. Similarly, if you have no stars and you lose a game, you also lose a rank, though you go to the lower rank with almost all of your stars

filled in. You also match with players at a similar rank to the one you are on. Additionally, there is a safety system built in—once you gain a "milestone" rank (5, 10, 15 or 20) you cannot drop below that rank no matter how much you lose, until the season ends. This allows some room for experimentation and safety in rank.

So, the system discretizes the process of gaining stars ranks, and also offers clear benchmarks in the form of a leveling system (the ranks). The system still has accumulated discrete awards (the stars) but there are far fewer of them, so they contain less detail and feel a little less judgmental. And they are notably wrapped in a hybrid leveling system. The evaluation now feels far less subjective than a normal ELO system, and can offer clear goals and benchmarks. You aren't just seeing a giant score go up and down, but are trying to advance up one more rank by getting one more star.

Splatoon 2[12] goes even one step farther. The system in Splatoon 2 resembles the normal grading system in school. You start with a C- grade, advancing to C, C+, B-, etc. all the way to a max rank of S+ (S is one level above A). Within each level, there is a bar that needs to be filled—each game gives you a certain number of points towards filling that bar. Additionally, the bar can be filled with cracks after each loss you take. There are three outcomes that can result from this system. If you completely fill the bar before it cracks 4 times, then you advance to the next rank, and your cracks and points are reset. If you gain 4 cracks and your points are below a threshold "ok" value, then you lose a rank and your cracks and points are reset. If you gain 4 cracks but are above a threshold value, then you remain at your current level, your cracks reset, and the bar loses a certain number of points (generally, it only lets you keep a certain amount of the points you gained above the threshold).

What's interesting about this is that Splatoon does have an ELO-like ranking system that it uses to match players within games (it also only matches you with players within your letter grade too). And the number of points you get in your bar also depends on the ELO-like system—play tougher opponents, and gain more points at the end of the match for a win. But when you are playing, you don't really notice the ELO system—you are mainly thinking about getting enough wins before you rack up 4 cracks. The competitive, subjective element that ranks you against other players kind of disappears, and you really only

think about a system in which you are competing against yourself to level up your own progress bar. It feels like a parallel to what runners often note—that running is really about beating your own previous best time. Even when you are in a race with other runners, you are often more concentrated on beating your best time than beating the other runners.

Image 8.10: A visualization of the post game rewards screen from Splatoon 2, from a recent match that I played on April 4, 2020. This shows all the elements described in the text. The bar is about half full, and has surpassed the "ok" line indicating that even if the bar is not completely filled before 4 losses hit, the character won't lose a level. There are 2 cracks right now, from 2 prior losses at this level. And the level is current rank "A-".

An additional note is that the rank system changes a bit for the most highly ranked players in the S+ level, and that at that level, the system does more closely resemble traditional ELO-like systems, as well as prominently featuring a leaderboard. The rewards system therefore feels quite different for the 99% of players with ranks between C- to S+, and the 1% of top players in the beyond-S+ system. Considering the research that rewards structures that emphasize competitiveness demotivate players,[13] this is an interesting decision—one system for the highly skilled (and likely highly competitive), and a different system for everyone else. And also considering that in education we are interested

in reaching the 99% of learners better, the system Splatoon uses for the 99% is the one most relevant to us.

The systems employed in Hearthstone and Splatoon 2 feel like the future of any multiplayer system that either explicitly or implicitly has to rank people against one another. By discretizing a system of advancing up levels and making advancing be a personal battle of gaining enough wins in some short period of time, the leveling system manages to feel like a general marker of one's competence without feeling like a subjective evaluation of one's worth. Especially with Splatoon 2's system, it feels very much like the classic Final Fantasy experience points system. Win games, to gain points, to go up ranks, to play more difficult opponents. It suddenly becomes a bit more like an intrinsic reward loop, rather than the typically extrinsic ELO-based reward.

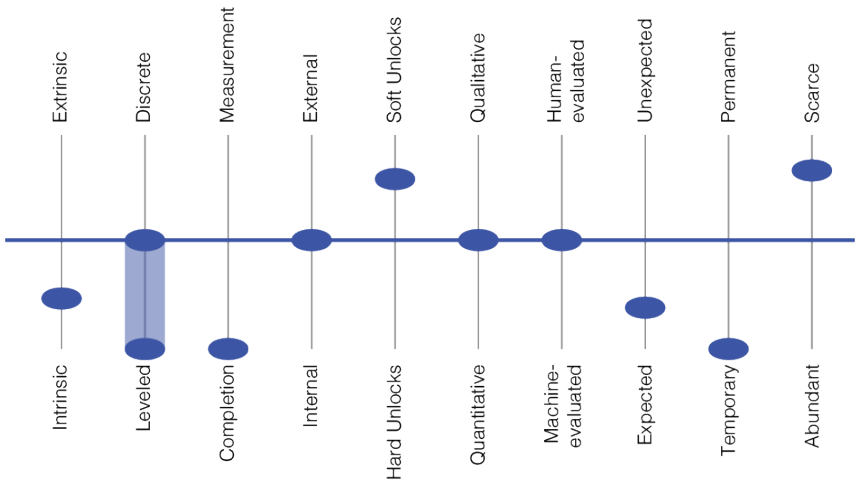

Image 8.11: A visualization of where the multiplayer matchmaking reward systems from Hearthstone and Splatoon 2 fall on the taxonomy.

I offer these examples in this chapter because in education, we currently have a strong need to rank students relative to each other, and that need doesn't seem to be going away soon. In a sense, school then acts as a "competitive multiplayer ladder." There's some clear parallels here—in school, we have the grade on an assignment, which is

a measurement based reward. That assignment grade is morphed into a class grade, or even broader a GPA, which feels like an accumulated reward that acts as a measurement of your ability. A similar system exists in ELO reward structure—your performance after a match results in a measurement-based reward of how many points you gain or lose. This then adjusts your overall rank, which is an accumulated discrete reward that acts as an overall measurement of your ability. If classic, measurement-based grades are analogous to the more traditional measurement-based ELO system, then the interesting question to ask is what discretized school reward system is most analogous to the discretized system offered by Splatoon 2 and Hearthstone?

TAKEAWAYS:

- Competitive environments that need to rank players/students can still set up an intrinsic-reward-like structure to appropriately motivate play. The system can balance non-judgmental feedback, extrinsic ranking of players, with a reward system that fosters inward-looking, reflective views of a player's skill improvement.

These examples have hopefully shown some promising directions in how intrinsic reward structures, different design choices on the taxonomy, and more meaningful rewards can be incorporated in real educational settings. The next two chapters will generalize some principles and changes to school that must be made to put this fully into practice.

9. How Do You Make Good Rewards?

Hopefully by this point you are convinced that intrinsic rewards are a good idea, have worked well in games and maybe have a place in education. (If you aren't convinced, then kudos to you for reading this far anyways). With this chapter, I want to determine the right steps forward for using rewards in education. I'll go through some best practices and recommendations for the design of reward systems in this chapter mostly working from theory, and then describe a practical longform example of putting these ideas into practice in the next chapter.

The Purpose of Education

One of the main difficulties in making recommendations for educational systems is that often educators don't completely agree on the purpose of education. Without understanding the purpose of education, it becomes really hard to recommend a system that can best reinforce that purpose. This may sound odd because surely everyone agrees that the purpose of education is to help students learn.

But what is learning? It's a general term that everyone agrees is important, but it means different things to different people. Does it mean gaining a body of content knowledge? Does it mean developing key literacies? But then which literacies: math, reading, technology, science, etc.? How about key problem solving skills? Or an ability to create or make artifacts? Or maybe an epistemic knowledge of what it means to be a professional (i.e. a scientist, or a writer)? Or maybe it simply means developing a lifelong love of learning, or developing interests? Maybe it involves developing socio-emotional skills?

The truth is every professional in education probably believes that several, if not all, of these things listed above should be included in

the purpose of education. Which then makes any design problem for education quite difficult, and often filled with tradeoffs. A design that most effectively helps students learn content knowledge might in turn fail to develop socio-emotional skills. Or a design on maker-education that focuses on the ability to create might miss out on some key domain literacies that aren't involved in that specific making activity. Sometimes, this is simply a matter of not being able to do two different things at once, and time-limitations cause some purposes to be prioritized over others. Other times, the most efficient design for one purpose can actually act against another purpose. It's this second case that is most concerning to us in thinking about design frameworks.

I don't really have a general solution towards this problem, except to acknowledge that different designs will further different purposes. As best I can, I will try to clearly contextualize design decisions in which educational goals they most foster, and acknowledge which goals they might most limit in consequence. The result is that there isn't really a single answer to the problem "How do you design effective reward structures for education?" You should choose a different reward design depending on what kind of student learning you want to foster the most.

Audiences in Education

In addition to thinking about different purposes of rewards in education, there are also different audiences that rewards can be designed for. The educational badging movement discusses different stakeholders that must be accommodated when designing badges.[1] We can borrow this idea and think about three stakeholders for reward systems: the learner, the credentialer, and the value provider.

The learner is relatively simple—they are the person gaining the reward by completing some criteria. Counting them as a stakeholder acknowledges that they both need a reason to gain a reward, and that they have to engage with whatever system is developed in the process of gaining those rewards.

The credentialer is typically considered the teacher, but it doesn't always have to be. This is the person who "credentials" the reward, or

verifies that the learner has accomplished what needs to be done to complete the reward. This makes the most sense to think about in a human-evaluated system, but even in a machine-evaluated system, the original designer could be considered the credentialer. In some after-school or informal learning environments, it could often make sense that a credentialer might be a separate person or team from the lead instructor.

The value provider is a relatively unique concept that we haven't discussed much, and is often very specific to the educational badging world, because it's so taken for granted in the standard school system and because it's irrelevant to most commercial games. This stakeholder is the person who gives the reward value by acknowledging that reward as being valuable for some purpose. This can be one of several people—for high school learners, this could be college admission boards that allow rewards to be counted as a part of an application. It can be employers that accept rewards with resumes when considering job applicants. It can be the Department of Education that allows rewards from informal environments to count towards diploma requirements. In the standard school system, the school itself is essentially the credentialer—the rewards offered in schools are clearly tied into elements that the school owns and provides value towards: get good grades, so that learners can graduate.

With these definitions, we can then think about what it means to design rewards for these three groups of stakeholders. In general, games typically design for the player (or learner) first, as most games have a single stakeholder. But in schools, we typically rate the value provider needs above all else, and design for those stakeholders first. What this means in practice is that in schools, rewards must be able to accurately assess learner's performance (which is a stakeholder need), and then, in whatever design space is leftover, can we address learner motivation.[2]

Let's break this down a little more. In our school system, as we discussed previously, most reward structures are fully entangled in our assessment structures. For example, look at GPA, diplomas, high stake tests, and passing/failing classes. All of these systems are designed for value providers—or they are meant to signal whether a learner should move on to the next step in the educational system. This means for a learner, these rewards feel entirely extrinsic, permanent, and often with

external value. They feel that way because they were explicitly designed that way, to accommodate value provider needs which are external to learner needs.

To really redesign schools to insert intrinsic rewards, we would need to design the system for learners first, and consider what might motivate learners the most—whether or not such a system offers high value for assessment and value providers. The term "ludic assessment" has been introduced to think about this idea.[3] If we develop assessments that are ludic or playful, they will equally value 1) fostering learning, 2) motivating students and 3) assessing student learning accurately. Such an assessment system implicitly needs to balance multiple constraints—it's impossible to achieve all three value propositions (learning, assessment and motivation) at once. But, reducing assessment value to increase learner motivation is a hard pill for many to swallow. In multiple contexts where I've participated in discussions of ludic assessments, they often quickly turn to deep seated constraints in the educational system as a reason for why assessments can't be made more playful.

Here's another way of looking at the situation. Gibson et. al (2013), describe 3 main purposes for badges specifically, but this can be applied to educational reward structures in general: 1) map progress and foster discovery, 2) signal reputation beyond the community where it is earned, and 3) incentivize learners to engage in pro-social behaviors. Note that purpose 2 is mainly for value providers, and purpose 3 is mainly for learners (purpose 1 applies to all three audiences as everyone would like to see maps of student progress). It's interesting to think about how these three goals map onto the reward taxonomy, which can in large part explain the various skepticism noted by those in the field.[4] In short, all of the badge skeptics describe how the final two purposes seem to be in conflict—allowing reputation to be easily signaled beyond a community will corrupt motivation, or that cultivating motivation will result in a badge system that can't easily achieve reputation outside the badge community. This is the tension between "local" and "global" rewards, described in terms of stakeholders. Local badges are primarily for learners and are about learner motivation. Global badges are primarily for value providers and are about signalling reputation beyond the community. This is the root of the conflict.

Let's go through these three purposes one by one, and map them onto our taxonomy.

First, a system that allows a mapping of progress most notably needs to be a permanent and expected system of rewards. Beyond that, the system can take many different forms.

Second, a system that signals reputation beyond the reward system most notably needs to be external in value, meaning the rewards have meaning outside the rewards system itself. To be easily recognizable, it also needs to be a discrete reward. Additionally, all of this gears the reward towards potentially being quite extrinsic in nature.

Third, a system that motivates learners (to engage in pro-social behaviors) is most prominently an intrinsic reward structure. It may also contain elements of a leveled reward built in an unlocking structure that may have internal value to the experience.

You can see that the second and third goals are mostly in direct conflict with each other, and that conflict centers around the premise of this entire book: extrinsic rewards vs. intrinsic rewards. Any design decision made to better achieve the second goal is a design decision that does a worse job at achieving the third goal, and vice versa. The only real way to balance the two together is to build a relatively complex and sophisticated hybrid badge structure that can achieve both design elements (and therefore goals) in different levels of the system. This, quite frankly, is hard to do well in any practical setting (but see the Gamestar Mechanic example in Chapter 8 for an example).

As can be seen, thinking about stakeholders can be a useful way to explicitly think about what purposes a reward system should serve for whom, and how designing for one stakeholder can compromise value for a different stakeholder. Choices about stakeholder and purpose of the reward system need to be explicitly considered and decided upon, in order to inform what reward design is best.

Design Recommendations for Schools

Below I offer some general guidelines on what badge designs typically reinforce or discourage different learning goals.

Principles for specific content and skill areas

Content areas that have clear progression can benefit from hard-unlocking intrinsic reward structures.

Math and reading are two great examples of content areas that have a well-defined progression of knowledge that advances and builds upon itself. In a sense, these content areas can form a natural intrinsic reward with a sort of soft unlocking structure. Do math, so that you understand simple concepts well, so that you can do harder math, so that you can understand harder math concepts, so that you can do even harder math, etc. In this case, to create more well-ordered problem solving and to offer clearer feedback on progress, it can be helpful to translate the progression into an explicit designer-made hard unlocking system. This can help students understand when they have mastered a concept before they can move on to the next, harder concept. Many math games already employ this principle in their level design, for instance like Dragonbox.[5]

Content areas that have discrete concepts can be represented by discrete rewards, or hybrid discrete-leveled rewards

Science or Social Studies are typically represented by discrete interrelated areas of knowledge, like cellular biology or ancient American civilizations. This can naturally lend itself towards a discrete reward system. But as many of these areas have greater and greater layers of depth at which they can be understood, it can be worth embedding a leveling system within the discrete badges. Meaning, maybe you can have a level 1, level 2, or level 3 understanding of cell biology. This can help give the discrete concepts independence from each other, while still noting that there is a depth of knowledge to gain in each area.

Skills or concepts that cross content areas can be served by accumulated discrete rewards

There are certain skills or high level concepts that bridge across subject areas. In the next Generation Science Standards, these are explicitly called out and referred to as “Cross-Cutting Concepts.”[6] As

an example, “Scale Proportion and Quantity” is a concept that bridges multiple areas of science. You can talk about the difference in scale between cells, tissues, organs and individuals, or you can talk about the size and relative distance of planets in our solar system, or you can discuss the order of magnitude difference in timescales between weather-related phenomena and climate-related phenomena. These concepts cross content, but are also embedded in content. They are often practiced as you do content-based exercises, and in this case a recommended reward is to use an accumulated discrete reward system. These skills are repeatedly practiced in different contexts and so a system that can send points to different cross-curricular categories as different activities are completed can accommodate these kinds of skills and concepts well. Imagine that as someone completes an activity, a discrete content-based award might be given (e.g. a “Mastering the Solar System” badge), while a series of points are given to a separate system to acknowledge concepts (e.g. experience points are awards for the “Scale, Proportion and Quantity” attribute).

Permanent, expected and abundant rewards can foster any content or skill based rewards

Every content area is a little different, but for the most part if you are targeting content knowledge, you want the rewards to be permanently gained once acquired, you want the rewards to be expected and well known in advance, and you want the ability for every student to be able to potentially gain every piece of content or skill. You could compromise one or more of these items, but that should probably be because you are balancing some other competing learning goal along with content knowledge.

Principles that avoid fixed mindsets and promote growth mindsets

Carol Dweck has outlined a compelling view of the difference between fixed mindsets (“I have a fixed ability that does not change”) and growth mindsets (“My ability can grow over time with effort”). There are also compelling reasons to believe that for educational purposes, a growth mindset is healthier and leads to better learning.[7] And

additionally, the kind of feedback that a student receives has a direct influence on affecting their mindset. As rewards are a very direct form of feedback, their design can have deep influences on a student's mindset.

So what rewards discourage fixed mindsets and foster growth mindsets?

Qualitative, human-evaluated, measurement-based rewards often reinforce fixed mindsets. They get worse when they are external and have an extrinsic reward structure.

That was a mouthful, but it's a relatively simple idea. When you are asking a person to offer a subjective evaluation of another person's ability, this always feels a little judgmental. When it is further related to something outside the experience itself, and the reward itself is extrinsic in structure, then it feels even more judgmental. Now, sometimes this situation is unavoidable—if the skill itself needs to be evaluated by an outside expert along subjective criteria, then this is just what needs to happen in that situation. Many certifications (e.g. getting your driver's license) are structured to work this way—big test at the end that is conducted by an expert to see if you are qualified to do the thing you need to do. Of course, there are other models for giving certifications—apprenticeships, for instance, offer another model, or the Boy Scouts approach to how and when merit badges and badges for rank are awarded. But those alternate models are certainly less efficient and more costly than the basic model. It will take some real effort in your evaluation approach to prevent the basic model from fostering a fixed mindset and destroying motivation and interest in an activity, but it can be done. You are just starting from a hole that you need to dig yourself out of.

It is worth noting that the typical cycles of a traditional classroom (i.e. do homework to get a grade from a teacher) exactly fit the criteria described above. And it's also interesting how the Boy Scouts were able to maintain the qualitative and human-evaluated elements while moving away from the other taxonomy elements.

Measurement rewards can promote growth or fixed mindsets depending on if the measurement is perceived as feedback or a judgment. The permanence of the reward plays a key role.

I spent a fair bit of time discussing the nuances of how measurement criteria can interact with qualitative or quantitative rewards to foster fixed mindsets in Chapter 2. I offered the idea that completion-based rewards can do well to avoid some of the more stigmatizing effects of qualitative rewards on growth mindset. Alfie Kohn offers basically the same idea, noting that if grades are turned in a binary (or completion-based) reward, much of their stigma is removed.[8] But in contrast, Deci and Ryan[9] had noted that measurement based rewards offer a greater potential to offer feedback on progress, which can seem contradictory.

Much of this really depends on how the measurement is received. If the reward feels judgmental and is permanent, then it feels like the reward is commenting on your innate ability, and serves to reinforce that you are innately good or bad at a task. If the reward feels like feedback and can be changed with later attempts, then it directs attention to how your effort on the task did or did not work out and how it can be improved in the future, which reinforces the idea that your ability can change over time. The former is similar to how grades on assignments typically work in school—the latter is how the 3 star system at the end of an Angry Bird level operates.

Leveled, permanent, quantitative rewards that are embedded in strong intrinsic reward structures with soft unlocks are great for growth mindsets.

If your goal is explicitly to foster a growth mindset, then I'd suggest a permanent reward that grows in levels as you advance, to really reinforce the idea of skill growth. Keeping it quantitative avoids any feeling of judgment. Soft unlocks will also reinforce how your skill has improved over time to allow your progress in the system. And an intrinsic reward loop helps focus attention on the activity itself, rather than your ability on the activity. All of these feel beneficial for promoting growth mindsets.

Temporary, machine-evaluated rewards are a great idea for promoting growth mindsets.

To go in what seems like the opposite direction, I'd also suggest one of the potentially most powerful ways to foster growth mindsets is with one of the lesser used reward types—temporary, machine-graded rewards. These are the typical status indicators you might see in a game when you are doing particularly well. The machine-evaluated nature of the rewards is necessary to take away the judgmental factor and to easily track the temporary state of the reward. But I think a reward that appears or disappears based on how you are acting in the moment seems best attuned to relate your skill level to your actions and not to a fixed aspect of your personality.

Principles that allow rewards to offer meaningful feedback

Feedback is an exceptionally important idea in the use of rewards. On one hand, this is very much related to designing good formative assessment, which in essence are rewards intended to give a student feedback on progress. From another perspective, Deci and Ryan[10] have noted that rewards that offer feedback are more able to foster intrinsic motivation in an activity.

Measurement-based rewards or leveled rewards can offer meaningful feedback.

Feedback is important not just in the context of goal setting but in education in general. Measurement rewards very explicitly offer feedback in the way the reward is presented. Leveled rewards also offer a lighter form of feedback in that they generally measure progress in how many of the small reward bits you have gained, and can be even more powerful feedback when embedded in an intrinsic reward loop. In both these cases, execution is important, and it's up to the designer to ensure that the rewards are valuing the right aspects of an activity.

Avoid discrete, completion-based, machine-graded rewards if you want your rewards to offer feedback.

The main types of rewards for which feedback is most difficult to

encourage is a discrete, completion-based reward. Their simplicity and directness leaves no room for feedback. This is what the most basic achievements in games look like: those achievements may do some things well, but one of those things is not offering feedback on performance.

Human-evaluated, qualitative rewards can offer feedback—but there's a lot of nuance in the execution that determines if it is done well or meaningfully.

If there is a measurement-based award evaluated by a human, then a subjective element is introduced to that reward. The feedback indicates how well someone else thinks you have accomplished an award. Although it is valuable to receive feedback on performance from an expert who knows how to mentor well, there are several important qualifications to that statement. In cases where an expert can't be found, or the expert is not a good mentor, then the feedback offered can be counterproductive. So although humans can provide more nuance and detail in the evaluation process than machines, the value of the feedback depends on the person doing the evaluation itself.

There are ways to mediate this, mainly by being as transparent as possible on the criteria that allow the reward to be given, or the rubric used to grade the activity. This can in part help the activity from feeling less qualitative and more quantitative. Involving the student in defining the criteria or rubric itself can go even further in mediating the potential negatives and ensuring the reward proactively fosters goal-setting.

Principles that work in makerspaces

To develop creativity and open-endedness in your activity, aim for either qualitative, human-graded rewards or unexpected rewards.

There's two ways to go here, to basically get people to think outside the box. If you use a very loosely defined qualitative criteria that is defined by non-experts (ideally peers), you can get people to use their imagination to fulfill that criteria in ways that delight their friends. This is essentially a description of the core mechanics behind the card game Apples to Apples.[11] On the other hand, another way to explore creativity

within a more designed system or environment is to hide the rewards and make them unexpected. This forces experimentation and creativity amongst students in order for them to be delighted with the rewards.

Discrete, completion rewards offer simplicity and more frequent feedback at the cost of detailed feedback and potential for useful reflection.

A discrete, completion award is what we commonly think of when we think of rewards and badges. It's essentially "Do this thing, get this reward." It's simple, straightforward, and understandable. But the cost of simplicity is a lack of information about the performance. And that information, whether in the leveled rewards or measurement rewards, is what enables deeper feedback and reflection. Both of these variations can work, and perhaps a mix of two co-occurring systems might work best in a Makerspace.

In Makerspace environments, we might naturally gear towards discrete, completion-based awards. "Did you make a bird-feeder with wood? Great! You now have the woodworking badge." To some extent these are useful- they are both great markers of progress, great goal-setting devices, and can often be given quite quickly after an activity is completed, allowing for a more instantaneous reward. But one of the most useful activities students can engage in, especially in maker education, is critical reflection on their own work. Many reward structures can be built to accommodate that reflection, but the reward itself may be less discrete, completion-based, and more delayed to accomplish this.

Principles that support student goal-setting

Expected, permanent rewards best enable goal setting.

A well-designed and openly shared reward structure is one of the most classic and powerful ways to enable goal setting in students. This in many ways is directly stolen from the gaming world, in which openly shared achievements or skill trees allow players to set goals for themselves in gameplay. Goal-setting reward structures most prominently need to be shared transparently, and therefore are a very

expected award. They need to be permanent, and not change with time.

Intrinsic reward structures with an unlocking structure that is leveled helps with goal setting.

An intrinsic reward structure that involves leveling up through a hierarchy is particularly good at not just offering a clear goal to be achieved, but also helping motivate players towards that goal. An unlocking structure also helps by imparting a clear sense of accomplishment towards that goal.

Principles that foster a lifelong love of learning

Intrinsic reward structures foster a lifelong love of learning.

This one is pure and simple. By making the reward for doing the action be as simple as doing the action more, you motivate your learner to value the action rather than the content of the reward itself. When that action is learning, your students begin to value lifelong learning.

A mix of discrete and leveled goals, perhaps in a hybrid structure, is usually best.

There's a challenge here, in that you want to give students a clear sense of progress, but you also want to impart a sense that there is always more to learn and master. A leveling structure generally notes progress, and you almost want the system to sort of extend into infinity or at least over years if not decades of someone's lifetime, to allow for continual improvement. But a structure stretched out over such a long time frame can become repetitive and cause one to lose their sense of place in that system. Discrete rewards allow for a clear stop and acknowledgement of progress, while the leveled rewards allow a continual growth trajectory to extend beyond any one discrete award. The Boy Scouts badge system is one example of how to achieve this balance well. Although not an explicitly designed reward system, our original hobby example of knitting socks also fits: you are progressively getting better at the skill of knitting, while having individual moments of success when socks are completed at discrete moments in time.

Soft unlocking structures, or a hard unlocking structure with internal value, also promote love of learning.

Soft unlocking structures are great in that you are always aware of your own ability and its development over time, which creates an understanding and appreciation for that ability. A hard unlocking structure avoids that by creating an external system, but if you create a currency with internal value in the hard unlocking system, you can help redirect energy spent on the unlocks back into itself. Why unlock this next gate? So that I can gain the ability to work on this next, harder gate.

Principles that give students agency or freedom of play

Design more rewards than the typical student should be able to achieve.

There's a principle that works in games that hasn't explicitly been acknowledged, but rings true in almost all of the case studies in the previous chapter. Almost all game systems have more badges available than the average player can possibly gain on one playthrough of a game. Meaning, beating the game should not mean getting all the badges. This is true for two different reasons—first it's due to playstyles. Achievements in games often foster playing in different ways—on any given playthrough, you have to choose a particular way to play the game (thus enabling freedom of identity). Second, achievements are for the above and beyond, not the standard (thus enabling freedom of effort). They encourage playing that goes above a normal, beat this game, run. Both of these reasons serve to foster student agency or choice on a very basic level. This principle is different from enabling scarcity—every reward can still be abundant, or possible for every player to achieve, without it being true that every reward will be gained by every player over the normal course of the game.

Hybrid leveled/discrete structures that end up with branching tree reward systems offer students meaningful choice in rewards. This

can be further emphasized if the branches have internal value to the game.

Leveled systems are great, in that they can allow a lot of player agency, in allowing players to decide how far, or how many levels, they want to go to as their goal. But if you embed discrete badges into a leveled system and end up with a sort of skill tree, you can go a step further. Players can implicitly choose goals amongst a wide variety arranged in a clear display—see the Civilization III[12] skill tree for an example. Simple discrete rewards can allow for simplicity and clarity in goal setting, but that comes at the cost of meaningful agency.

A strong internal-value currency system in an intrinsic reward structure helps give choice.

Many times when there was some kind of freedom at the "reward" step of the intrinsic reward structure in Chapter 7, that was because the game featured some kind of currency. That currency was gained in the reward loop itself, and then spent at the reward stage. Many games were meaningful because there was ample choice in how to spend the reward.

With time and coordinated effort, we can find the right way to amplify education through appropriate use of rewards. The final chapter of the book will show one such attempt, to combine these ideas together into a concrete, comprehensive picture of an entire school system driven by intrinsic reward structures.

10. Do Intrinsic Rewards Have a Place in Education?

At this point I have described the value of intrinsic rewards and noted many interesting situations which have effectively used these rewards. But, our examples are far steps from a traditional school environment. Plus the evidence is quite clear that extrinsic rewards are still dominant in school—we have grades, high stakes testing scores, transcripts, GPA, and diplomas. And this doesn't even consider the minor moment-to-moment rewards used in some classrooms, like gold stars and stickers.

Whether intrinsic rewards belong in education actually boils down to a related question—should learning be an enjoyable activity? If learning should be enjoyable, then without a doubt we should figure out how to place intrinsic rewards in school. With intrinsic rewards, the value of the reward is derived from the action that you are doing. If learning is a meaningful action (or one that can actually be enjoyable), then intrinsic rewards are an ideal mechanism to amplify that enjoyment. If learning should be an unenjoyable task to be accomplished, then intrinsic rewards have less place in education, and extrinsic rewards are a good option.

I'm of the view that learning is an enjoyable activity. I've had fun learning in my own life, and have seen the joy of learning brought to students that I have mentored and taught. And as a consequence, I believe intrinsic rewards have an unrealized potential in education.

I'd even go a step further, and say that learning is one of the fundamentally enjoyable and meaningful activities in life. One of the three pillars of Self-determination Theory is competence. Learning is fundamentally about getting better at something, and therefore inherently builds up competence and is intrinsically motivating. Of course it is also satisfying when you have choice in your learning, and can share your learning in a real way with others, but even without

these two extras, learning should be inherently satisfying in that it builds up competence.

For me, the ultimate goal of education is not just to teach someone a few facts, but to impart a lifelong love of learning. Or in other words, to inspire a self-motivated learner. These are people who want to "learn for the sake of learning," which is the classic credo of the American liberal arts college. This is really just another way of saying "Learn, to be able to learn more." The self-motivated learner, the ideal we strive towards in school, views learning as an intrinsic reward. If this is a goal that we want to achieve, then why not create explicit intrinsic reward systems in schools that place learning as the central action.

At this point, we've also shown quite clearly that intrinsic rewards do work in games, mainly through communicating competence. Learning is also about competence, and so it seems that many of the successes in games should also be applicable to learning. Several authors additionally argue that games, at their fundamental level, are about learning.[1] Whenever you pick up a game, you are learning how to play the game, whether it's learning the controls of the game, understanding how the game's system operates so that you can more strategically manipulate the game's system, or becoming immersed in the mindset of the character you embody in the game. If games are about learning, and we agree that games are both enjoyable and have made good use of intrinsic rewards, then it seems logically that we should expect the same from the learning that exists in schools too.

So my answer is a strong yes, intrinsic rewards do have a place in education. But, how do we incorporate such rewards into educational systems? Let's put the principles into practice and design a new school based on the ideas in this book.

Designing a School based on Intrinsic Rewards

As a warning, it will become obvious that complying with the best design principles for a reward system requires changing the underlying structure of how a school functions. From the previous sections of the book, this shouldn't be too surprising—the school system is already

overloaded with reward structures. What might not be obvious at first glance is how deep these existing structures run.

Also, for those that might object that adding a reward structure will necessarily corrupt the purpose of education, I would also point to the above point—that our system already has rewards in its basic structure, so this is not so much an attempt to add a reward structure to school, as it is a way to redesign the ones that already exist in a way that mimics more meaningful, intrinsic reward design.

So let's start by imagining a school. Many schools have a set of values. So to start we will make a system of rewards derived on those values. For this, let's imagine a STEM-focused school with an emphasis on empowering their students and giving them voice. Their main values might be Scientific Inquiry, Leadership, Collaboration, and Communication.

Now we want a system with some constraints so that each of these skills will be recognized in a consistent way. But we also want flexibility, so that students and teachers can customize these rewards to their individual needs. We also want to allow students agency, and recognize that each student is likely gaining different skills to a different degree.

So let's make a set of four tags, one for each of the four values above. Every individual reward (aka, grades on assignments) that a student gets will be tagged with one of these four skills. Now you can readily see a portfolio being built up. That portfolio has four pages, one for each tag, and each page will display all the rewards that a student got with that tag. This feels a bit similar to the GameStar Mechanic system.

We also want students to feel progression, to level up these skills. So each individual reward will have a point system attached to it. We can use a standard kind of terminology, of Small, Medium, Large, and Extra Large achievements (aka "T-shirt sizing"). These can increase exponentially in value from 1, 3, 10, and 30 points. Except let's not call them points, let's refer to them as growth credits. Similar to the way the growth credits scale exponentially, the amount of effort required to get each reward should increase exponentially. Small rewards should be gained with a simple homework assignment that takes about one hour, Medium should be a more complex multiple-hour assignment, Large should be an essay or longer project that might take a week to complete, and Extra Large should be a multi-week or unit long project.

This is in a sense somewhat similar to the credit hour system

implemented in many US colleges. A certain number of hours must be gained in each required subject area before a student can graduate. The main difference in our example is that we are giving what might be called more micro-credentials, or things that happen within a class. We are also giving credit for active student work within a course, rather than simply completing a course. We are also giving credit for skills that cross-courses.

Of course, the subject area credits should matter too, and schools should require that students progress in certain content areas in addition to developing their four skills. We could tag individual rewards with content tags in a parallel system that mimics the skills system—but we can probably go simpler here too. Right now, most classes are arranged in subject areas, and let's assume we generally keep that class structure. In each class, students can gain one content-based credit for their engagement in that course. We'll give this a sort of completion-based criteria, and let the teacher decide what is needed to achieve the completion criteria for the course. So in each course, you'll have some minimal requirement you'll need to meet to gain the one content credit for that course.

Ok, so at this point, it doesn't feel like something that would be too different. Basically we are telling teachers that they should transform their curriculum into these microcredentials, which is work but not huge structural change. When students complete assignments, they get growth credits, which shows their growth on key skills. And when they pass a certain level of work in the course, they also gain a content credit for the subject area of that course. But, if we stopped here, we'd end up with a relatively ineffective badge system. It would feel a bit layered on top of the education, and have little intrinsic value. So let's get to the more difficult pieces next, and layer in the basic structure that will allow an intrinsic reward structure.

First, we need to talk about grading. It makes sense that teachers should be the ones evaluating these microcredentials that they have created for their class, but ideally it means we allow some flexibility in the grading process. Students should be able to submit whenever they are ready to submit, rather than by a specified deadline. Additionally, they should be allowed to submit multiple times. Teachers should really change the letter grade to a "pass or fail" or really an "achieved or not yet sufficient," mimicking the way assignments work in the Ward

Game. The grade should probably always be accompanied by written feedback on what has or has not allowed the student to achieve the reward. But this structure completely changes a teacher grading workload. Teachers will be grading different assignments at different times. Teachers will no longer grade all of the same assignments on the same day. This will take more of a teacher's time—but, that time is being spent supporting students with specific feedback on their chosen learning artifacts, or what can be called personalized learning. So that time is going to a good purpose, as was noted in the Ward Game.

But what about the usual letter grades? My recommendation is to just get rid of them (a point on which I fully agree with Alfie Kohn, to come full circle). They are a reward system that serves no purpose in this new structure, and will just be redundant with, and therefore cause tension with, the microcredentialing and growth credits system. If a school is serious about the microcredentials and growth credits being something that matters, then the school should be replacing the old thing that matters, aka the grading system. Additionally, based on the second set of changes described next, it will be increasingly difficult to even find a place for grades in the school structure. So let's get rid of grades at our fictional school.

Several times we've talked about the importance of choice and meaningful agency in this book, and this school restructuring can't avoid that issue either. If this rewards structure is going to work, then students need to have choice. There's two levels at which that choice needs to exist- one is within classes, and one is between classes. The second is easier to discuss, so let's start there.

Students should have choices about which classes to take. In one sense, this isn't new. Many colleges feature high levels of class choice. High school (and to some extent middle schools) also feature some ability to choose classes—but usually limited to an elective or two, or how far you want to advance in AP science or social studies, for instance. This proposal basically makes all schools look more like colleges—allowing more meaningful choices of what classes to take at every level of our school system. The school can still require a certain number of content credits that need to be completed, but can offer several different options for how to complete that content credit. In our school, based on what classes they choose, students will have different

opportunities to level up different skills. For example, one math class might focus really on applications to Scientific Inquiry, while another strongly emphasizes Collaboration through group work. Both fulfill the required math content credit, but each allows different opportunities for progress in the skill system. What this means is that the choice of classes is directly cued into our reward structure, as it affects what skills students can improve upon. Implied is the idea that every student might choose to level up a different set of those four skills, and that should be both accepted and encouraged. We'll come back to how that is recognized in diplomas in a second. Exactly how much class choice should be allowed is not exactly clear, and would depend on the amount of in-class choice that is allowed.

The within-class choice is what is a much bigger change. The goal is that students have a high degree of agency in how they choose to engage in a given class (i.e. Scot Osterweil's freedom to engage and freedom to experiment). If the goal of a class is for students to gain microcredentials that correspond to key learning, and we want students to have meaningful choice within a class structure, then there should be far more opportunities for micro credentials than a given student could complete. Which is basically a huge shift from how we run classes. This means that not every student will be doing the same thing at the same time. They might not even be doing the same things at different times—they will literally be doing different things in the same course. By the time the course ends, every student will likely have completed a different number and a different set of artifacts as a result of the course. It also means that the teacher should make more microcredentials available than every student should get—the right order of magnitude is probably around two times as many options as the average student should achieve in the course. This is essentially asking a teacher to create two times as many assignments for that course as they would normally use.

This may sound overwhelming for a teacher, but remember our conversation from grading—the teacher is taking more of a mentor role in the class, freeing their time for individual work with students. They are the initial game designer, setting the constraints of the course and how to engage with the material through a thoughtful series of microcredentials. Then they act as mentors to guide students through the "game" that they have created. This is somewhat consistent with

the principles of Flipped Learning, though a very different execution of the idea. It also has parallels in the game world of Dungeons and Dragons—teachers are essentially acting as dungeon masters.[2]

Now, this doesn't imply that students will *never* be doing the same thing at the same time, and I imagine a good class would have several of those moments. For instance, an English teacher might say "We are going to spend the next two weeks reading *To Kill a Mockingbird* together. There are 6 potential microcredentials that you can choose to complete during this reading—all of them are different sizes and related to different skills. I expect everyone to complete at least 1 of the microcredentials." Or it can be a bit looser like "We are going to spend the next two weeks reading *To Kill a Mockingbird*. You can choose to read the book with us, and aim for some of the microcredentials available from this book, or you pass on this book and continue on your projects from *1984*."

A Science class might look different. Perhaps here, a teacher begins the class proposing four big questions related to physics, and then a series of microcredentials that are related to exploring each question. Each student must choose to complete the path of at least two big questions, but they can choose which two. Two weeks later, a series of four new questions are offered, and students choose their new path. So at any given point in time, students are always working from the same set of four questions, but those questions change over time.

As a third example, maybe a math teacher decides to really break up their coursework and make all of their assignments transparent. Maybe they break each assignment into a set of a few questions, with each set being something that takes approximately 30 min each to complete (so all Small assignments) but they make 50 such assignments available for the course, and require that students complete at least 20 of them. They might then lecture on the material every day such that students are hearing the same lecture at the same time, but students have the ability to move forward or backwards in the assignments as they need, engage more deeply or more simply with the wealth of microcredentials available to them.

One feature of this new system that is completely absent from our existing system is that there is a freedom to engage in courses. If a student is really not into a course, they could do the minimum needed to advance through the course and get the content credit. But if they

are really into a course, they can go above and beyond, and do more than they might have done in that course in a more traditional structure. Students that go more deeply into a course are recognized with more microcredentials, growth credits, and skill development. This is somewhat analogous to the idea of "extra credit," except imagine that the majority of the course is extra credit.

The phrase "minimal requirements for content credit" is important—it is sort of assumed that students would be required to engage at some basic level to be enrolled in that course, which makes sense. But, I'd encourage even at this level to ensure that agency is included. There's a huge difference between "earn at least 5 small microcredentials and 1 medium microcredential to pass this course" and "everyone has to earn exactly these 6 microcredentials to pass the course, and then there's this extra stuff you can do on top if you are interested." Every course, subject, and teacher will be different (it's important to respect teacher agency!), and sometimes the latter might just be what makes the most sense—but it shouldn't just be the default option for every course.

This system will add a lot of openness into how students spend their time—to navigate this, each student should have a school adviser that they meet with with some regularity, and who helps them dedicate their time effectively. On a minimal level, the advisor should make sure the student is not simply doing the bare minimum in each course—although that's all that is strictly required for a course, the school-wide requirements for graduation should require students to do more than the minimum to graduate. So in at least some courses (in a well-designed system, maybe half their courses), students should be going above and beyond. To monitor this, advisors should have a dashboard that displays a student's current growth credits progress in all their courses.

Additionally, with the basic structure described above, there is now room to build intrinsic reward structures and unlocking structures in both the within-course and between-course structure. Within courses, we can create levelled assignments, such that initial stages of assignments need to be completed to unlock harder levels worth more points. Or, we can have a sort of sequential set of assignments, where one assignment unlocks further assignments in a sort of skill tree or gameboard. Maybe there is an initial assignment that is only about the

content and counts towards the content credit, and once completed unlocks 4 additional assignments that allow a student to apply that content towards each of the 4 skills. All would be examples of hard unlocking structures with clear intrinsic reward loops—complete assignments to unlock more difficult assignments, to practice the skills and content at a harder level, to complete even more assignments and unlock even more difficult assignments. In the math class example above, where the entire assignment structure was broken into 50 small assignments, we can also envision a more soft unlocking structure. The assignments link to the lecture, and probably mostly become possible to do once you hear the lectures. But you can also work ahead or advance beyond the lecture in the assignments list—the only thing preventing you from progressing forward is your own knowledge and skill.

We can also devise a between-class unlocking structure. This is rather easy to discuss, because the school system (particularly in college) is already filled with hard unlocks—they usually go by the name of course prerequisites. What's nice about this new system is increased flexibility in thinking about and defining prereqs. A teacher can still require that specific courses are completed as a prerequisite for a more advanced course. But they can also require that a specific number of growth credits are gained in those courses—so you have to overachieve in the prerequisite course to qualify for the next course. Or, they might require a specific set of microcredentials as a prerequisite (like, maybe you have to get two specific XL microcredentials in Biology to be able to take AP Biology, but those two microcredentials are not part of the minimum requirement for the Biology course). Or maybe they require a certain number of total growth credits in certain skills, or in certain additional tags that individual teachers may develop. Unlocking structures create an opportunity for teachers to collaborate on the connection between their courses through nuanced prerequisites. This also reinforces a nice further reward loop—complete courses and level up skills, so that you can unlock harder courses that let you practice the skills and content at an even more advanced level.

So let's talk about graduating requirements for a second. Remember in this school, there are 4 major skills, and each growth credit is tagged with one of those major skills. For a student to graduate, let's require that they must achieve a certain number of content credits, and a

required number of growth credits in each of the 4 skills, AND they must achieve a "specialization"—meaning they must achieve a higher number of growth credits in at least one of the skills. So maybe, 10 content credits are required in each subject area, 200 growth credits are required in each skill, and 350 growth credits are required in their specialty. The microcredentialing system should be designed so that achieving this goal is roughly equal to the same amount of work that students did in the old system—thus in an overall sense, this new structure is the same amount of schoolwork as the old structure, just arranged very differently. And it should also be designed so that an overachieving student might be able to achieve two specialties, but no student should be able to achieve all four specialties. To recognize "minors", a school might also want to recognize an intermediate level of a skill, like 275. Or, they might want to adjust the growth credit graduation requirements so that there is more specialization and less uniformity, so maybe 100 growth credits in each skill, and 500 growth credits in their specialty. Or they might couple certain growth credits threshold with requiring that at least two XL projects are completed in each skill, if they want to emphasize project-based learning and ensure every student has engaged with at least some minimum number of large projects.

It's worth noting that even if two students choose the same specialty, their path to that specialty will be very different. So even in a graduation path where two students look similar at a high level (they have the same number of overall growth credits in the 4 skills), they will be quite noticeably different at a closer look (they will have completed different courses and microcredentials).

There is likely another new role needed in the school of microcredential designer, which I think is what might normally be called an instructional coach. This new role should monitor the overall microcredentialing system. They should ensure that there are well more than 200 available growth credits that can be gained in each specialty, and that there is the right distribution of microcredential sizes for each skill. For example, Scientific Inquiry might have a ton of S and M microcredentials, whereas Leadership might only have a few XL microcredentials. Both might have the same number of total potential growth credits available, but the progression to get there will look very different, and it would be the job of the microcredential designer to

work with teachers to balance the system in a way that works for the school.

As one additional digression, it's easy to see how this might scale up to a district-wide initiative. In this case, the district might define a set of 4 key values that all schools in their district should focus their microcredentials around. But, they might additionally allow each school to customize the system to their unique needs by adding an additional 2 school-specific values of their own (so 6 total values at each school). The district might require that certain classes, and certain XL microcredentials, be offered in every school, but then allow schools or individual teachers to define the smaller microcredentials that work for them, and to fill out their curriculum with whatever works well for their schools' and their teachers' expertise.

It's worth noting that we are really leaning in a competency-based learning direction. In general, competency-based learning and competency-based assessment do tackle the problem of redesigning the rewards in school directly, and can involve even larger changes to the structure of school than the ones described here (such as allowing students to pass courses when they exhibit mastery, rather than when the course ends). I chose to describe an example that still involved regularly-timed classes with subject-specific teachers, to show an example that still achieved effective reward design, with as minimal as possible changes to the logistics of today's standard school.

Let's talk about where this fits on the taxonomy. Let's start with the assignments within classes. These are now clearly discrete, completion-based rewards—there's an assignment you need to complete, and once it is done it's complete. It can be resubmitted, but once the assignment is completed, that completion is permanent. Some of the dimensions depend on the particulars of the class and the assignment (remember, we are leaving a lot open to the teacher to define in a way that works for them). But there is strong internal value to this hybrid reward system, in how completing assignments feeds into the growth points and content credits system, as well as potentially unlocking further assignments in the class. In this larger hybrid structure, the assignments are definitely cued into a long-reaching intrinsic reward loop—complete assignments, to gain content credits and growth credits, to unlock further learning opportunities that allow even more credits to be gained.

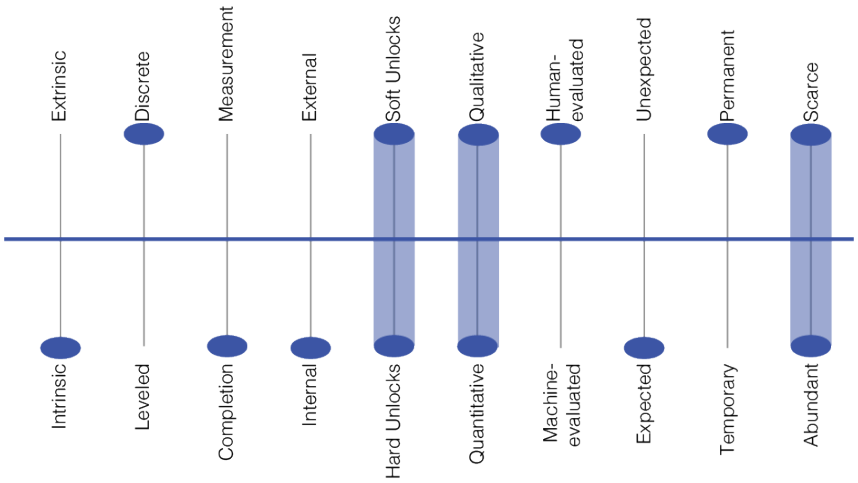

Image 10.1: A visualization of the assignment completion reward structure for our imaginary school.

Now let's talk about the growth points. This is a classic accumulated discrete reward. These end up having a similar feel to the Gamestar Mechanic experience points—they are again completion-based and permanent. They add up together over time, feeding up into the graduation requirements and course unlocking requirements. They have a mix of internal and external value, both feeding back into the different reward systems, but also acting as an external indicator of how you progressed through the system that can be used for graduation requirements and listed on resumes and college applications. It's also not strongly linked to an intrinsic or extrinsic structure, sitting in the hybrid structure between both systems. They are mainly abundant, though there is a time limitation in that a student must choose which they will and won't get, creating a small degree of scarcity.

Now, the content credits. These are again accumulated discrete rewards, though adding up to a much lower number than the accumulated discrete rewards. They are also completion-based, and permanent once acquired. They are in many ways similar to the growth credits, but with a stronger internal value and a stronger tie in to the intrinsic reward loop. They also have a very hard unlocking structure with course requirements.

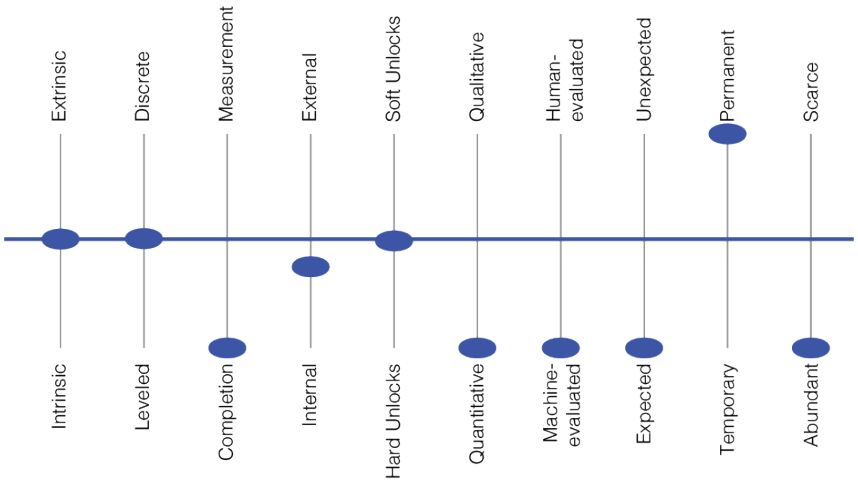

Image 10.2: A visualization of the growth points reward structure for our imaginary school.

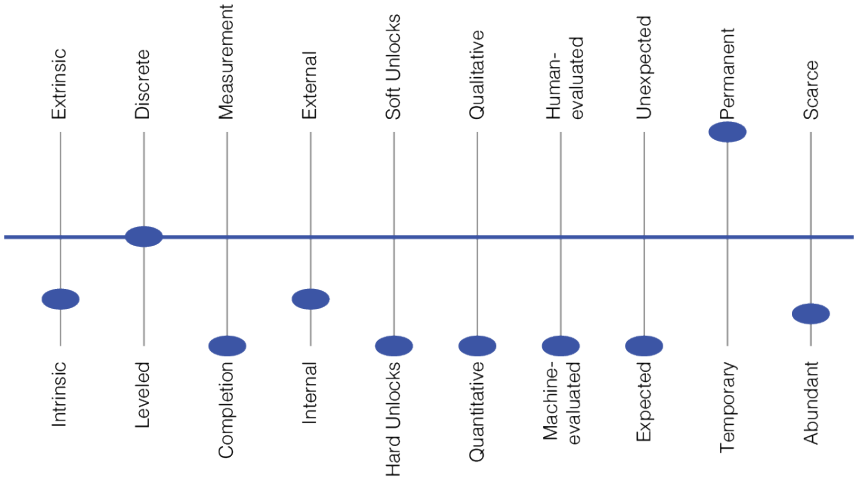

Image 10.3: A visualization of the content credits reward structure for our imaginary school.

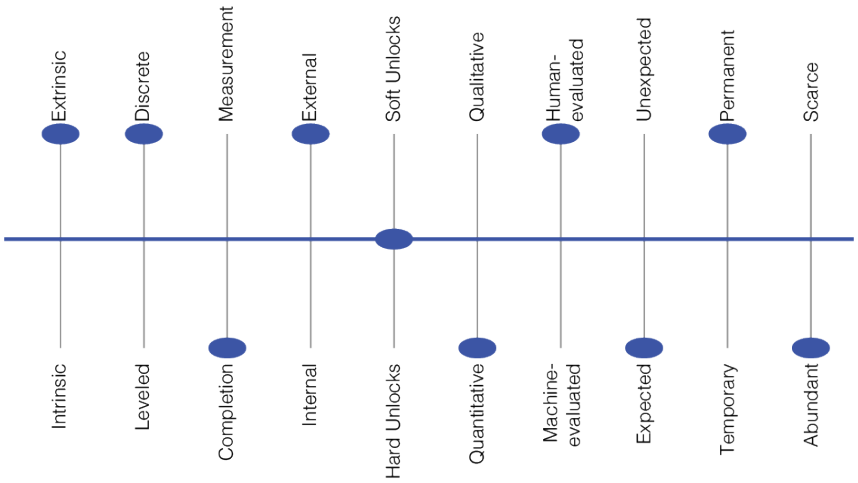

Image 10.4: A visualization of the diploma system for our imaginary school.

Finally, let's discuss the diploma system. The diploma system is a discrete, completion-based reward—once enough growth points and content credits are gained, a diploma is awarded. Because of this system, it has extremely clear quantitative criteria that determine when it is awarded. The diploma is sort of the end of the loop, and in the context of this school has a clear external value, towards either gaining a job or being accepted into the next school level (in contrast to how we looked at the diploma in Chapter 4 and Figure 4.8, where the diploma was considered in context of a student's entire educational journey).

At this point, it's easy to see how other rewards can also be layered over this core system. The schools could have sponsored surprise challenges that suddenly appear at certain times of year. These could be structured as contests, to create some scarcity. The reward from these could be growth credits, but it could also involve a separate one-off reward that confers some internal or external value that is separate from the reward system. These can function like the mystery cards in the Ward Game, that confer extra school benefits (like extra floor passes or late passes, or maybe that allow a student to skip a required M assignment in one course that is needed for content credit). Or they could function like the Multiplayer Classroom's envisioned "Skillz" points system, and give students powers that they might not normally

have to help them complete their work better (like bringing in outside resources for a challenge that aren't normally allowed, or being able to turn an individual assignment into a group assignment to collaborate with teammates for a combined score).

Separate from challenges, there's a lot that teachers can do within courses to change what the rewards look like. There are many lessons from the Multiplayer Classroom that can involve adding scarcity or group-based points that enable challenge. The system is designed so that teachers can create unique experiences for their classes that can still easily and clearly align with the larger reward structure. As long as the assignments reward growth credits that are attached to a certain tag or value at the school and the course has a clearly defined content credit criteria, the teachers have freedom and agency to customize their classrooms in any way that they'd like.

What's important is that this system allows clear linking between individual actions that students complete on a day to day basis to the overall goal of school. This clarity of goal, and sense of progress and accumulation towards the goal, is something uniquely enabled by designing good, intrinsic reward structures. Because assignments that students complete level up in a single reward system, they can see this sense of progress. The way they go through a class, the way they progress between classes, and the way that they build up towards their graduation requirements all rely upon a clear system with the message of "learn, to learn more."

Student examples

Finally, let's describe the story of two example students to see what an educational trajectory can look like from a student perspective in this new system. In this example, we'll imagine a high school that uses the four values we described above. The school requires a minimum number of growth credits in all four skills to graduate, and a higher number of points in one specialty, and a basic number of content credits in each subject.

Chelsea

Chelsea is a junior in high school. She is very interested in engineering, and appreciates the opportunity to choose her own path towards an engineering focus in high school. In her freshman year, she went to one of the student showcases that featured work from Mrs. V's "shop" class. The students use what they learned to design solutions to problems in their school. She had a senior interview her and work with her to design a new banister system that allowed freshman to see better during school assemblies. Ever since, she has been interested In taking Mrs. V's class herself, and helping out younger students herself.

Chelsea has been working with her advisor, Mr. H., to chart a path that helps her get to that point. She's been more focused on what skills she needs in math and physics class, but they've had many conversations about how Leadership and Scientific Inquiry are related to being an effective engineer. Chelsea is shy and has always found Leadership intimidating and nerve-racking, but has taken courses recommended by Mr. H. to develop that skill track. Scientific Inquiry is something she has pursued with a lot more enthusiasm, and is on track to gain enough points by the middle of her senior year to have a specialization in Scientific Inquiry. She is talking to Mr. H. this week about which path to pursue in her senior year. She could pursue the specialization in Scientific Inquiry, but a lot of the advanced courses in senior year are more about running science experiments and learning chemistry and biology, which isn't too appealing to her. Likewise, she could pursue a Leadership specialization, but that would also involve a series of courses that are mostly not in science.

As she reviews her trajectory with Mr H., she decides to be very honest about her goals to invest fully in Mrs. V.'s course. They review the course-if she gains every microcredential available to her in Mrs. V's course, which takes a lot of effort, she will still fall short of her graduation requirements without taking other courses. They both then learn that Mrs. V's course has more credits available in Collaboration than in any other skill. Her course is not only about designing solutions, but about doing so in teams with others. Looking at her skill portfolio, Chelsea's Collaboration total is not that far behind her Leadership total. Many courses she took to level up her Leadership skill also involved a lot of teamwork, which she did appreciate—without realizing it, she

gravitated towards the tasks that were more geared towards collaborating with others than taking a leadership role. With Mr. H., they identify some steps she can take with the FIRST Robotics team (an extramural club she is already involved with) to stop working towards leadership and start earning more collaboration points, which allows Chelsea to breathe a sigh of relief—she's been very anxious about trying to move into the leader role on the team, but really enjoys participating on the team. She also knows her classmate Tacora is really interested in leading the team next year, and didn't want to have to fight with her for leadership.

They identify two other science courses that also strongly involve teamwork and collaboration, and have the potential to earn a lot of Collaboration points. Chelsea finds both courses more exciting than the Leadership or Scientific Inquiry tracks that have been laid out. In fact, one of them is about collaborating with a local college on some of their field research in a local lake. She hadn't paid as much attention to the course before because it was in biology, but once she reads the description and sees what is available to do in the course, she finds the course really interesting. She really wants to meet some of the real scientists and see what their work is like. And she likes the idea of working in a large team to collect data to help someone with research, especially around a lake that everyone knows has gotten too polluted recently. She knows this will also help her community.

She works out with Mr. H. a new plan to have a specialization in Collaboration. It will take a lot of focus in her coursework, but if Chelsea pays attention to what assignments she chooses in the next few courses, she'll be able to easily get to her graduation requirements, while still being able to explore her interests in engineering and helping others.

Patrick

Patrick is a sophomore, and he's been having difficulty at the STEM high school. He has always gotten good grades in math and science, and that's the main reason why his parents picked out this STEM academy for him. But the level of choice he has is overwhelming. He knows how to play the game of school—he is assigned homework, he

studies, he gets it done, and then passes the test. But here, there are not many tests, and no one really tells him what to do.

He's been meeting with his academic advisor, Mr. H. Patrick always completes work well and grasps the content, but he doesn't engage in a lot of additional assignments and materials. Patrick tends to respond to the bare minimum required in every course to get content credit, not because he is trying to do the bare minimum but because that's what he understands the course to be about. Many courses have required assignments that are necessary to complete the course, but on their own don't award many points towards the graduation requirements—it's the "non-essential" work that involves applying that content that gives the opportunity to build up lots of growth credits. But Patrick more just sees the required assignments as the main part of the course, and the other work as fluff. Mr. H has tried to encourage him to do a bit more than just the bare minimum, in one of the four specializations at the school. Because of Patrick's aptitude for math and science, he's been encouraged to pursue the Scientific Inquiry track. But Patrick saw what a lot of the science projects were last year at the science fair—they were a fair bit different from the ones in middle school. They were very open-ended: every poster board looked different. Every experiment was different. There wasn't a clear formula to follow. He wouldn't even know where to start. He's interested in learning about all the science there is in the world—there so much to know. He rather spend his time learning more about what other people have already figured out, than trying to do a science project himself, about something someone else probably already knows the answer to.

He went into Mrs. F's Biology class, which has been one of his favorite classes so far. She's set up a system where every single homework assignment is laid out ahead of time. You can see from the start all of the work required in the course. Although Mrs. F does give lectures, the homework allows you to see what is coming and work ahead of schedule if you like. Patrick really likes this challenge, of trying to figure out the assignments before Mrs. F gives him the answers in lecture. He uses his textbook and lot's of searching on the internet to try to find the answers ahead of time, basically learning on his own. Mrs. F does have several optional science experiments that you can do, which really help level up your Scientific Inquiry skills. Patrick isn't really excited about

those, but knows those experiments are the whole reason he's in the class, to do that "extra" work beyond the required assignments.

Today, Mrs. F. pulls him aside. She starts by asking him if he has started on any of the experiment-based assignments, and he admits, no, he hasn't. She then mentioned that she knows he's been working ahead—he always seems to be at least one week ahead: how does he do that? He mentions how he likes to challenge himself to see if he can figure out the content before she even goes over it in class, and it's a fun challenge. Mrs. F then says she has an idea, if he's interested. They can pick out a topic that is 3 weeks ahead of them. He can complete that homework ahead of time on his own, and then he can prepare a lesson on the topic to teach the rest of the class. Patrick thinks about it and finds it pretty interesting. He does like talking about science with others. That was the best part of the science fair in grammar school. But he says maybe he shouldn't do it, because Mr H. says he is supposed to focus on Scientific Inquiry. Mrs. F. says that maybe they can discuss it with the three of them together.

During the next advisor check in, Mrs. F. joins them and talks about her offer. She would make an XL credential for Patrick, that would be in the Communication track. Patrick is very excited about this, more excited than Mr. H. has seen him yet, even though he seems to say he doesn't want to do it because it won't help him graduate. But then Mr. H. says, maybe he should shift his focus to Communication instead of Science Inquiry? They lay out a new course list that can help get him there. It involves a mix of Science courses that look like Mrs. F's and are a bit more lecture heavy, but with opportunities to practice communicating science. They also lay out a Speech and Communication Course as well as some English Literature Courses that emphasize verbal presentations. Patrick generally likes the plan, and so they try it out.

A few months later, Patrick has now taught two different lectures in Mrs. F's course, earning him two XL microcredentials, and the admiration of his classmates. None of them even knew he was working ahead of the material—after his lecture, many started asking him for help on the assignments. He ended up forming a "office hours" support during his free period, to help them all study together. Mrs. F also created a L credential in Communication and Leadership to acknowledge this effort. And Mr H also talked to his other

teacher—Patrick's math teacher didn't give a lot of lectures, but offered to let Patrick teach a subject that he knew many students were struggling with in their group projects, earning Patrick another XL Communication badge.

At the end of his sophomore year, Patrick had completed more than half of the required credits for a specialization in Communication, and is well on his way to achieving a specialization in Communication by the time he graduates. Mr. H has been sharing some of the unique assignments that Patrick has achieved with other teachers at the school, and some of them have added similar assignments to their course list, to help other students gain Communication credits through teaching peers. With these additions to the courses, Patrick has even more options available to follow his interests.

11. Conclusion

In the very first chapter, I noted that in games, the rewards are often well-designed, meaningful, intrinsic rewards whereas in education the rewards are often poorly-designed, meaningless, extrinsic rewards. And that in essence is the main point of the book. The gaming industry has been an innovator in the design of playful reward systems. Although education has experimented with several different kinds of rewards, they all often sit in the same, extrinsic reward, design space.

Rewards are a system design problem, and it is hard to fundamentally alter how rewards are used in schools without altering how schools function. The last chapter ended on a high note—pointing to one example of how an educational system can follow the best practices of reward design. There are, I'm sure, other possibilities to pursue in this design space, and this is not a one-size-fits-all scenario. In fact, I'm sure there are schools around the world already following these best practices to various degrees.

Part of the issue is related to how we see assessment in education. We typically view assessment as something disconnected and separate from the learning process—but it's certainly not something disconnected from the system of schools. When assessments become high stakes, we essentially make those assessments a powerful, extrinsic reward. And at that point, to pretend like the assessment is not affecting the learning process, or is not affecting how we structure schools, is to be naive. There's a reason we tend to teach to the test (in the U.S. at least)—it's the best way to get the most important rewards that have been laid out in the education system.

When you stop thinking of high-stakes standardized tests as an extra thing we added to schools to help us measure student learning, and you start thinking of it as one of the most powerful rewards we have placed on the school system, the framing of the problem changes. An unavoidable necessity becomes a design constraint. We can still require that we measure student learning, but do high-stakes, extrinsic rewards really offer the best way to do so?

For me, this issue is important. Schools can be better: they can be more meaningful, more interesting, and more effective places for

students to learn. And they can achieve all of these things at the same time. But it requires change, and it requires learning from fields that have already been innovating in this space. I hope two outcomes come out of this book.

First, I hope this provides a guide for educational developers and teachers. For those making learning software and lessons plans, I hope this offers a way to be more thoughtful about reward and incentive systems. I hope that when those systems are developed, they are a conscious choice to fall in certain parts of the framework offered here. I also hope that in-school and out-of-school educators use this framework to redesign their classroom in whatever ways make sense to them, whether it's their grades, assignments, GPA's, lesson plans, courses, pre-requisites, or whatever else. As much as rewards are omnipresent in the digital system that educational developers are making, the classroom itself is an interactive place, and much of what happens in classrooms goes beyond students interacting on their own with a website. To be truly effective, the rewards in both the digital learning platforms and social interactions need to be rethought.

Second, I hope this provides a guide for educational researchers. We are far past the time where it is ok to be asking "do rewards benefit learning?" or "do games benefit learning?" The details of implementation matter far more than whether something is being used. As a basic first step, research papers in this space, especially ones focused on student engagement or motivation, should at least describe the design of the game and/or reward system in enough detail that meta-analyses can make sense of what design choices are most effective. Even better, the reward system should be described according to some framework, like the one offered in this book. Or best of all, the studies themselves should experiment with reward design, and move away from simply testing reward vs. no reward scenarios.

I'm sure there is nuance I haven't captured in this book. I bet there's some dimensions of reward design that are more important than the ones I included here. I'm almost certain there are best practices with rewards that are left out of this book. But I hope with a thoughtful and concerted effort from educators, educational developers, and researchers, we can have a better understanding in the future of how to design the rewards systems in our schools.

In many parts, this book is a highly theoretical exercise. But we

shouldn't lose sight of what's really at stake. This is about children, and the education that they deserve to have.

Acknowledgements

There are many people to thank over many years, for the delightful conversations, presentations I have listened to, papers I have read, and games I have played, that have all in some way, great or small, contributed to the ideas in this book. To name a few that have been particularly helpful, Scott Price has been a thoughtful colleague throughout this entire process, whether as an interviewee, reviewer, or general thought partner. Naming the number of ideas that were hashed out over a Friday tea and kava session that later appeared in this book would probably be an infinite game—and one day I will hopefully come to understand what that has to do with intrinsic rewards. Lucas Blair's taxonomy was the first that I felt really made sense out of the confusing mishmash that ends up as game design, and Lucas's foundational work was an essential framework I knew I could build upon. I'm indebted to standing on his shoulders from the very start. I appreciate Lucas's thoughtful conversations over the years and the full review of the book. My copy-editor Maggie Jaris has been a huge support—the words just sound better after she goes through them. The staff at ETC Press, Brad King and Drew Davidson, are always helpful and a pleasure to work with.

To name a few others who have been also particularly helpful over the years in forming these ideas: Lee Sheldon, Juan Rubio, Kristen Dicerbo, Barry Joseph, Rafi Santos, Dixie Ching, Daniel Hickey, Yoon Jeon Kim, Liz Owen, Tara Chklovski, Karina Linch, Allisyn Levy, Brendon Trombley, Michael Gi and Will Jordan-Cooley.

And as always, I'd like to thank Hannah for her constant ideas and critiques, and her unending love and support.

Glossary

Abundant Reward – A reward that everyone engaging with the reward system is able to achieve. This dimension is a spectrum with many gradations depending on how abundant or scarce the reward is. For the other forms of this dimension, see Scarce Reward.

Accumulated Discrete Reward – A reward formed of discrete elements that add together into an overall score or total. In these rewards, the sum total of the reward that you have collected is usually the most relevant aspect of the reward. For the other forms of this dimension, see Discrete Rewards and Leveled Rewards. For examples of this reward, see Points and Experience Points.

Achievements – Achievements are the classic digital form of badges, first popularized on the XBox Live system but now incorporated into most games and gaming platforms. They are most commonly a discrete reward achieved for accomplishing some criteria in the game. See also Discrete Reward and Badges.

Adventure games – A genre of games in which some kind of environment must be explored, and that act of exploration works as an intrinsic reward loop. Early in the game industry there were pure adventure games, but today most games are action-adventure games (they involve some kind of battle or fighting process in the act of exploring, with any of the modern Zelda or Mario games being good examples of the genre). If the fighting builds up some kind of leveling system, then the game can be called an Action-Adventure-RPG. These three genres blend into each other in practice, and it's hard to draw exact lines between them.

Assessment – A method used in educational systems to understand what students know. This can have multiple different audiences and

purposes. See also Formative Assessment, Summative Assessment, Playful Assessment or Game-based Assessment.

Badge – One of the classic reward structures, with Levels and Points. Badges are the classic example of a discrete reward—there is a single object that you gain. Badges were first popularized in the Boy Scouts, with physical badges that can be attached to the scout's uniform. See also Discrete Reward.

Completion-Based Reward – A reward that displays if someone completed a task. This is related to the criteria for which the reward is given. Completion-based rewards have a binary criteria, of completed or not completed. For the other form of this dimension, see Measurement-Based Reward. For sub-forms of this reward, see Performance Contingent Completion Reward and Non-Performance Contingent Completion Reward.

Discrete Rewards – A reward that is binary: either one has or doesn't have the reward. For the other forms of this dimension, see Accumulated Discrete Rewards and Leveled Rewards. For an example of this reward, see Badge or Achievement.

Educational badging movement – A movement that originated in the after-school space to award students microcredentials for learning wherever it occurs, in a standard format that can be easily transferred between systems. On a technical level, most badges are OBI compliant, meaning they fit into a technical framework developed by Mozilla which can be easily passed between systems. An OBI compliant badge allows for "meta-data" to be passed with the badge, which can include the student artifacts used to gain the badge. In its most idealized form, which was never fully achieved, these badges could be transferred or used for credit on resumes or in schools. For more info, you can explore the Mozilla Hives, Cities of Learning or Summer of Learning initiatives, which were attempts to realize this vision.

Elo Rating System – One of the earliest forms of points systems used to match and rank players in competitive gaming environments. Originally developed by Arpad Elo, a Hungarian mathematician, to be used for Chess, the Elo rating systems or others ones derived from it form the basis for the majority of today's competitive matchmaking and rating systems. The Elo Rating System is a prime example of an accumulated discrete reward that is measurement-based and scarce.

Expected Reward – A reward which someone knows that they will receive before they receive it. For the other form of this dimension, see Unexpected Reward.

Experience Points – Experience points are the standard currency in many RPGs, and form the basis of the intrinsic reward loop in those games. Like Points, experience points are an accumulated discrete reward, but one that usually builds into a leveling system. Gain enough experience points, and you advance up a level. Some newer games shorten this to "XP". See also Points or RPG.

External Reward – A reward whose value lies in some aspect that is external to the experience. Although this carries a lot of overlap with Extrinsic Rewards, there can also be Intrinsic Rewards in which progressing up loops has some external value, in addition to its internal value (such as the diploma system). For the other form of this dimension, see Internal Reward.

Extrinsic Reward – A reward that has the structure "Do this to get that." Specifically, when you use the nagging 3-year-old test, you end up with a "just because" answer, or a dead end. This is the classic form of rewards that is most often referenced by default, particularly when referencing the negative aspects of rewards. For the other form of this dimension, see Intrinsic Reward.

Fixed Mindset – The view that ability or intelligence is a fixed attribute

that you are born with, and cannot be improved with time or effort. Fixed mindsets tend to be correlated with worse learning outcomes, and are reinforced with feedback that points to someone's ability (e.g. "You are so smart to stack a tower of blocks so high"). See also Growth Mindset.

Formative Assessments – Assessments primarily meant to provide feedback to a learner, to allow them to understand their progress and level of understanding. Formative assessments are often used in or during the learning process. As a general term, they do not need be limited to students in a classroom, though that is their common connotation. For example, an end of unit exam can act as formative assessment to a teacher on how well they are teaching. Or temporary in-game rewards can act as formative assessments to a player on how well they are playing. These are primarily contrasted with Summative Assessments. They are also analogous to Playful Assessments.

Four Freedoms of Play – A theory developed by Scot Osterweil that defines when an experience contains enough openness or freedom to engage someone in play. The four freedoms are freedom to experiment, freedom to fail, freedom to try out identities, and freedom of effort.

Gamification – The application of game-based designs to non-game settings. When defined this way, it is an extremely large term, but in practice it is most commonly used to describe game-inspired marketing techniques. In Education, the term Game-Based Learning is more commonly used to describe effective incorporation of game-based learning principles in the learning process.

Game-Based Assessment – An attempt to use a game as an assessment tool. Game-based assessments have been proposed as both Formative Assessment (as learning games are played they can report on student learning) and Summative Assessments (an item on a standardized test can be a game). There is a current debate on whether

a game that is used for assessment, and particularly for a summative assessment, is still a game—it is argued that because the game in this context is no longer playful (See Four Freedoms of Play), it is no longer a true game.

Game-Based Learning – Using games or game-based principles in learning contexts. This term often refers to a holistic approach to incorporating games into a learning setting in a way that uses effective pedagogy modeled by the game industry. It is most commonly contrasted with Gamification.

GPA – Stands for Grade Point Average, this is a measurement commonly used in high schools and colleges to describe how well a student is doing across all of their courses. It is a cumulative measurement-based reward that is measured relative to peers.

Grades – The most common reward structure used in traditional schools. Generally, grades tend to be permanent, measurement-based, discrete rewards. These three elements taken together in this combination are almost never found in today's games, despite being omnipresent in today's schools.

Grinding – A gaming term that describes the process of repeatedly doing a mundane action to gain experience points and go up levels (which amounts to repeating cycles of an intrinsic reward loop).

Growth Mindset – The view that ability or intelligence is an attribute that can be improved with time or effort. Growth mindsets tend to be correlated with better learning outcomes and increased persistence, and are reinforced with feedback that points to the results of someone's efforts (e.g. "You built a really tall tower out of blocks, look at how tall it got"). See also Fixed Mindset.

God View Simulation Game – A sub-category of Simulation games in

which the player has a zoomed out view of the world, and can take actions to manipulate the system at large. SimCity or Civilization are two examples of the genre.

Guild – A structure in multiplayer games where players can band together to form groups that collaborate and achieve game objectives (and the accompanying rewards) together. Guilds are most common in MMOs.

Hard Unlocking Reward – An Intrinsic Reward in which you are primarily prevented from advancing through further loops by gates set up by the game designer. This is a way to create well-ordered problem solving and scaffolding. For the other form of this dimension, see Soft Unlocking Reward.

Human-Evaluated Reward – A reward whose criteria is evaluated by a person. For the other form of this dimension, see Machine-Evaluated Reward.

Internal Reward – A reward whose value lies in some aspect that is internal to the experience. Although this carries a lot of overlap with Intrinsic Rewards, there can also be Extrinsic Rewards in that they do not loop back to the action, but do convey inward-facing value (such as achievements in games which reward in-game currency). Note that with Pervasive Games, it can sometimes be hard to draw the boundary between what is inside and outside the experience—in these cases, almost everything can be argued to be within the game and therefore an internal reward. For the other form of this dimension, see External Reward.

Intrinsic Reward – A reward that has the structure "Do this to do this more." Specifically, when you use the nagging 3-year-old test, you end up looping back to the original action such that you can say that you are doing the action to do more of the action. This reward is

omnipresent in the video game industry. For the other form of this dimension, see Extrinsic Reward.

Leveled Reward – A reward that exists to various degrees or levels. The levels of this reward can be progressed through, usually in stages. It is distinguished from Accumulated Discrete Rewards in that it's not about a total, but about advancing through sequential stages. It is also distinguished from discrete, measurement based rewards which can give you a rating based on your performance–although the rating can have various degrees or levels, one does not advance through them. For the other forms of this dimension, see Accumulated Discrete Rewards and Discrete Rewards. For an example of this reward, see Levels.

Levels – One of the classic reward structures, with Badges and Points. Levels are the classic example of a leveled reward—there are multiple ranks or levels that can be attained. See also Leveled Reward.

Ludic Assessment – A form of assessment that is meant to be seamlessly embedded in an engaging, playful learning experience. Although embedding assessment in games (e.g. Game-Based Assessment) is one way to achieve a playful assessment, not all game-based assessments are playful assessments, and playful assessments can also be incorporated in non-game contexts. Notably, a playful assessment is incorporated into a game in a way which does not compromise the game's ability to still be a game, unlike some game-based assessments. Many playful assessments have been developed for Maker Education in specific. Playful Assessments are generally Formative Assessments, but they also intend to blur the line between Formative and Summative Assessments. They primarily intend to find a way to measure what students know without disrupting the learning process, which is sometimes called an embedded assessment or stealth assessment: right now traditional formative and summative assessments are both disruptive to the learning process.

Machine-Evaluated Reward – A reward whose criteria is evaluated by

an automated system. For the other form of this dimension, see Human-Evaluated Reward.

Maker Education – A movement in education to teach students valuable skills through learning how to make things, whether that be circuit boards, computer programs, or woodworking. It is analogous in its design principles to Game-Based Learning.

Measurement-Based Reward – A reward that displays how well someone completed a task. This is related to the criteria for which the reward is given. Measurement-based rewards have some measurement of how well someone did at the task, usually in the form of a score or grade. For the other form of this dimension, see Completion-Based Reward.

MMO – Stands for Massively Multiplayer Online, a genre of games that refers to platforms where you can play and interact with thousands to millions of other players. The most common type of MMO fits into the RPG genre.

Non-Performance Contingent Completion Reward – A subform of Completion-Based Rewards in which the completion criteria for the reward is given out simply for showing up, and is unrelated to exhibiting performance or skill. These rewards are sometimes called "participation rewards" and are generally not recommended for common use. See also Performance Contingent Completion Rewards and Completion-Based Rewards.

Permanent Reward – A reward that once gained can never be taken away. This dimension is a spectrum with many gradations depending on how hard or easy it is for a reward to be taken away. For the other form of this dimension, see Temporary Reward.

Performance Contingent Completion Reward – A subform of

Completion-Based Rewards in which the completion criteria for the reward is dependent on exhibiting some kind of competence or skill. By demanding some form of performance, these rewards share some attributes with Measurement-Based Rewards, in that both convey some kind of measurement of skill. For practical purposes, most completion-based rewards that are considered throughout this book are performance contingent rewards, since Non-Performance Contingent Completion Rewards are generally not recommended for common use. See also Non-Performance Contingent Completion Rewards, Completion-Based Rewards, and Measurement Rewards.

Pervasive Games – A term that has multiple debated definitions, but generally means a game which expands its play space, or "Magic Circle", to intersect with the real world, or which defies having a boundary to the game such that it's difficult to determine when the game ends and the non-game real world begins. Live Action Role Plays (LARPS) or Augmented Reality games are two well-referenced examples of pervasive games. These games complicate the idea of Internal Rewards vs. External Rewards, since that taxonomy dimension depends on having a firm definition of where the boundary of the game ends.

Points – One of the classic reward structures, with Badges and Levels. Points are the classic example of an accumulated discrete reward—points are usually tallied towards some kind of total. See also Experience Points, Accumulated Discrete Reward.

Puzzle Games – A genre of games that involves solving discrete puzzles, usually to advance up levels to harder puzzles. The act of advancing through levels can form the basis for an intrinsic reward loop. Tetris and Candy Crush are two examples of the genre.

Qualitative Reward – A reward that is evaluated along qualitative criteria. For the other form of this dimension, see Quantitative Reward.

Quantitative Reward – A reward that is evaluated along quantitative criteria. For the other form of this dimension, see Qualitative Reward.

Real-Time Strategy Games – A genre of games that involves collecting and allocating resources to control a map and/or defeat an opponent. It is a subgenre of Strategy Games, and is contrasted with "Turn-Based Strategy Games" such that you and your opponent are acting at the same time as time moves along continuously, rather than taking turns. The act of collecting resources to develop more units to expand your reach to gather even more resources forms an intrinsic reward loop. Warcraft is an example of this genre.

Reward – Any kind of goal-setting, tracking or advancement layer added to an experience or activity. This is meant to be an inclusive definition of many types of rewards and reward-like things, thus noting that designing rewards is an activity with deep historical roots that crosses many realms of human life.

RPG – Stands for "Role-Playing Games", or a genre of games that commonly refers to video games that feature characters which go up levels by gaining experience points, from either defeating monsters or completing quests. This characteristic of the genre forms an intrinsic reward loop. Final Fantasy is an example of this genre.

Scarce Reward – A reward that only a few people engaging with the reward system are able to achieve. This dimension is a spectrum with many gradations depending on how abundant or scarce the reward is. For the other form of this dimension, see Abundant Reward.

Simulation Games – A genre of games that involves interacting with some kind of simulation of a real system. These can be of many varieties and cross genres, such as puzzle games or strategy games. The simulation can be of a physical system, like laws of gravity, or of a

historical situation, like World War 2. See also God-View Simulation Game.

Soft Unlocking Reward – An Intrinsic Reward in which you are primarily prevented from advancing through further loops by your own skill or strength. This is often considered an elegant design that effectively calls attention to how one is gaining competence. For the other form of this dimension, see Hard Unlocking Reward.

Speedrun – A specific category of gameplay in which a player attempts to beat a game as fast as possible. Speedrun can either be a verb, referring to the act of playing the game with this self-selected goal, or be a noun, referring to the video evidence of the gameplay. Speedruns are often ranked by time of completion in a meta-reward structure.

Standardized Testing – A classic form of a Summative Assessment, that intends to measure what students know according to a set of widely accepted academic standards. These tests typically occur once a year and have high stakes for students attached to them. As such, they come off as quite judgmental, and suffer the negative consequences associated with judgmental, permanent, measurement-based rewards, such as demotivating students and fostering Fixed Mindsets. See also Summative Assessment.

Strategy Games – A genre of games that involves making strategic choices to either defeat a scenario or defeat an opponent. These are often primarily contrasted with Action games, which rely on reflexes rather than carefully considered choices. This is a large genre with several sub-genres, such as Real-Time Strategy Games and Tower Defense games. Many but not all strategic games have some kind of Intrinsic Reward loop around playing well to gain more power to play even better, until you've gained so much power you can win the game. Chess and Go are both very old examples of this genre.

Summative Assessment – An assessment which intends to make a judgment about how much knowledge a student has. The primary audience for these assessments are the value providers in educational systems, or those that need to give out educational credentials like passing courses, getting diplomas, or accepting students into college. These assessments usually occur after the learning process is over, and accordingly are not meant to help students learn better. They are primarily contrasted with Formative Assessments. See also Standardized Testing for an example of a summative assessment.

Temporary Reward – A reward that can be lost, taken away, or spent after being gained. This dimension is a spectrum with many gradations depending on how hard or easy it is for a reward to be taken away. For the other form of this dimension, see Permanent Reward.

Tower Defense Games – A genre of games that involves placing towers along a path to shoot down incoming enemies that march towards you. The goal is to defend your base by preventing any enemies from reaching it. The act of using resources gained from defeating enemies to upgrade your defenses and defeat more enemies forms an intrinsic reward loop. Kingdom Rush is an example of this genre. This is considered a sub-genre of Real-Time Strategy Games.

Unexpected Reward – A reward which is awarded to someone as a surprise. There are a few different dimensions of unexpected rewards- at the most extreme someone might not know that reward even exists until they receive it. In less severe unexpected rewards, someone might know that a reward exists, but might not know when, where or how they might gain that reward, leaving some surprise in the act of gaining the reward. For the other form of this dimension, see Expected Reward.

Unlockables – In many games, powerups are unlocked over the course of playing the game (most commonly in action-adventure games) that confer some kind of benefit on the player. These unlockables act as a sort of discrete reward with internal value. When cued into an Intrinsic

Reward structure, unlockables can form a Hard Unlocking Structure. See also Adventure Games.

Notes

Chapter 1

1. I took a little liberty with the phrasing for simplicity's sake, which in its original form is "Do this and you'll get that." Kohn, A. *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes.* Boston: Houghton Mifflin, 1993.
2. Square. *Final Fantasy*. 1987. Nintendo Entertainment System.

Chapter 2

1. David, R. *The Ancient Egyptians: Beliefs and Practices*. Brighton: Sussex Academic Press, 1998.
2. Boy Scouts of America. "History." Accessed March 8, 2020. https://www.scouting.org/programs/venturing/about-venturing/history/.
3. Boy Scouts of America. *Handbook for Boys*. New York: Doubleday, 1911. 23-46.
4. Boy Scouts of America. "Guide to Advancement 2019." Accessed March 8, 2020. https://www.scouting.org/resources/guide-to-advancement/.
5. Boy Scouts of America. "Merit Badges." Accessed March 8, 2020. https://www.scouting.org/programs/boy-scouts/advancement-and-awards/merit-badges/.
6. Boy Scouts of America. "The Merit Badge Program." Accessed March 8, 2020. https://www.scouting.org/resources/guide-to-advancement/the-merit-badge-program/
7. Boy Scouts of America. "Mechanics of Advancement in Scouts BSA." Accessed March 8, 2020. https://www.scouting.org/resources/guide-to-advancement/mechanics-of-advancement/boy-and-varsity/.

8. Michelin Guide. “About us.” Accessed March 8, 2020. https://guide.michelin.com/en/about-us.
9. The “9 Pin Zhi” system was referenced in a book dating to the 3rd Century AD, called the *Classic of Arts* and written by Handan Chun.
10. Just, T. *U.S. Chess Federation’s Official Rules of Chess 6th Edition*. New York: Random House, 2014. You can also see: Wikipedia. “Chess Rating System.” Accessed March 8, 2020. https://en.wikipedia.org/wiki/Chess_rating_system.
11. Butts, A. *Scrabble*. 1933. Competitive Scrabble player ratings can be looked up at https://www.cross-tables.com/.
12. Valve and Hidden Path Entertainment. *Counter Strike: Global Offensive*. 2012. Steam, Xbox 360.
13. Blizzard Entertainment. *Overwatch*. 2016. Microsoft Windows, Xbox One, Playstation 4.
14. Nameco. *Pac-man*. 1980. Arcades.

Chapter 3

1. A readable version of Blair’s taxonomy is: Blair, L. “The Cake Is Not a Lie: How to Design Effective Achievements.” Gamasutra, April 27, 2011. https://www.gamasutra.com/view/feature/134729/the_cake_is_not_a_lie_how_to_.php. You can also go to Blair’s Ph.D. thesis at: Blair, L. “The use of video game achievements to enhance player performance, self-efficacy, and motivation.” Ph.D thesis, Central Florida University, 2011. In Electronic Theses and Dissertations, 2004-2019: 1827. https://stars.library.ucf.edu/etd/1827
2. Rovio Entertainment. *Angry Birds*. 2009. iOS.
3. See Chapter 3, note 1, https://www.gamasutra.com/view/feature/134729/the_cake_is_not_a_lie_how_to_.php.
4. Blizzard Entertainment. *World of Warcraft*. 2004. Windows, MacOS.
5. Herbert, D “peppy”. *osu!stream*. 2015. iOS.
6. Halfbrick Studios. *Jetpack Joyride*. 2011. iOS.
7. Riot Games. *League of Legends*. 2009. Windows and MacOS.
8. Capcom. *Megaman 2*. 1988. NES.

9. See Chapter 3, note 1, https://www.gamasutra.com/view/feature/134756/the_cake_is_not_a_lie_how_to_.php.
10. Csikszentmihalyi, M. *Flow: The Psychology of Optimal Experience*. New York, NY: Harper Perennial, 1991.
11. See Chapter 1, note 1, pg 204-205.
12. For one example, this study (which is discussed in depth in the next chapter) highlights the importance of the feedback messaging when a reward is given. O'Rourke, E., K. Haimovitz, C. Ballweber, C. Dweck, and Z. Popović. "Brain points: a growth mindset incentive structure boosts persistence in an educational game." *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014): 3339-3348.
13. Konami, Bemani. *Dance Dance Revolution: Extreme 2*. 2005. Playstation 2. Though most of the notes from DDR remain true across the series, minor differences do exist from game to game. I am basing my notes on the version with which I am most familiar, which is the 2005 release for Playstation 2.
14. International Skating Union. "ISU Figure Skating Media Guide 2019/20." Accessed May 2020. https://www.isu.org/media-centre/guides/media.
15. See Chapter 3, note 1, https://www.gamasutra.com/view/feature/134744/the_cake_is_not_a_lie_how_to_.php.
16. PopCap Games. *Alchemy*. 2001. Windows, later for other platforms.
17. Armour Games. *Achievement Unlocked*. 2008. Browser-based. Accessed at https://armorgames.com/play/2893/achievement-unlocked.
18. Retro Studios, Nintendo. *Metroid Prime*. 2002. GameCube.
19. Valve. *Team Fortress 2*. 2007. Steam.
20. Joseph, B. "My Beef with Badges." *Connected Learning Alliance* (blog). March 13, 2014. clalliance.org/blog/my-beef-with-badges/
21. Knight, E. "More Beefs." *World of E's* (blog). 2014. erinknight.com/post/82103788980/more-beefs/embed
22. Casilli, C. "The Myth of the Lightweight Badge."*Persona* (blog). February 2, 2014. carlacasilli.wordpress.com/2014/02/26/the-myth-of-the-lightweight-badge/.
23. Firaxis Games. *Civilization III*. 2001. Windows and Mac OS.
24. Square. *Final Fantasy 7*. 1997. Playstation.

Chapter 4

1. Deci, E. L., R. Koestner., and R. M. Ryan. “A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation.” *Psychological Bulletin* 125, no. 6 (1999): 627-68.
2. Deci, E. L., R. Koestner, and R. M. Ryan. “Extrinsic rewards and intrinsic motivation in education: Reconsidered once again.” *Review of Educational Research* 71, no. 1 (2001): 1-27.
3. Abramovich, S., C. Schunn, and R. Higashi. (2013). “Are badges useful in education?: It depends upon the type of badge and expertise of learner.” *Educational Technology Research and Development* 61, no. 2 (2013): 217-232.
4. Filsecker, M., and D. T. Hickey. “A multilevel analysis of the effects of external rewards on elementary students’ motivation, engagement and learning in an educational game.” *Computers & Education* 75 (2014): 136-48.
5. Barab S., T. Michael, T. Dodge, R. Carteaux Jr., and T. Hakan. *Quest Atlantis*. 2005. PC.
6. See Chapter 3, note 10.
7. For four examples:
   Vogel, J. J., D. S. Vogel, J. Canon-Bowers, C. A. Bowers, K. Muse, and M. Wright. “Computer gaming and interactive simulations for learning: A meta-analysis.” *Journal of Educational Computing Research*, 34, no. 3 (2006): 229-243.
   Ke, F. “A qualitative meta-analysis of computer games as learning tools.” In *Handbook of Research on Effective Electronic Gaming in Education,* edited by R. Ferdig, 1-32. New York: IGI Global, 2009.
   Sitzmann, T. “A meta-analytic examination of the instructional effectiveness of computer-based simulation games.” *Personnel Psychology*, 64 (2011): 489-528.
   Wouters, P., C. van Nimwegen, H. van Oostendorp, and E. D. van der Spek. “A meta-analysis of the cognitive and motivational effects of serious games.” *Journal of Educational Psychology* 105 (2013): 249-265.
8. For two examples:
   Young, M. F., S. Slota, A. B. Cutter, G. Jalette, G. Mullin, B. Lai, Z. Simeoni, M. Tran, and M. Yukhymenko. “Our Princess Is in Another

Castle: A Review of Trends in Serious Gaming for Education." *Review of Educational Research March* 82, no. 1 (2012): 61-89. Zhonggen, Y.. "A Meta-Analysis of Use of Serious Games in Education over a Decade." *International Journal of Computer Games Technology* (2019).
9. Clark, D. B., E. E. Tanner-Smith, and S. S. Killingsworth. "Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis." *Review of Educational Research* 86, no. 1 (2016): 79-122.
10. Dichev, C. and D. Dicheva. "Gamifying education: what is known, what is believed and what remains uncertain: a critical review." *International Journal of Educational Technology in Higher Education* 14, no. 9 (2017): 1:36.

Chapter 5

1. See Chapter 1, note 2. Here's a Final Fantasy video of a basic battle scene, with rewards given after each battle: https://www.youtube.com/watch?v=0wi6rZDciQo. The first leveling up occurs around 10:20.
2. Square Enix and Jupiter, 2007. *The World Ends with You*. 2007. Nintendo DS. Here's The World Ends with You trailer, showing snippets of battle sequences throughout the demo https://www.youtube.com/watch?v=mMPfZBIK47c. 1:35 shows some view of inventory management with pins and items.
3. Hemisphere Games. *Osmos*. 2009. PC and app-based distribution. Here is an Osmos trailer showing core gameplay: https://www.youtube.com/watch?v=jrzhlTn1_ds.
4. Supercell. *Clash Royale*. 2016. All mobile platforms. The clan feature was introduced as a later update to the game in 2018.
5. See Chapter 3, note 16. The Gun Mettle campaign was introduced as a later update to the game in 2015.
6. King. *Candy Crush Saga*. 2012. Facebook and app-based distribution.
7. PopCap Games. *Bejeweled*. 2001. Originally for browsers.
8. Nintendo. *Tetris*. 1989. NES. Tetris has a complicated history, and has appeared on many different platforms throughout history. The

original version can be attributed to a single developer, Alexey Pajitnov, although I cited the version I played and have the most experience with. A sample gameplay of the original NES version can be found at https://www.youtube.com/watch?v=CvUK-YWYcaE. The level up from the level 1 to level 2 occurs at 1:29.

9. See Chapter 3, note 7. A complete Mega Man 2 playthrough (in speed run form) can be found at https://www.youtube.com/watch?v=KLrmg1A6EKY. At 2:07 you can see the first boss being defeated and the first power being gained.
10. See Chapter 3, note 15. A complete Metroid Prime playthrough can be found at https://www.youtube.com/watch?v=xzTotHdersg. The main game occurs at around 15:30—from there you can watch as more areas of the map are unlocked as the player explores. At 21:30, the first powerup is acquired, which unlocks new areas of the map.
11. Blizzard Entertainment. *Warcraft: Orcs & Humans*. 1994. PC. Blizzard has made its original games available for purchase and play: https://www.blizzard.com/en-us/games/legacy/. The two sequels that were released, Warcraft II and Warcraft III for short, also contain the same basic reward loop as the original, but with additional added complexity.
12. Ironhide Game Studio. *Kingdom Rush*. 2011. Browser and iPad, later for multiple app-based and console-based distribution markets. There have been sequels produced after the original, which also contain the same basic reward loop.
13. Square. *Final Fantasy X*. 2001. Playstation 2. The game has recently been remastered, and the remastered version is easily available to play, but contains several changes. Most important here are the changes to the Sphere Grid, which affect how you choose to level up your characters. There is the option in the remastered version to use the "original" sphere grid or the new "expert" sphere grid. Both allow interesting choice in how to level up characters, and illustrate the point in the text.
14. Maxis. *SimCity 2000*. 1993. PC. Here's a SimCity 2000 video of basic gameplay: https://www.youtube.com/watch?v=7zY-hN2b2rg.
15. Zinga. *Farmville*. Zynga, 2009. Facebook plus other distribution markets.
16. Bethesda Game Studios. *Fallout Shelter*. 2015. Multiple app-based

and console-based distribution markets.

Chapter 6

1. Nicholson describes the danger of designing gamification systems for the organization, rather than the player, which feels like another phrasing of this same point. Nicholson, S. "A RECIPE for Meaningful Gamification." In *Gamification in Education and Business*, edited by L. Wood and T. Reiners, 1-20. New York: Springer, 2014.
2. Ryan, R. M., and E. L. Deci. *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York: Guilford Press, 2017. See pg. 530.

Chapter 7

1. See Chapter 6, note 2.
2. See Chapter 4, notes 1 and 2.
3. Nintendo. *Wii Fit*. 2007. Nintendo Wii.
4. You can find many examples by searching for "[Name of game] speedrun", but websites like www.speedrun.com/ track the best speedruns for each game.
5. Team Cherry. *Hollow Knight*. 2017. PCs initially, later on consoles.
6. See Chapter 2, note 14.
7. See Chapter 5, note 15.
8. Meier, S. "Interesting Decision." March 2012. Game Developers Conference (GDC), San Francisco, CA. Video Recording, 01:00:40. www.gdcvault.com/play/1015756/Interesting.
9. Nintendo. *Mario Kart*. 1992. Super Nintendo Entertainment System.
10. Osterweil, S. "The Four Freedoms of Play and Learning." YouTube Video, 00:05:13. https://youtu.be/-66lm9T4bNk.
11. See Chapter 5, note 12.
12. See Chapter 5, note 8.

13. See Chapter 5, note 4.
14. See Chapter 4, notes 1.

Chapter 8

1. GameLab. *Gamestar Mechanic*. 2010. Browser-based.
2. Price, S., interview with the author. February 8, 2019. Edited transcript here http://kmiklasz.blogspot.com/2019/02/reward-structures-in-gamestar-mechanic.html.
3. See Chapter 3, note 17.
4. Darvasi, P., interview with the author. August 21, 2019. Edited transcript http://kmiklasz.blogspot.com/2019/08/rewards-interview-with-paul-darvasi-big.html.
5. For one such deep dive, see Darvasi, P. "The Ward Game: how McMurphy, McLuhan, and MacGyver might free us from McEducation." In *Teacher Pioneers, Visions from the edge of the map*, edited by C. Williams-Peirce, 70-105. Pittsburgh: ETC Press, 2016.
6. Sheldon, L. *The Multiplayer Classroom: Designing Coursework as a Game*. Boston: Course Technology PTR, 2012.
7. See Chapter 3, note 3.
8. See Chapter 8, note 6, pg 43.
9. See Chapter 8, note 6, pg 146.
10. See Chapter 8, note 6, pg 149-150.
11. Blizzard Entertainment. *Hearthstone.* Blizzard Entertainment, 2014. Microsoft Windows, macOS, iOS, Android.
12. Nintendo. *Splatoon 2*. 2017. Nintendo Switch.
13. See Chapter 4, note 4.

Chapter 9

1. Derryberry, A., D. Everhart, and E. Knight. "Badges and Competencies: New Currency for Professional Credentials." In *Digital Badges and Education: Trends, Issues and Cases*, edited

by. L. Y. Muilenburg and Z. L. Berge. 12-20. New York: Routledge, 2016.
2. See Chapter 6, note 1.
3. Conversations about "ludic assessment" have also occurred under the name "playful assessment." See Miklasz, K. "Playful Assessment Explained!." BrainPOP Educators. 2016. Video, 11:02. https://educators.brainpop.com/video/playful-assessment-explained/. and Kim, Y. J. and L. Rosenheck. "Tch Talks 28: Playful Assessment." Teaching Channel. 2018. https://learn.teachingchannel.com/blog/2018/07/20/tch-talks-28-playful-assessment.
4. See Chapter 3, notes 17, 18, and 19.
5. WeWantToKnow AS. *DragonBox*. 2012. iOS.
6. National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press, 2012. See especially Chapter 4.
7. Dweck, C. S. *Mindset: the new psychology of success*. New York: Random House, 2006.
8. See Chapter 1, note 1.
9. See Chapter 4, note 1.
10. See Chapter 4, note 1.
11. Hasbro. *Apples to Apples*. 1999. Board Game.
12. See Chapter 3, note 20.

Chapter 10

1. Several other authors have pointed to how games are fun specifically because they help you learn, such as: Gee, J. *Good Video Games + Good Learning*. New York: Peter Lang, 2007. and Koster, R. *A Theory of Fun for Game Design*. Sebastopol: O'Reilly Media, 2014.
2. See Scott Price's ignite talk at https://www.youtube.com/watch?v=nC2nyqb4eV8, minutes 24:47-30:02 Price, S. "Teach Like a Game Master, Design like a Teacher." Connected Learning Summit, Cambridge, MA, August 2018. and see Darvasi, P. "Five

best practices teachers can learn from dungeon masters." KQED. MindShift, Accessed February 6, 2020. https://www.kqed.org/mindshift/53553/five-best-practices-teachers-can-learn-from-dungeon-masters.

About the ETC Press

The ETC Press was founded in 2005 under the direction of Dr. Drew Davidson, the Director of Carnegie Mellon University's Entertainment Technology Center (ETC), as an open access, digital-first publishing house.

What does all that mean?

The ETC Press publishes three types of work:peer-reviewed work (research-based books, textbooks, academic journals, conference proceedings), general audience work (trade nonfiction, singles, Well Played singles), and research and white papers.

The common tie for all of these is a focus on issues related to entertainment technologies as they are applied across a variety of fields.

Our authors come from a range of backgrounds. Some are traditional academics. Some are practitioners. And some work in between. What ties them all together is their ability to write about the impact of emerging technologies and their significance in society.

To distinguish our books, the ETC Press has five imprints:

- **ETC Press:** our traditional academic and peer-reviewed publications;
- **ETC Press: Single:** our short "why it matters" books that are roughly 8,000-25,000 words;
- **ETC Press: Signature:** our special projects, trade books, and other curated works that exemplify the best work being done;
- **ETC Press: Report:** our white papers and reports produced by practitioners or academic researchers working in conjunction with partners; and
- **ETC Press: Student:** our work with undergraduate and graduate students.

In keeping with that mission, the ETC Press uses emerging technologies to design all of our books and Lulu, an on-demand publisher, to distribute our e-books and print books through all the

major retail chains, such as Amazon, Barnes & Noble, Kobo, and Apple, and we work with The Game Crafter to produce tabletop games.

We don't carry an inventory ourselves. Instead, each print book is created when somebody buys a copy.

Since the ETC Press is an open-access publisher, every book, journal, and proceeding is available as a free download. We're most interested in the sharing and spreading of ideas. We also have an agreement with the Association for Computing Machinery (ACM) to list ETC Press publications in the ACM Digital Library.

Authors retain ownership of their intellectual property. We release all of our books, journals, and proceedings under one of two Creative Commons licenses:

- **Attribution-NoDerivativeWorks-NonCommercial:** This license allows for published works to remain intact, but versions can be created; or
- **Attribution-NonCommercial-ShareAlike:** This license allows for authors to retain editorial control of their creations while also encouraging readers to collaboratively rewrite content.

This is definitely an experiment in the notion of publishing, and we invite people to participate. We are exploring what it means to "publish" across multiple media and multiple versions. We believe this is the future of publication, bridging virtual and physical media with fluid versions of publications as well as enabling the creative blurring of what constitutes reading and writing.