Finding the beat: Cycles of expertise in rhythm games

Kevin Miklasz, BrainPOP

Abstract: I have experienced an interesting puzzle when playing rhythm games: gameplay on a song usually proceeds from being so complex that I do not even know what I am doing wrong, to being so fluent that I can play the song without conscious effort. Thus, I get better at the game without knowing how that improvement occurs or what it looks like. To better understand the development of my own rhythm game literacy, I downloaded four songs on the popular rhythm game *Jukebeat*, and recorded all of my gameplay on those four songs over a period of nine months. From this recording I observed how quantifiable measures of my performance and improvement in the positioning of my fingers and compared with my self-perceived gameplay skill. Along with observations and reflection of my gameplay recordings, I also present a theoretical framework for understanding the development of rhythm game literacy.

Introduction

Literacy is a fundamental aspect to learning. Literacy takes many forms, but generally involves interpreting meaning from sensory inputs. The process of interpreting meaning can be quite complex: it often involves more than just knowing definitions, but rather having a situational or systemic knowledge (Gee, 2007). Literacy thus involves "embodied intelligence," or having a well developed understanding of the contextual nature of symbols developed through actions, or embodied experiences. Embodied intelligence is built up from multiple sessions of practicing and reflecting on that practice, or what can be referred to as cycles of expertise (Gee, 2007). Squire goes so far as to describe game literacy as particularly embodied in the interactivity of a game, and thus is most directly represented as performance expertise (Squire, 2008).

Rhythm games are often considered to involve practice with musical literacy. Musical literacy generally involves understanding the timing of notes in meter and beats as described in Lerdahl and Jackendoff's Generative Theory of Tonal Music (Lerdahl and Jackendoff, 1996). Professionally trained musicians are known to perform better than non-musicians in understanding and interpreting the timing of both visual and auditory signals (Ramsayer et al., 2012). Of additional interest, coupling physical movements to beats have been found to increase musical literacy (Manning & Shutz, 2013), indicating that movement is useful in developing musical expertise and there is a potential use for rhythm games to develop genuine musical literacy. On the other hand, rhythm games do not offer an exact parallel to the way music is performed (Miller, 2009; Arsenault, 2008) and evidence that skills transfer from rhythm games to general musical literacy has not been found (Gaydos, 2010). Emergent timing and event timing are recognized as two distinct skills, the former involving the coordination of fluid and continuous movements and belonging to the realm of the athlete, and the latter involving discrete and regular events and belonging to the realm of the musician (Janzen et al., 2014). In this light, games are more similar to sports than music performances, potentially explaining why attempts to show increases in event-timing musical literacy from gameplay have been unsuccessful. This paper will mostly avoid this tricky issue by recognizing that game-based musical literacy is increasing (i.e. a player's scores in rhythm games increases over time with practice), and concern itself with understanding how this game-specific literacy develops—whether or not a more generic and transferable music literacy is also developing. In this way, rhythm games are simply treated as a convenient case study for understanding the development of a specific, context-dependant expertise.

For that purpose, rhythm games are a particularly useful case study for several reasons. First, rhythm games have clearly defined cycles of practice, namely replaying songs. Second, the game offers a clear mode of performance to express the mastery gained, thus providing an embedded assessment of mastery (Shute, 2013). Third, a player's score in a rhythm game can be considered a close analogy to a quantifiable measure of literacy.

This well-played example plans to investigate a simple issue- how does literacy expertise develop over repeated cycles of gameplay? To answer this general question, the author focuses specifically on rhythm games. I have noticed from my gameplay that the development of such literacy seems to be far from a regular, linear process. It involved the development of several, functionally separate literacies, which each seem to develop in jumps and spurts. The end result is that gameplay on a song usually proceeds from being so complex that I do not even know what I am doing wrong (i.e. lack of literacy) to being so fluent that I can play the song without conscious effort (full literacy through embodied intelligence). The transition between these two states happens so subtly that I am not quite sure when and how the transition occurs, nor am I able describe what exactly changed in my gameplay to cause this increase in performance. This well-played session is a conscious investigation into how exactly this unconscious transition from low to high literacy occurs in rhythm games. This is achieved through video recordings

and journaling throughout several months of my mastery of four new songs.

For the GLS presentation, I present the theoretical framework, clips from the actual video recordings of gameplay over time, and a close analysis of my journals and videos. Additionally, I will conduct a live playthrough for the audience of two songs- one in which I have undergone many cycles of practice and developed a high level of song-specific expertise, and one which I will download shortly before GLS and will never have played before. The audience can compare and contrast playthroughs of the two songs to offer any additional insights into how cycles of expertise affect the difference in my gameplay of those songs.

Methods

Author background

I am relatively experienced in several forms of the rhythm games genre, including *Dance Dance Revolution*, *Elite Beat Agents*, *Guitar Hero*, *Rock Band*, *Osu Stream*, and *Jukebeat*. In all of these games, I progressed from a beginner to some moderate to high level of expertise. This study focuses on *Jukebeat* which proved one of the easier games to record and analyze, but I believe that the patterns described for *Jukebeat* likely also hold true for other rhythm games. I can currently pass most songs on *Jukebeat* at a level 9 difficulty, but have yet to pass any songs on level 10 difficulty. Thus I am at an advanced level of literacy in *Jukebeat*, but still have room to grow in expertise.

Description of *Jukebeat*

Jukebeat (Konami, 2011) is a freemium game available on the iPad and iPhone. It comes preloaded with three playable songs, but has many “4 song packs” available for purchase through the in game store. Each song has three levels of difficulty, and each level has a further rating from 1-10 allowing for a more absolute metric of difficulty that can compare various songs to each other.

Songs are played on a 4x4 grid of square buttons. Players press one or more of the buttons in sequences as a song plays. A visual cue appears about one second before a player is suppose to strike the button, cueing them into the intended timing (Figure 1). One of three feedback animations occurs after a player hits a button, to indicate whether the player hit the note in perfect, near perfect, or far from perfect timing (Figure 1). Players are not penalized for taps on empty buttons. Notes can occur individually or in groups that must be pressed together. A typical *Jukebeat* song on a high level of difficulty involves coordinating the movement between 3-4 fingers on each hand.

Like most rhythm games, the point system awards more points for the closer you are to the beat, but also weights the score for each note by your “combo,” or the number of consecutive prior beats hit in perfect or near perfect timing. Thus the score accounts for both individual accuracy on a note, and repeated accuracy across notes. Individual notes are worth different points on different songs, such that the more notes a song has, the less points each note is worth to ensure the maximum possible score on any song is 1,000,000. Thus the system is weighted in such a way that scores between different songs feel comparable. The game also awards a letter grade for various final scores: less than 700,000 is an F, above 700,000 is a C, 800,000 is a B, 850,000 is an A, 900,000 is an S, 950,000 is SS, and 1,000,000 is SSS. Typically the only way to score an SS or higher is to get a full combo on a song.

I rarely get scores above an A on any particular song- at the point where I can regularly achieve A’s on a song, the song tends to lose my interest. I am most engaged and interested in a song when trying to move my score from an F to a B.

Setup and analysis

In June 2014, I downloaded four new songs from *Jukebeat*’s store. These were songs I had never heard nor played before. The songs also captured the range of my current skill levels: three were ranked at level 9, and one was ranked at level 10. Based on my current expertise level, I would expect to master three of the songs and struggle with the final one.

I built a device to record my hands and the screen as I played *Jukebeat*. I did not use screencapture because I was interested in the motion of my hands and fingers in particular, and if they might show any subtle changes over the cycles of practice. I recorded every playthrough of the four downloaded songs, over a period of nine months. I did play more than just those four songs, but only recorded playthroughs of those songs. I continued my natural play cycles with the game, which usually involved playing the game intensely for a few days to weeks, then putting it down in favor of other games for a few weeks to months, then returning again.

I analyzed several features of my play in the recorded videos. These include easily quantifiable things like total song score and scores during particularly challenging sequences of notes. But it also includes tactile information about which fingers were predominantly used both during the song as a whole and during particularly challenging sequences. I also observed what caused me to miss sequences of beats, whether I was moving my fingers in the wrong positions, or moving them at the wrong timing.

Theory

Categorizing literacies

In thinking about my gameplay over the years, I believe that there are three primary skills involved in doing well at any rhythm game. The first is literacy, or making sense of stimuli acquired through senses. Second is coordination, which involves translating inputs into outputs. It still involves a sense-making activity, but involves coding inputs from one or more sources into a suitable output, usually muscle movement. “Muscle memory” is another word for this. Third is physical finesse, and involves the physical action required to complete the desired output. Conceptualizing the motion that you would like to take is not the same as actually achieving that motion and the desired end result, which describes the difference between coordination and finesse.

Although these three skills were described based on reflections of my own gameplay, there are clear parallels with the conceptions of musical literacy described in the introduction, particularly in the conception of emergent timing skills of professional athletes (Janzen et al., 2014). Also, Squire’s concept of game literacy as being a performance expertise (Squire, 2008) encompasses all three of these skills as part of a single game literacy, as all are required to exhibit performance in the game.

In rhythm games, I believe that for practical purposes, we are not being stretched to the limits of finesse. Besides *Dance Dance Revolution*, most rhythm games do not require a vast amount of physical exertion, and do not require movements that the average person is incapable of performing. What we lack is the coordination, or the muscle memory, to execute these movements fast enough, or the literacy required to accurately read visual and audio cues.

Literacies in *Jukebeat*

I will now refer to *Jukebeat* in particular, though these same literacies would likely apply to most rhythm games. In *Jukebeat*, there are three primary skills that must be perfected to advance one’s performance: Visual literacy (VL), Tactile Coordination (TC), and Audio Literacy (AL).

1) *Visual literacy* is about being able to make sense of note patterns as they come up. It is the earliest skill learned in the game—you need to be able to understand notes before being able to respond to them. This literacy has different levels of competency- the notes become harder to read at more difficult levels both because there are more notes and because they move faster, requiring you to improve your VL. Once VL is attained at a particular level of difficulty though, it is easily transferred to other songs at that same difficulty level.

2) *Tactile coordination* is the reflexes and finger agility required to respond to particular note patterns. It involves making sense of visual and audio clues to produce muscle movements. TC can actually be viewed as a series of different minute skills, rather than one big skill. Being able to hit different types of sequential patterns requires different motor actions, and therefore each sequence requires its own TC. This is akin to being able to read the letter “A”, but not yet understanding the letter “B.” Additionally, being able to string multiple sequences together is an additional skill, just as being able to recognize the letter “A” is different from being able to read the word “ant”. Understanding where one finger leaves off in one sequence and how to connect it to the first note of the next sequence is an additional level of TC. Based on my experience, TC’s seem to be highly transferable- a discrete TC gained for one sequence in one song readily applies when that same sequence is played in other songs.

3) *Audio literacy* is about being able to read the metrical structure of a song. This is actually one of the last skills needed to play the game well, despite being the one most commonly associated with the game genre. At higher level songs, VL informs where you should move and AL informs when you should move (with your ability to actually move in the desired sequences determined by TC). AL is on the one hand extremely song specific. AL gained for lower levels on one song often travels up to and improves performance on higher levels of that same song. In general though, AL is its own higher-order skills that develops over time across many songs, and can allow you to grok the beats of new songs faster. But, a part of it is always song specific, and your song-specific AL will always improve the more that you practice a particular song, no matter how much of an expert you are.

Cycles of Expertise

Rhythm games offer multiple opportunities for repeated cycles of practice. First, any given song repeats certain sequences of notes throughout the song, which gives you a chance within a song to practice that sequence multiple times. Second, the songs themselves are clearly meant to be replayed, giving the opportunity repeat that song multiple times. Third, songs at equal difficulty offer opportunity to practice playing at that difficulty in multiple ways, with equally challenging but different note sequences.

From my experience, I would suggest that there are four distinct levels of expertise that a player progresses through the more that they play a rhythm game:

Level 1: When you first start playing rhythm games, you really are just practicing VL. You play the songs better when you use your VL to learn both when and where to hit a beat. TC skills are pretty minimal, there are not really even sequences yet, the notes are played so far apart that each motion to hit a beat feels separate from the next motion. TC at this point just involves getting the timing of single notes right. The AL skills are pretty nonexistent and are not really even being practiced. Although notes are being played on a beat, they are being played so slowly that you induce their timing visually more than auditorily.

Level 2: Once your VL becomes somewhat advanced, you can progress to the next level of songs, where the idea of sequences, or series of notes played on the half or quarter beat, becomes prominent. This challenges both your VL and TC, as you now need to think about several motions happening in close repetition. Muscle memory of sequences starts to be built, and TC is undergoing the most improvement at this stage (though VL is still becoming more advanced too). At this point, AL is still irrelevant, as the sequences happen in enough isolation from each other that VL still informs the timing of the sequence more than AL, and the sequences are short enough that AL is not needed to keep you on beat.

Level 3: Once your TC has mastered basic 3-note sequences, you can progress to songs where sequences become faster and longer. Smaller sequences previously learned must be chained together, in sequences that can be 5-15 notes long. The VL task becomes more challenging, and less about reading individual notes as much as seeing patterns of sequences and letting your muscle memory move from one sequence to the next. You no longer see the notes as individual beats, but you read them visually as sequences. TC is constantly strained, and fingers will actually begin to tire over the course of several songs, building up finesse to some degree. These long repetitions of notes, and the increased speed of the songs and the speed at which the notes pan across the screen, means that it becomes increasingly difficult to infer timing visually. Visual pattern recognition still informs *what* sequences of muscle movements should be enacted, but audio cues start to inform *when* those movements should be enacted, and how to remain on beat over a 10-15 note sequence. In my opinion, it is at this level that the game becomes fun, and this is when you begin to really flex your AL.

Level 4: In the highest level songs, it is primarily about AL. VL is still continuing to be strained by some especially difficult songs, but for the most part this skill is fully formed and most songs are completely readable. The player has also built an extensive muscle memory library of TC's, which continues to be added to and challenged by each new level of song. But songs at this level are simply impossible to be played correctly if audio cues are not used to infer beat timing. At this level, it is fully incorporating TC with AL that most determines performance.

From this hypothesized progression of skills, one can see that the main literacy that most influences one's performance changes as one's skill level changes, starting with VL, then moving to TC, then to AL. This also means that AL is only practiced in rhythm games in a highly complex way that must be fully integrated with other visual and tactile skills. This is an interesting comparison to most of the musical literacy tests described in the introduction, which test that literacy in a highly simplified, abstract manner (e.g. Ramsayer et al., 2012).

Hypothesizing from theory

Based on this theoretical framework, I have several hypotheses about how my performance would progress, depending on which skill is being strained the most in a new song. These hypotheses helped guide how I observed the motion of my fingers in my recordings.

H1: If Audio Literacy is most constraining performance, timing should be off for beats, but fingers should be moving in the correct sequences. This timing should get fixed with time, and be the primary factor behind performance improvement. This improvement is only seen over the number of repeated playthroughs of that song.

H2: If Tactile Coordination is most constraining performance, then one would hit the beats at the right time, but in the wrong positions, or to happen in the right position but always with a delay due to higher processing time to ex-

ecute the maneuver. Over time, the positioning and timing should rectify itself as the appropriate muscle memory is built up.

H3: If Visual Literacy is most constraining performance, then there should be trouble inferring both position and timing due to general cognitive overload. Improvement should proceed from random to more purposeful motion. That random motion may or may not be on beat.

Gameplay Observations and Reflections

In looking at my gameplay records, several patterns are apparent. First, my performance has increased over time, and the increase has been somewhat linear, though with a lot of variability (Figure 2). The quantifiable increase in whole-song ability is certainly more regular that I expected it to be. Apparently, the feeling of mastery occurs much more suddenly than the measurable score of mastery. I gained mastery in two of the songs (i.e. achieved at least a B level rating), and despite my focused attention on the issue, still found my feeling of mastery to appear subtly and thoroughly nearly all at once, without exactly knowing when and how it occurred.

Second, I have found a particular stumbling block in one of the songs, which has allowed for a very interesting case study. The song featured a kind of note sequence for which my muscle memory was unable to cope. Over time I learned that the sequence needed to be played with my thumb and middle finger, rather than with my index and middle finger (which was my natural tendency when first encountering that song). This forced me to explicitly unlearn a given TL or muscle memory, and relearn a new technique to succeed in the song. Challenging automated routines with new sequences is a vital process in Gee's cycles of expertise (Gee, 2007), and actually led to a drop in my performance from my initial playthroughs. I perceived my performance improvement as a simple 2-step transition (index+middle finger to thumb+middle finger), but in fact the recordings revealed 5 different finger configurations I employed through the learning process, which was unknown to me before reviewing the recording. Additionally, I did not transition through those sequences in a regular way- although over time the proportion of the more "advanced" configurations" used on a song increased, in any given playthough I would employ between 2-4 different configurations at different points throughout the song. It was only after I started achieving the most advanced configuration with some regularity that I gained self-percieved efficacy.

Additionally, I have noticed that while struggling to learn that new sequence, I underperformed in the sequences immediately before and after the troublesome sequence, indicating that the additional mental energy required to correct the sequence was distracting from my ability to correctly play sequences that I could otherwise easily handle. This also caused highly irregular play performances from session to session, with scores fluctuating quite a bit in successive playthroughs (Figure 2). This also strongly indicates a cognitive overload on TC.

Third, I observed for two of the songs that there was a significant segment of the song that I noticeably increased my performance on, and that increase was in a large way contributing to my final score. Yet in both of these songs, it was not apparent to me that either I was struggling so much with this particular sequence, or that such a drastic improvement had occurred on that sequence. The improvement was a combination of improved timing with some novel repositioning of fingers (so a combination of AL and TC). But in both cases my self-perceived efficacy on the song as a whole only occurred after the sequence that I was unaware I was struggling with had been mastered. Reflecting on my own play, I mostly perceived the section as one in which I did adequate on initially, and showed a modest improvement on over time.

The general sense is that my TC was being most constrained by the songs, and I saw a lot of confirmation for H2. There was in some sense a sharp and binary transition through discrete finger configurations, a clear indication of TC issues. But the transitions occurred in a disjointed and gradual progression that involved significant retrogression. I would not suddenly move from configuration 2 to 3, but would in contrast still be employing configuration 2 and 3 in a song while simultaneously perfecting configuration 5. I found that my learning was filled with discrete steps that transitioned gradually over time. My learning and progress also seemed generally misaligned with my self-perceived efficacy. But perhaps this cycle between success and failure, progress and retrogression, is how game literacy expertise does and should develop during repeated cycles of play.

Figures and Tables

Figure 1: An example of my fingers in the midst of gameplay on a song. Beats appear as the colored green and white shades, and must be tapped when the shades fully cover the square. The rainbow circle is the feedback response indicating that a note was recently tapped with perfect timing.

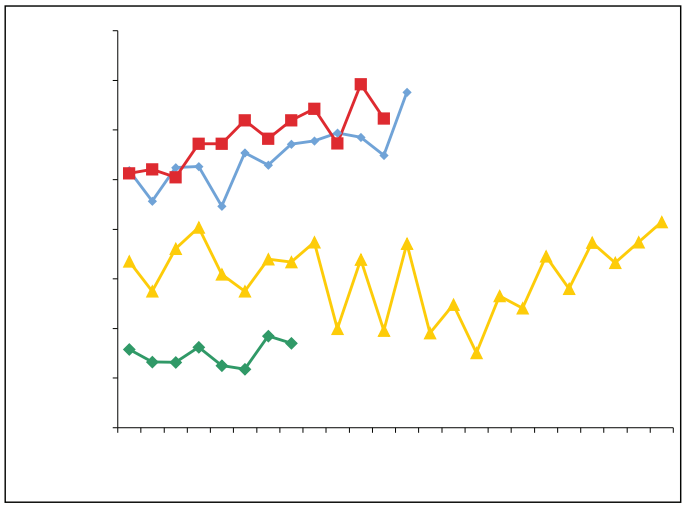

Figure 2: A graph of the author's gameplay performance over time. In this graph, the x axis shows the number of repeated playthroughs of the song, which occurred over a period of 9 months at unequal time intervals.

References

Arsenault, D. (2008). Guitar Hero: "Not Like Playing Guitar At All"? *Loading…* 2(2), 1-7.

Gaydos, M. (2010). Rhythm Games and Learning. *ICLS 2010 Conference Proceedings*, IL: University of Illinois at Chicago

Gee, J.P. (2007). Good Video Games and Good Learning. New York: Peter Lang.

Janzen, T.B, Thompson W.F., Ammirante, P., & Ranvaud, R. (2014). Timing skills and expertise: discrete and con-

tinuous timed movements among musicians and athletes. Frontiers in Psychology, 5, 1-11.

Konami. (2011). *Jukebeat* [iOS videogame]. Tokyo, Japan: Konami

Manning, F. & Shutz, M. (2013). “Moving to the beat” improves timing perception. *Psychonomic Bulletin & Review*, 20, 1133-1139.

Miller, K. (2009). Schizophonic Performance: Guitar Hero, Rock Band, and Virtual Virtuosity. Journal of the Society for American Music, 3(4), 395–429.

Rammsayer, T.H., Buttkus, F. & Altenmuller, E. (2012). Musicians do better than nonmusicians in both auditory and visual timing tasks. *Music Perception*, 30(1), 85-96.

Shute, V. J. & Ventura, M. (2013). Measuring and supporting learning in games: Stealth assessment. Cambridge, MA: The MIT Press.

Squire, K. D. (2008). Video-game literacy: A literacy of expertise. From *Handbook of research on new media literacies*. New York, NY: MacMillan, 635-669.