The best of both worlds

The development of an evidence-based game concept

By Shengnan Chen

Introduction

I come from a psychological research background, and I became really fascinated by serious games because of their power to motivate people to learn and to change behaviors. Even better, they are fun to play and to make. After I finished my Research Master's in Behavioural Science at Radboud University, I started to do an internship at Grendel Games. During my time here, I am learning to do game design and investigating how we can integrate research into the serious games development process.

Research is often conducted after serious game development has been finished. Although that type of research is necessary to prove the effectiveness of the serious games, it does not help to improve the game if it is found to be ineffective, which is unfortunately true in many cases. Introducing research early and making it an integral part of game development would reduce the risk of spending all the budget on a serious game that will prove to be ineffective in the end. Therefore, we have started to explore an approach for research-based game development (which I will briefly introduce in the next section).

There has been a lot of literature on research during serious game development(e.g., Jiang, McKanna, Calabrese, & El-Nasr, 2017; Katsikarelis, 2015; e.g., Olsen, Procci, & Bowers, 2011; Procci, Chao, Bohnsack, Olsen, & Bowers, 2012a). However, the existing literature mostly comes from academia. I have not been able to find documentation of the integration of research into the whole cycle of game development done by serious games companies. Therefore, I would like to share what we learned at Grendel Games by applying academic knowledge to game development in a business context. The insights are not meant as instructional, but rather a springboard for critiques and

improvement. We are trying to get *the best of both worlds* – using the power of research and the expertise of a game development company to make a fun and effective serious games. These insights may be interesting for serious game companies, researchers, or clients who are interested in making a serious game.

We will start from this article, in which we look at the first stage of game development: creating a game concept. I will use a current project as an example. In this project, we are working with the local hospital (MCL) and the Health-tech startup Fennema & Zantema to make a game for chronic pain relief. I will write about the steps we took from research to the game concept and the lessons learned. In this project, the clients are also the domain experts, so I will refer to them as clients/domain experts.

Research-based game development

First, I will briefly describe the research-based game development approach that we propose. We are currently testing this approach using our game development projects. This approach is based on a review of the literature about serious games development and interviews with staff who have been involved in research projects with games developed by Grendel Games.

Katsikarelis summarized the following four phases for serious games production: concept phase, pre-production phase, production phase, and post-production phase (Katsikarelis, 2015). In the concept phase, a game concept is devised based on requirements analyses; in the pre-production phase, a prototype is created and revised iteratively based on playtesting feedback; in the production phase, the game is created and improved iteratively; in the post-production phase, the production team provides support for the game.

The goals of research for each phase is different. Research at the concept phase aims to have an evidence-based game concept which is devised to have a measurable impact (Kato, 2012). During the pre-production phase, the goal of research is to ensure that the core mechanics are effective in terms of both entertainment and education (Rassweiler, 2012). During the production phase, research aims to gather feedback about the usability, engagement, and learning of the game in order to continually make improvements (Procci, Chao, Bohnsack, Olsen, & Bowers, 2012b). Finally, during the post-production phase, the final game is tested for its efficacy and effectiveness in

terms of the learning goal (Kato, 2012). In addition to one-off research studies, in-game metrics can also be implemented to continually monitor the effectiveness (Serrano-Laguna, Manero, Freire, & Fernàndez-Manjòn, 2017).

Based on these different types of research needed for each production phase, we propose the following five steps:

1. Concept phase: making sure that we have a solid, evidence-based game concept *("How can the game make a measurable impact?")*
2. Pre-production phase: testing the core mechanics using prototypes *("Does the core mechanics work for the entertainment and education goals?")*
3. Production phase: testing the usability, engagement, and learning iteratively *("Is the game easy to understand, fun, and effective?")*
4. Post-production phase 1: doing efficacy and effectiveness research of the final game *("Is the game efficacious and effective in reaching the learning goal?")*
5. Post-production phase 2: tracking metrics after game is released *("Is the game effective in the real life?")*

This article will focus on the first part – the concept phase. I will outline the steps that we took to develop the game concept in the next section.

Steps for creating an evidence-based game concept

Define the goal of the game

Defining the goal of the game is the most important step before creating a game concept. Ideally, we want a goal that is specific, time-bound, measurable, feasible, and relevant (Kato, 2012). We also want all stakeholders to agree on this goal. We started with: "Reduce pain intensity by at least XX% for Dutch adult chronic pain patients immediately, X months and X months after they play the game."

We filled in the blanks using evidence in the literature. Using the decrease of pain intensity as an example, I will demonstrate the deliberation we went through to balance feasibility and relevance. Firstly, I chose to specify a percentage decrease instead of aiming for statistical significant changes because it is more relevant for patients (Dworkin et al., 2008). Statistical significance is still important and will be included in our effectiveness research but achieving statistical significant is a result of a combination of the size and variance of the treatment effects and the sample size (Dworkin et al., 2008). Therefore, it does not always mean clinical significance.

A 10-20% reduction in pain intensity is considered by patients to be minimally important, a reduction of 30% or more is considered to be moderately important, and a decrease of 50 or more is considered to be substantial (Dworkin et al., 2008; Farrar, Young Jr., LaMoreaux, Werth, & Poole, 2001). A previous study using a video game intervention to treat chronic pain patients found a 33% decrease in pain intensity from baseline to post-test (Jones, Moore, & Choo, 2016). Therefore, we decided to aim for a 30% decrease, which is both relevant for the patients and feasible according to previous research.

Do a literature review

To make sure that we have a scientific list of requirements for the game, we did a literature research together with the clients/domain experts, which resulted in a document called "Problem Analyses". The whole process of drafting, getting feedback, and finalizing the "Problem Analyses" took one month. Based on the structure recommended by a previous article (Kato, 2012), we defined the problems that the game tries to solve, the goal of the game, the theoretical background, and the existing treatment methods (including games and non-games). A second part of the "Problem

Analyses" proposed candidate outcome measures and the preliminary planning of the efficacy and effectiveness research.

An example of the existing treatment methods that we found from the literature was Cognitive Behavioural Therapy. It focuses on people's beliefs about pain, and how their thoughts affect their behaviour (Eccleston, Williams, & Morley, 2009). We found that among all the psychological treatments, it has the most evidence and is the most effective (Eccleston et al., 2009). Therefore, we decided to include Cognitive Behavioural Therapy as a treatment method in the game. We only reviewed psychological treatments. We justified this point in the "Problem Analyses", but I will not present the details here because it is beyond the scope of this article.

Extract requirements for game concept from the theoretical review

From the "Problem Analyses", we extracted 20 requirements for the game concept. I will illustrate how we did that with one of the 20 requirements. We found that the variability of context can maximize the transfer of the new association learned in exposure therapy based on previous research (Craske, Treanor, Conway, Tbozinek, & Vervliet, 2014). From there, we extracted a requirement that the game should have a varied combination of exercises, scenarios, and environments.

When we had the list of requirements for game concept, we sent it to the clients/domain experts and asked them to rate the priority of each requirement. Seven out of the 20 requirements were rated as "this must be done", six were rated as "high priority", two were rated as "low priority", and four had uncertain priority due to insufficient evidence.

Create a game concept

Together with the game designer Remi, we had a brainstorming session and came up with 15 game concepts. We started from the requirement that we thought was the most important – the game should provoke movements that the patients usually avoid – and jotted down whatever concepts that came to mind. We did not filter the concepts using other requirements. Our ideas ranged from cheerleader management to a *Last Guardian* type of game.

After we had an initial list, we ran the concepts through the 20 requirements that we derived from the previous step and selected the game concept that met the most requirements. It met 18 out of the 20 requirements.

Remi and I then met with Gerard, who is the programmer and the product owner for the game (i.e., product manager), to further discuss this concept. Gerard pointed out that the main gameplay (i.e., collecting) in the concept is not the most fun and suggested another gameplay (i.e., puzzle). It was more fun, allowed more individual variations for patients, and is suitable for people of different ages, so we decided to use a puzzle-based gameplay instead. We did not go for an action-based gameplay because we thought it would not be suitable for older patients.

We also revised the game concept so that all the requirements that were deemed mandatory and highly important by the clients/domain experts were reflected in the game. Some requirements, while not deemed important by them, were still included because we thought they had the potential to distinguish the game from previous games for chronic pain.

When we had the first version of the game concept ready, we showed it to some of the colleagues. A colleague suggested that the game concept did not have an end goal so it may not motivate the players. We were out of time to incorporate his feedback although it was a valid point.

Ask clients/domain experts to evaluate the game concept

The clients/domain experts looked at the first game concept and had a meeting with us. There was a large discrepancy between their ideas of the game and the initial concept. The clients had done multiple testing with chronic pain patients using existing games, and they had obtained new insights about what the patients liked. They suggested four requirements that they thought were the most important according to the testing. The main discrepancy was that from our literature review, we reached the conclusion that behavioral experiment and cognitive distraction were the most important treatment methods and we had a game concept based on that. However, the clients/domain experts thought that cognitive diffusion was the most important treatment method that should be included in the game. As a result, we had to change the entire concept.

The meeting was very efficient, and we agreed on a new game concept within an hour.

Revise game concept

We replaced the entire concept with the concept that the clients/domain experts proposed. The setting, gameplay, and story were changed. The clients/domain experts suggested an action-based gameplay and the game development team took the suggestion. The clients/domain experts approved this version of the concept, and the development team was happy with this concept too because it was an accurate reflection of the new requirements, and we could make the new gameplay fun more easily. The new setting and story fit well with the new gameplay.

The first reason for this radical change was that the clients/domain experts were much more specific about the requirements after they saw the first concept. They narrowed down the list of requirements to four requirements, and only two of these requirements came from the 13 requirements that they deemed highly important or mandatory before reading the game concept. The other two requirements were new.

The second reason was that the clients/domain experts came with new information that disproved the assumption we had about the target audience. We previously assumed that an action-based gameplay would not be suitable for the target audience, but the clients/domain experts had tested an action-based game with the target audience and found out that they liked it.

The third reason was that we mis-judged which requirements were the most important because we did not have the domain knowledge and clinical experience to make the judgement.

We did not follow through our first concept, and it may seem like we gave the clients/domain experts too much power in the creative process. That is a risk that we should be cautious about in our work, but in this case the changes were justified because they were based on new evidence and a more accurate interpretation of the literature. The only thing that we could have done better was to also participate in the testing that the clients/domain experts ran with the target audience to make sure that we could look at the testing through the game development team's perspective.

Lessons learned

Nobody has made a game like this, so we cannot just use a magic formula. Going through all the steps above made sure that we have an evidence-based game concept that has been tested by the clients/domain experts and approved by the clients/domain experts. Nonetheless, there are many improvements we can make to the process:

Combine the domain experts' opinions with the literature review

We spent a long time on doing the literature research, but all of the 20 requirements for the game concept came from the section about existing treatment methods (games and non-games). Out of the 20 requirements, only two ended up in the final game concept. The academic in me had the tendency to do a thorough review, but the business world is much more result oriented than academia. The output : input ratio seemed really low using the thorough literature review method.

I think the reason is that although I could find a long list of effective treatment methods for chronic pain from the literature, I did not have enough domain knowledge in chronic pain to judge which ones were better and more important. A lot of time could have been saved by asking the domain expert to point us to the most effective treatment methods in the beginning of the literature research. Then we can focus on finding evidence to support those specific treatment methods and finding measures for the outcomes of those treatment methods.

On the other hand, even though at the surface it seemed like a thorough literature review did not produce a lot of results, I still think it is necessary for a researcher within the company (or an external research partner) to go through this thorough review. Although domain experts have a lot of practical experience, they may not be aware of all the research evidence. For example, while writing this article, I came across a Cochrane systematic review of psychological therapies for managing chronic pain (Williams, Eccleston, & Morley, 2012). Systematic review is a synthesis of all the existing evidence and it is the highest level of clinical evidence (Phillips et al., 2009). I found out from the systematic review that Cognitive Behavioral Therapy (which we are using in the game) has small positive effects on disability and catastrophizing, but no impact on pain or mood compared to active control (Williams et al., 2012). Based on this evidence, I would advise to change the goal of the game to reducing disability and catastrophizing instead of reducing pain. Furthermore, a thorough literature review not

only suggests *what* treatment methods to use, but also provides the rationale for *why* we use those methods. In summary, we should ask the domain experts to guide us through the literature review, but at the same time we should back up their opinions with research evidence (whether we ask them for the evidence or look for it ourselves) because expert opinion is the lowest level of clinical evidence (Phillips et al., 2009).

Participate in the testing with the target audience together with the clients/domain experts

Our clients/domain experts did a great thing – they did many testing with patients to evaluate the different treatment methods. From those testing, they narrowed down the treatment methods that worked the best. It would have been better if people from the game development team also participated in the testing so that we could also observe what worked the best in terms of gameplay. We may have reached different conclusions based on our perspective.

Elicit feedback from within the company as soon as we had a draft of the game concept

The first time someone within the company who was not on this project saw the game concept was when we had the complete first game concept. He provided very insightful feedback, but we did not incorporate that because we did not have enough time. In the future, we should show the draft of the game concept as soon as we have a rough draft **to make use of the team's intuitions and expertise in making fun games.**

Add a step to filter the clients/domain experts' feedback before revising the game concept

In this case, changing the game concept radically **based on clients/domain experts'** feedback was justified because the new concept adhered to the new evidence, and gameplay suggested by the clients/domain experts was considered by the game development team to be more fun than the gameplay in the first concept.

However, this could go wrong sometimes. I interviewed Tristan, the Game Design Director at Grendel Games, about Garfield Count Me In, a game that the company made. In the beginning, the domain experts insisted that the calculation exercises should not have any gameplay elements because the target audience, children in

special education, can be easily distracted and have difficulty completing the exercises. Although Tristan struggled with the dilemma that having calculating exercises without gameplay elements would make the game boring, he went with the domain experts' suggestion because he trusted their expertise. Later, when the game's development was almost finished, he heard from another domain expert that it would not be problematic to incorporate gameplay into the calculation exercises for this group of children.

In the future, we should add a step to filter domain experts' feedback. In this step, we should separate the functional requirements from the creative suggestions. In other words, we should follow the clients/domain experts' requirements about the goals of the game but be more cautious when they suggest how the game should look, how the gameplay should be, or how the educational elements should be incorporated into the gameplay. If we are in doubt, we can get the opinions of other domain experts and the other members of the company or test the ideas with the target audience. If the clients/domain experts do not like the game concept, but it is effective and liked by the target audience, we should stand our ground.

Expect a communication curve and communicate more frequently with the clients/domain experts

Figure 1. Communication curve.

From the initial ideas that the client came to us with to the current game concept, there have been multiple iterations. During every iteration, we discovered discrepancies between what we thought the clients/domain experts wanted and what they actually wanted. It feels like the discrepancy is getting smaller and smaller as we talk with each other more and more. Eventually we will reach the Bull's eye – a game concept that we are both happy with (see Figure 1). Some parts of the curve are avoidable. For example, if we could get the requirements from the clients/domain experts and clarify them earlier, and if we could involve the clients more while making the first game concept, we could have made a game concept that was more on-point at the first time. Other parts of the curve were inevitable because the clients/domain experts may not be able to tell us what exactly they want until they see the concrete game concept. It indicates that when we plan for the development of a serious game, we should take the communication time into account, and communicate with the clients/domain experts earlier and more frequently.

Next step

With the background research and feedback from clients/domain experts, we have a solid game concept now. The next step is to bring the concept to life by making a prototype! We will run testing of the prototype and share the insights in another blog post. Stay tuned!

References

Craske, M. G., Treanor, M., Conway, C. C., Tbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10-23.

Dworkin, R. H., Turk, D. C., Wyrwich, K. W., Beaton, D., Cleeland, C. S., Farrar, J. T., . . . Zavisic, S. (2008). Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *The Journal of Pain, 9*(2), 105-121.

Eccleston, C., Williams, A. C. D. C., & Morley, S. (2009). Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane Database of Systematic Reviews, 2009*(3). doi:10.1002/14651858.CD007407.pub2.

Farrar, J. T., Young Jr., J. P., LaMoreaux, L., Werth, J. L., & Poole, R. M. (2001). Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain, 94*(2001), 149-158. doi:10.1016/S0304-3959(01)00349-9

Jiang, R., McKanna, J., Calabrese, S., & El-Nasr, M. G. (2017). Iterative design and testing for the development of a game-based chlamydia awareness intervention: A pilot study. *Games for Health Journal*, 205-216.

Jones, T., Moore, T., & Choo, J. (2016). The impact of virtual reality on chronic pain. *PLoS ONE, 11*(12), e0167523. doi:10.1371/journal.pone.0167523

Kato, P. M. (2012). The role of the researcher in making serious games for health. In *Serious Games for Healthcare: Applications and Implications* (pp. 213-231). Hershey, PA: IGI-Global.

Katsikarelis, T. (2015). *Serious Games Production: State-of-the-art, State-of-the-practice and Potential Requirements Engineering Benefits.* (Master's thesis), Utrecht University,

Olsen, T., Procci, K., & Bowers, C. (2011). *Serious games usability testing: How to ensure proper usability, playability, and effectiveness.* Paper presented at the Design, User Experience, and Usability. Theory, Methods, Tools and Practice. DUXU 2011, Berlin.

Phillips, B., Ball, C., Sackett, D., Badenoch, D., Straus, S., Haynes, B., . . . Howick, J. (2009). Oxford Centre for Evidence-based Medicine – Levels of Evidence (March 2009). Retrieved from https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/

Procci, K., Chao, A., Bohnsack, J., Olsen, T., & Bowers, C. (2012a). Usability in serious games: A model for small development teams. *Computer Technology and Application*, 315-329.

Procci, K., Chao, A., Bohnsack, J., Olsen, T., & Bowers, C. (2012b). Usability in serious games: A model for small development teams. *Computer Technology and Application, 3*(4), 315-329.

Rassweiler, T. (Producer). (2012, October 17). Features: Rapid Prototyping: Tips for Running an Effective R&D Process. *Gamasutra*. Retrieved from https://www.gamasutra.com/view/feature/179501/rapid_prototyping_tips_for_.php?page=1

Serrano-Laguna, A., Manero, B., Freire, M., & Fernàndez-Manjòn, B. (2017). A methodology for assessing the effectiveness of serious games and for inferring player learning outcomes. *Multimedia Tools and Applications, 77*(2), 2849-2871. doi:10.1007/s11042-017-4467-6

Williams, A. C. D. C., Eccleston, C., & Morley, S. (2012). *Psychological therapies for the management of chronic pain (excluding headache) in adults*. Retrieved from