COWBOYS, ANKLE SPRAINS, AND KEEPERS OF QUALITY: HOW IS VIDEO GAME DEVELOPMENT DIFFERENT FROM SOFTWARE DEVELOPMENT

Authors:
Dr. Emerson Murphy-Hill
Dr. Thomas Zimmermann
Dr. Nachiappan Nagappan

Felix Wang,
David R. Cheriton School of Computer Science

UNIVERSITY OF WATERLOO | FACULTY OF MATHEMATICS

PROBLEM TO BE SOLVED

Research question

- RQ:

  How is Video Game Development Different From Software Development?

- Answer (Most Likely):

  It's different in some ways but similar in others.

- Why is it still important then?

  - Rarely studied

  - Only 3/116 software projects studied in the past 2 years at major SE venues were games.

  - Game developers hold themselves apart from formal SE.

Meanings of this paper

- The first study that systematically and empirically investigated the differences between traditional SE development and video game development.

- Implications:

  - If they are different -> Tools and practices for traditional SE development are not applicable to Game development -> New direction for research + impact on education.

  - If they are similar -> Game developers can benefit from the established and validated tools and practices of traditional SE development.

NEW IDEA

Methodology

- Qualitative interviews:

- 14 interviewees (till saturation) found on LinkedIn.

- at least 2 years of Game development, and at least 2 years of non-game development in the past 10 years.

Methodology

- Quantitative Surveys:

- 5-point Likert scale survey consisting of 28 statements.

- 364 participants working in Microsoft on Games, Office, and Others.

Interview Results

- SWEBOK Topics #1: ***Software Requirements***
- Perceptions:

- Functional requirements are for non-game development. Games are fun and subjective, so no strict requirement is needed, only high-level goals.

- Game's user experience is different from non-games, and unfulfilled requirements in Games are less problematic.

- Requirements come from multiple sources. Gamers, as the ultimate client of games, have less say in their requirements.

Interview Results

- SWEBOK Topics #2: ***Software Design***
- Perceptions:

  - Less up-front thought on architecture.

  - Game's lifespans are hard to predict, so it's hard to determine if reusability is important.

  - Game development rewards creativity.

Interview Results

- SWEBOK Topics #3: ***Software Construction, Tools, and Methods***
- Perceptions:

  - Project-specific performance tuning cannot be reused.

  - Games focus more on innovation than similarity.

  - There are some reuse on game engine/general SE tools, etc.

Interview Results

- SWEBOK Topics #4: ***Software Testing and Quality***
- Perceptions:

  - Games are less testable, because:

    (1) The state space of games is too large.

    (2) No clear definition of what constitutes "correct behavior".

    (3) Non-deterministic due to AI/Multi-threading/Distributed computing, etc.

    (4) Automated testing is fragile to frequent changes.

    (5) Software tester is more expensive than human testing.

Interview Results

- SWEBOK Topics #5: ***Software Maintenance***
- Perceptions:

  - Maintenance is delayed due to:

    (1) Less management buy-in.

    (2) New content is enough for product release, no need to modify and extend program behaviors.

  - Cloud-based platforms like Steam may enhance game maintenance.

Interview Results

- SWEBOK Topics #6: ***Software Configuration Management***
- Perceptions:

  - Games have significant amounts of content, and configuration management is "chaotic".

  - Part of the problem is due to a lack of code review.

Interview Results

- SWEBOK Topics #7: ***Software Engineering Management***
- Perceptions:

  - People in game management positions tend not to have technical backgrounds.

  - It's hard to communicate engineering issues to non-engineers.

  - Non-engineers don't respect engineering activities because they have no immediate impact.

Interview Results

- SWEBOK Topics #8: ***Software Engineering Process***
- Perceptions:

  - Agile is a good fit because of the game's unpredictability.

  - There's a lack of process.

  - Imposing control may oppress creativity.

  - Game dev has much more pressure on releasing software on time.

Interview Results

- General Work Topics #1: ***Problem Solving and Skill Variety***
- Perceptions:

  - Game dev has distinct technical challenges.

  - People's understanding of "fun" is subjective.

  - Game developers require a wider variety of skills.

Interview Results

- General Work Topics #2: ***Autonomy***
- Perceptions:

  - Game developers must have a high level of autonomy.

Interview Results

- General Work Topics #3: ***Specialization and Independence***
- Perceptions:

  - Communication and conflict resolution are required to bridge interdisciplinary gaps.

  - Specialists like *sociologists, anthropologists,* and *economists* are required.

Interview Results

- General Work Topics #4: ***Interaction Outside the Organization***
- Perceptions:

  - Game developers tend to have a stronger tie to the customer.

Interview Results

- General Work Topics #5: ***Knowledge of Results***
- Perceptions:

  - Profitability and game rewards won are good indicators of results.

  - Game developers have a “celebrity status” when the game they developed goes popular.

Interview Results

- General Work Topics #6: ***Significance and Experienced Meaningfulness***
- Perceptions:

  - Game developers find meaning in their work by knowing how many people use it.

  - Games are entertainment, it feels meaningful to create positive emotional experiences for others.

Interview Results

- General Work Topics #7: ***Physical Demands and Work Conditions***
- Perceptions:

  - The video game industry is notorious for long work hours.

  - Developing motion-based games is physically demanding.

Survey Results

- Confirmed that, in Game development:
  - Less clear requirements.
  - Use Agile more.
  - Creativity is valued more.
  - The ability to communicate with non-engineers is valued more.
  - Require a more diverse team.
  - People are more impressed by Game dev.
- Disconfirmed that:
  - Engineers' likelihood to move into management shows no difference.

POSITIVES

Positives #1: Methodological Triangulation

- Combined qualitative data (interview) with quantitative data (Likert-scale survey).
- The authors themselves admit that the interviews and surveys individually provide limited insights.
- But combining them together:
  - Cross-validation.
  - Increased the depth and breadth of insights.

Positives #2: Soundness of Interview Analysis

- In an interview study, we emphasize **interpretive depth** and **emerging patterns** rather than the counts or distributions of occurrence.
- Even a single participant's account can sometimes be analytically meaningful, even if it's rare.
- Unlike some papers, which sacrificed the soundness of their research methods for some quantifications like "(22/77) participants perceived…", the author focused on **conceptual richness** and the **theoretical relevance** of responses, and discussed some meaningful and interesting but rare responses.

NEGATIVES

Negative #1: Unsoundness of Survey Analysis

- Took the mean of Likert-scale responses, which is a controversial indicator.
- The number of Likert scales is an ordinal number, with no numerical meaning.
- Can someone guarantee the interval between "Strongly Agree" and "Partially Agree" to be the same as the interval between "Partially Agree" and "Neutral"?
- While it's acceptable, it departs from strict survey analysis methodology.

The two Effect Size columns indicate the difference in means between Games and Office in the first subcolumn and the difference between Games and Other in the second subcolumn. For example, the mean response to S13 for game developers was a score of about 4.5 (between "Agree" and "Strongly Agree") whereas the mean response for Office developers was 3.3 (between "Neutral" and "Agree"); as a consequence, the effect size is 4.5 – 3.3 = 1.2. Effect sizes are additionally colored on a gradient from blue to orange; blue colors means game developers were more likely to agree with the statement and orange colors mean they were less likely to agree.

Page 9

Negative #2: Disproportionate Survey Participants

- Participants in Games (145), Office (100), and Other (119).
- The paper is using the Wilcoxon rank-sum test, which will have more statistical power in classes with more samples and less statistical power in classes with fewer samples.
- Thus, in this paper, the results contrasting Games with other domains are statistically more robust and confident, but the findings could be interpreted with less confidence within smaller groups.

	Games	Office	Other
Mean years at Microsoft	4.4	7.1	5.1
Mean years of development experience	10.7	11.0	8.8
Number of engineers	113	61	82
Number of testers	32	39	37

Page 3

- Participants in Games are 45% more than in Offices, possibly overshadowing subtle effects in smaller groups.
- The authors did not identify this threat in their Limitations section.

Negative #3: Suspected Experience Bias in Interview Participants

- This paper guarantees participants have at least 2 years of experience in Game dev and at least 2 years of experience in non-Game dev.
- But there may be biased participants who are more senior in game development and less senior in non-game development (e.g. developers with 8 years of game dev + 2 years of non-game dev).
- Game developer's mindset.
- This will potentially introduce bias: (1) We shouldn't be restricted by unit tests, (2) We shouldn't be restricted by considering reusability, (3) We shouldn't care about gamers' requirement that much, etc.

		Games	Non-Games
Median years of development experience		8.5	8.5
Number of interviewees with "extensive" experience in…	Programming	10	12
	Design	6	5
	Management	7	4
	Audio/Visual	2	3
	Testing	3	5

Page 3

- These perceptions may be biased if there's a disparity between each participant's experience in Game Dev and Non-Game Dev.
- There should be at least two rows: "Median years of game development experience" and "Median years of non-game development experience" in this table to guarantee this, so as to justify its credibility.

FUTURE WORK

Lab-based Controlled Experiment on Unit Tests in Game Dev

- RQ:
  Will using unit tests in Game Development with MVC architecture enhance the quality of games?
- We organize a coding session for 40 participants to complete two game dev coding tasks (P1 and P2) of a game's Model part, with or without writing unit tests.
- Group A: P1 first (no unit tests), then P2 (at least 70% statement coverage).
  Group B: P1 first (at least 70% statement coverage), then P2 (no unit tests).
  Group C: P2 first (no unit tests), then P1 (at least 70% statement coverage).
  Group D: P2 first (at least 70% statement coverage), then P1 (no unit tests).
- We measure the quality of code by each participant eventually.

RATING

4.5/5 (GOOD PAPER OVERALL)

An extremely novel paper, which identifies the underexplored domain in software engineering, opened the research perspective on a series of game-related studies, including game dev practices, empirical standards on gamification, etc. But it contains minor imbalances and is not exhaustive enough.

DISCUSSION POINTS

1. What caused the disparity that the interview study suggests that people in game management positions tend not to have technical backgrounds, but survey studies disconfirmed that? Which one should we believe in?
2. In section 4.1.8, it says “interviewees reported being under significantly more pressure to release the software on time for games than for non-game software”. Do you think this is inherently why game dev is different from non-game dev, or is it the consequence of a lack of process?
3. In section 4.1.3, it says “there is less code reuse in games … because reuse implies similarities between software, yet games emphasize innovation”. Do you think traditional non-game devs are pursuing similarities rather than innovations? Do you think there is a trade-off between similarities and innovations?
4. After reading this paper, do you think the game development is more “art”, more “science”, or more “engineering”?