Deliverable Evaluation Report

1. Performance Testing Report

1.1. Question:

Does the application maintain an acceptable average frame rate during gameplay and destruction simulation and therefore satisfy Learning Goal 3 “Maintain gameplay performance whilst dynamically simulating environmental destruction”? By what extent does the application satisfy Learning Goal 3.

1.2. Hypothesis:

The application maintains an average frame rate of of approximately 30fps not dipping below 30fps in its 1%low measure.

1.3. Build Specifications:

Final built project packaged utilizing UE5 “Shipping” package mode.

- 9 Walls
    - 3 brick
    - 3 wood
    - 3 concrete
- 3 projectiles

Hardware test was conducted on:

- CPU: AMD Ryzen 7 5800X
- RAM 32GB DDR4
- GPU GeForce RTX 3070

1.4. Sample:

10 x 30s performance captures during gameplay.

1.5. Method:

1.5.1. Methodology:

Manual performance testing utilizing NVIDIA FrameView to generate a 30s performance benchmark report conducted by lead developer.

1.5.2. Procedure:

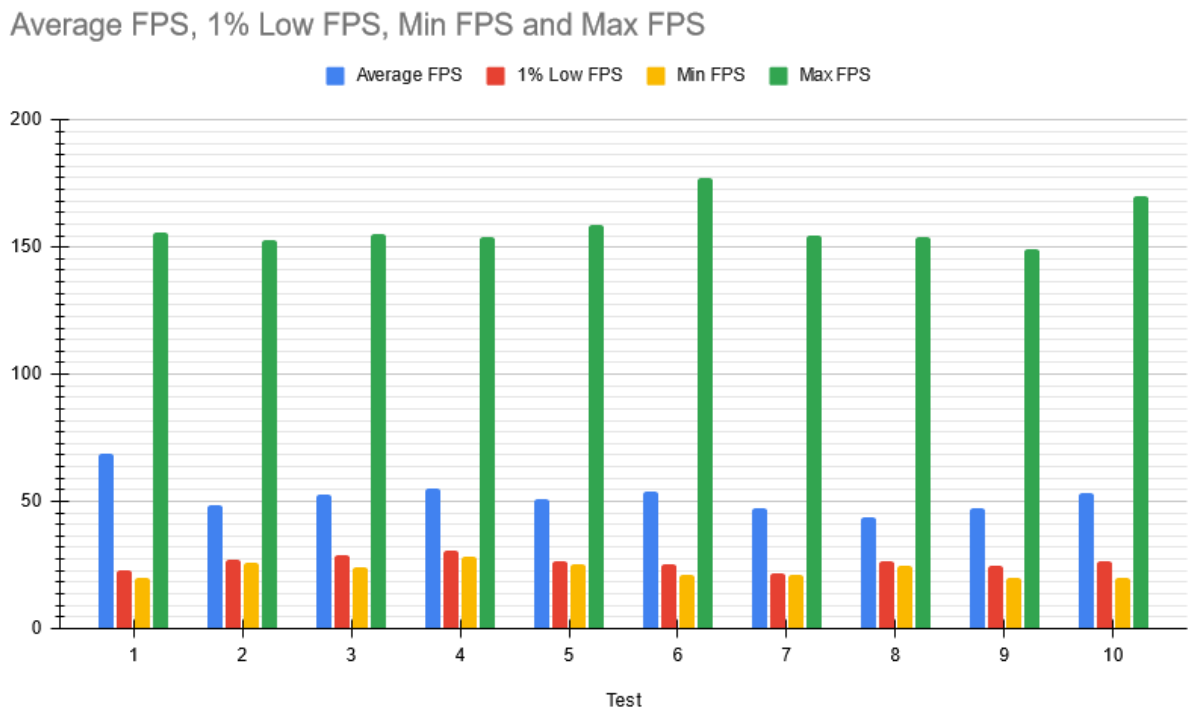
1. Tester starts game
2. Tester enters the gameplay level.
3. Tester selects projectile 2
4. Tester begins performance capture
5. Tester fires 1-3 projectile 2s into each wall to ensure comprehensive destruction of each wall.
6. 30s capture ends
7. 30s capture records results into csv table

1.6. Results:

(Figure 1 Performance test results table)

Test	Average FPS	1% Low FPS	Min FPS	Max FPS
1	68.96	22.83	19.99	155.71
2	48.76	27.32	25.68	152.95
3	52.75	28.92	23.96	154.92
4	54.83	30.49	28.25	154.09
5	51.01	26.62	25.12	158.56
6	53.7	25.06	20.92	177.00
7	47.55	21.95	21.19	154.84
8	44.01	26.32	24.59	153.87
9	47.03	24.38	19.90	149.21
10	53.03	26.58	19.97	170.09

(Figure 2 Performance test results graph)



1.7. Discussion:

The two most important measurements returned from this test are the average FPS measurement and the 1%Low measurement. These two measurements combined give the greatest overall insight into the average performance of the application. Whilst MaxFPS and MinFPS provide less important insights into the peaks and troughs of performance and generally indicate areas of unusually high or unusually low activity within the application.

When observing figures 1 and 2 the average fps remains high between 45 and 60fps indicating a relatively performant average experience in general gameplay. However the significant distance between the average fps measure and the 1% low measure indicates the presence of significant frame drops throughout the testing session. These lower frame rates are likely to occur during aspects of the game that are the most hardware intensive, in this case this is likely to be upon the destruction of a mesh as it fractures into smaller pieces. This is likely to present itself as stuttering during the first frames following a mesh's destruction as the hardware begins tracking and calculating a significant spike in physics objects and collisions.

1.8. Conclusion:

In conclusion, the results gathered have demonstrated that whilst the average performance of the application does maintain an acceptable average fps, the moderate temporary performance drops during some of the more complex destruction simulations that create frame stuttering means that I would be amiss to conclude that the application has successfully achieved Learning Goal 3. Whilst the application comfortably maintains an average fps of over 30 the significantly comparative drop to the applications 1% lowe that frequently falls well below 30fps denotes a likelihood of stuttering within the application during play.This inability to provide a smooth and consistent frame rate would significantly hamper the play experience in a larger game and would only be worsened as more destruction of this type is implemented within a level. These results therefore conclude that the application whilst able to maintain a grater than 30fps average cannot maintain a greater than 30fps 1%low and therefore the hypothesis is not supported. Additional performance optimization would be required for this application to effectively achieve learning goal 3.

2. User Testing Report

2.1. Question

How do players perceive the various destruction simulations within the application? Do the simulations behave how players expect or are they cartoony in nature?

2.2. Hypothesis

That the destruction simulations will behave how the players expect them to more often than not.

2.3. Build Specifications

Final built project packaged utilizing UE5 “Shipping” package mode.

- 9 Walls
    - 3 brick
    - 3 wood
    - 3 concrete
- 3 projectiles

Hardware specifications: These tests were conducted on a wide variety of PCs.

2.4. Sample

8 users with a variety of experience with First, and third person games that incorporated environmental destruction.

2.5. Method

2.5.1. Methodology:

Users were to play a built copy of the game filling in a survey during and after their testing session.

2.5.2. Procedure:

1. Testers opened the game level
2. Testers opened the survey
3. Testers were instructed to complete the survey as they interacted with the level
4. Testers were advised they could restart the level any time in order to try different projectiles or destroy the walls more than once.
5. Testers were instructed on the games controls and HUD
6. Testers were provided an explanation of the different sections of the survey.
7. Testers were left to complete the gameplay test and survey.
8. Survey results were collected.

2.6. Results

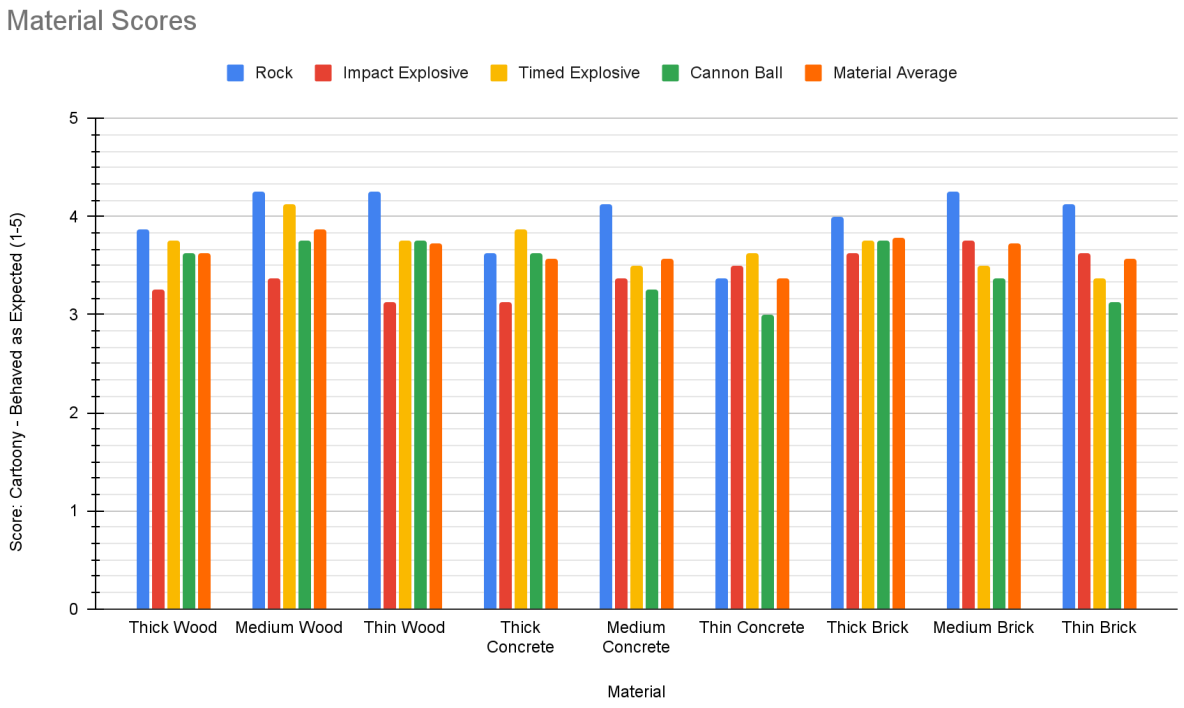
Interaction Rating Results:

(Figure 3 Interaction Rating Results Table)

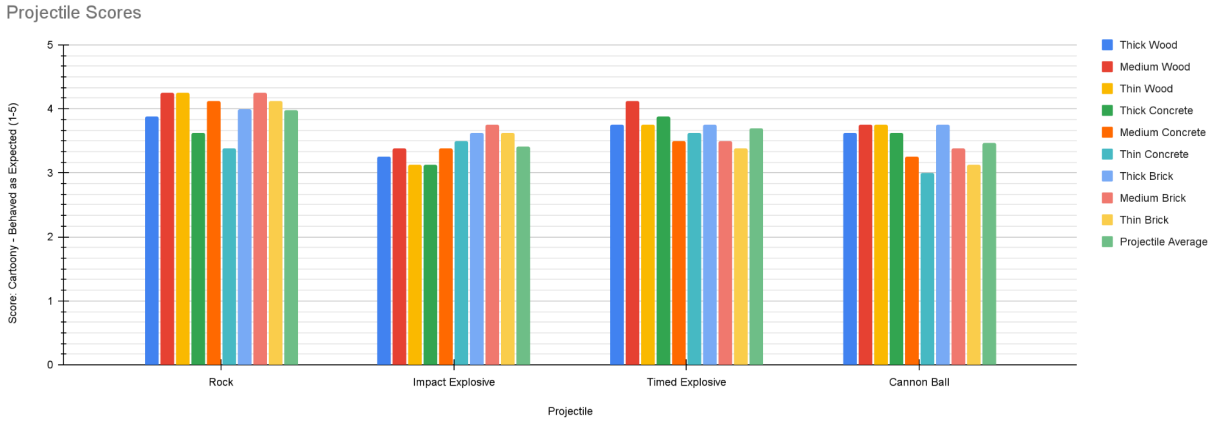
Wall x Projectile Avg Scores	Rock	Impact Explosive	Timed Explosive	Cannon Ball	Material Average
Thick Wood	3.875	3.25	3.75	3.625	3.625
Medium Wood	4.25	3.375	4.125	3.75	3.875
Thin Wood	4.25	3.125	3.75	3.75	3.71875
Thick Concrete	3.625	3.125	3.875	3.625	3.5625
Medium Concrete	4.125	3.375	3.5	3.25	3.5625
Thin Concrete	3.375	3.5	3.625	3	3.375
Thick Brick	4	3.625	3.75	3.75	3.78125

Medium Brick	4.25	3.75	3.5	3.375	3.71875
Thin Brick	4.125	3.625	3.375	3.125	3.5625
Projectile Average	3.986111111	3.416666667	3.694444444	3.472222222	

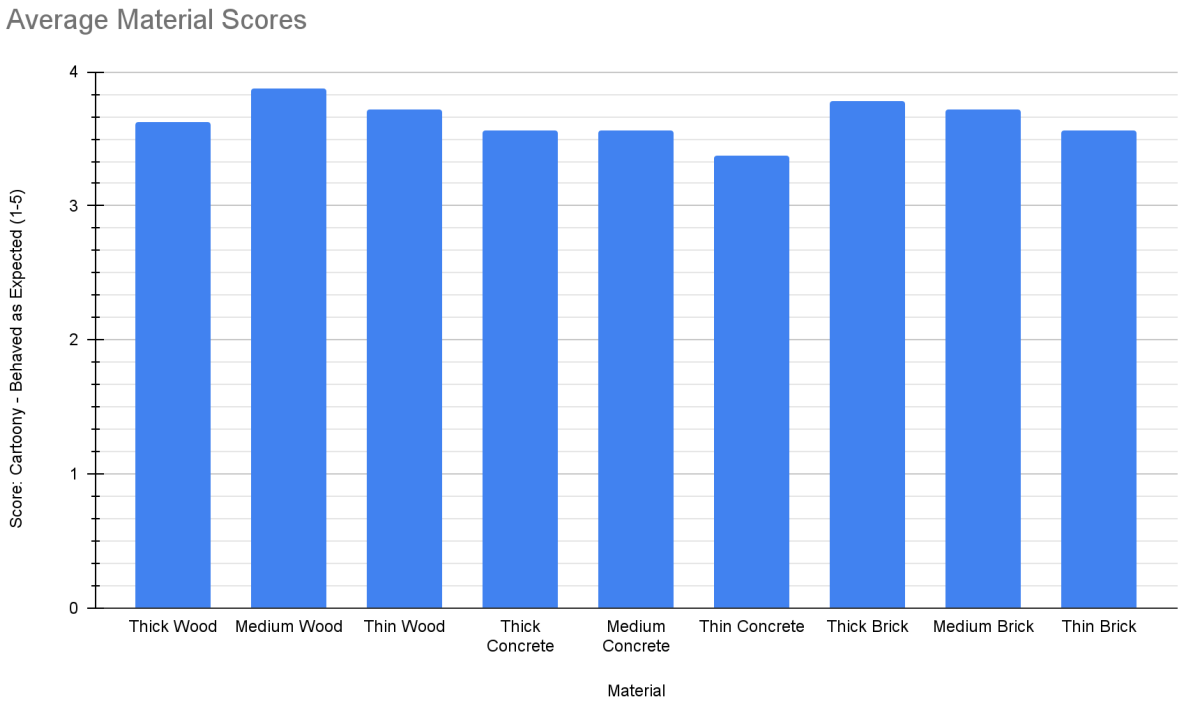
(Figure 4 Material Rating Scores Graph)



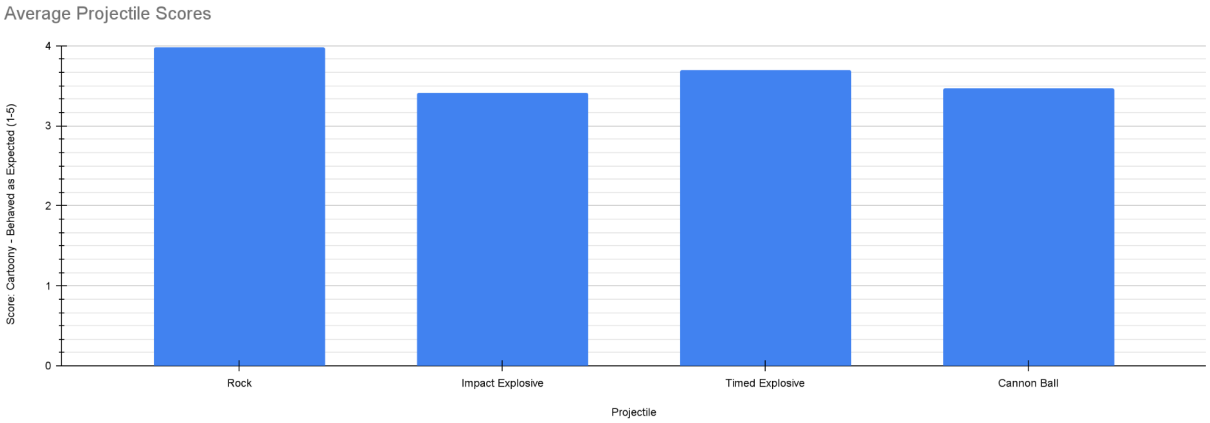
(Figure 5 Projectile rating scores graph)



(Figure 6 Average Material Rating Scores Graph)

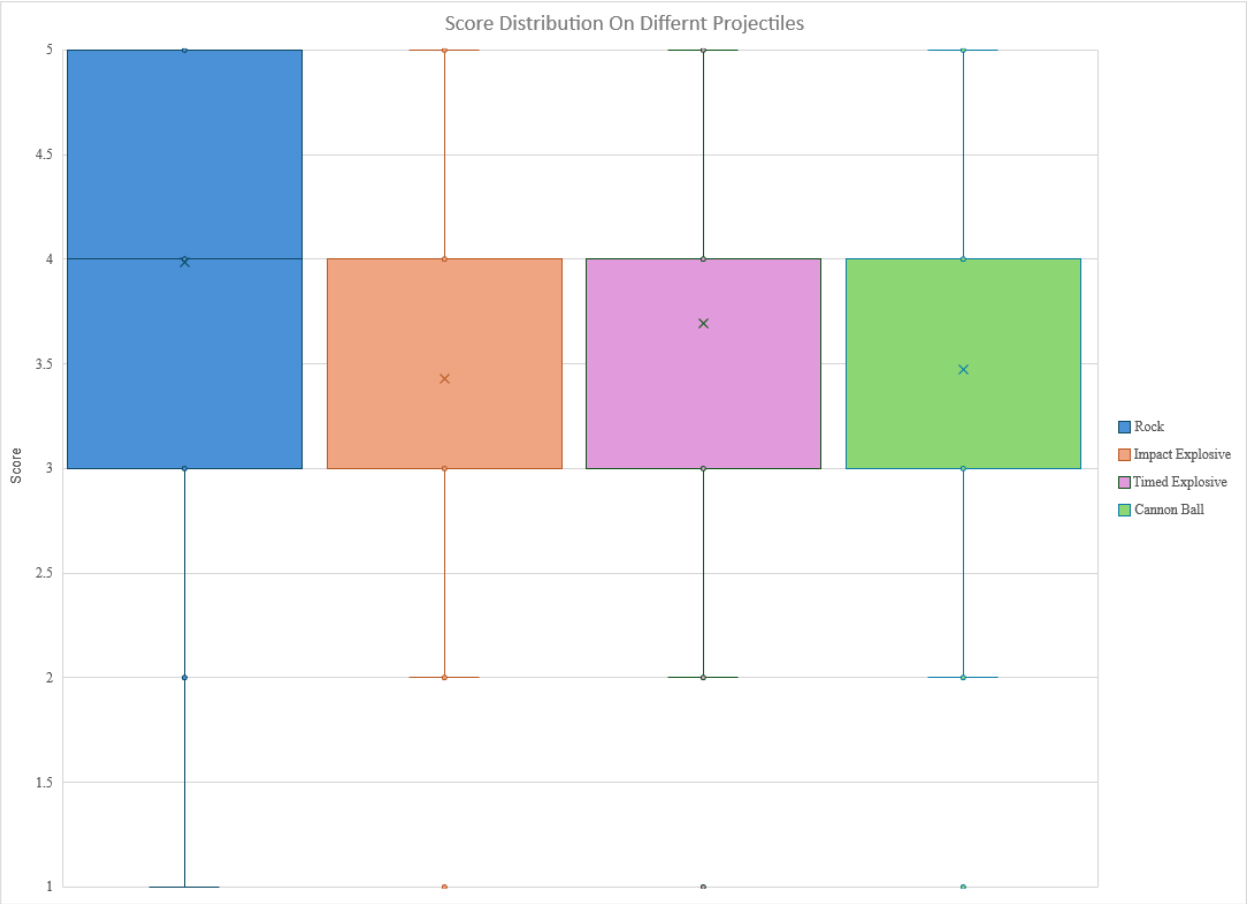


(Figure 7 Average Projectile Rating Scores Graph)



(Figure 8 Material rating score distribution)

(Figure 9 Projectile rating score distribution)



Long Answer Question Results:

- **Question:** Please describe how you felt about the destruction of wooden (brown) surfaces overall. Did it feel realistic or strange? if so how?
    - The response to this question was largely positive
        - “Very realistic, wood splintering felt good.”
        - “Wood felt good especially with how the panel breaks at the base due to it being pushed. Thicker panels felt cartoony especially with explosives breaking predictably.”
        - “Realistic in that it simulated planks of wood, and splintering at impact points. Strange in how explosions would destroy the wall in the same way as solid projectiles (big chunks of splintered wood)”
    - A number of responses mentioned what brought their rating down was that the wood tended not to break at the point of impact but instead at other areas.

        - “Could maybe splinter more on the contact point, instead of spreading the impact all the weigh vertically... otherwise ok”
        - “The splintering was largely realistic apart from the ends that began to break off first rather than the area of impact, as well as the residual energy of the shards spinning around and being launched so far”
- **Question:** Please describe how you felt about the destruction of concrete (grey) surfaces overall. Did it feel realistic or strange? if so how?
    - Response to concrete material was mixed
        - “Realistic”
        - “Strange in that solid projectiles would usually cause the base to break before blowing a hole through the surface especially with cannon ball (projectile 4). Realistic in that an angled hit from a solid projectile would bounce off without much damage, and explosion damage was mostly as expected.”
    - A number of different opinions on what didn’t feel correct or right about it
        - “thin walls feels strange as it behaved like paper meshes, but thicker walls feel real as cannonballs get stopped straight away.”
        - “The concrete felt the least realistic in how the whole wall collapsed. Although the initial impact felt grounded in reality, it often shattered from the bottom and then the rest of the wall just flipped over as one unit”
        - “Concrete sometimes felt a bit "cake" like, but otherwise good. Could be the shading.”
- **Question:** Please describe how you felt about the destruction of brick (orange) surfaces overall. Did it feel realistic or strange? if so how?
    - Had a higher number of generally positive responses stating that it broke apart as expected
        - “probably the most accurate of the 3 surfaces”
        - “Felt good on the thicker walls, the thin walls felt cartoonish except for the rock and grenade.”
        - “They look realistic”
        - “Bricks broke realistically into large chunks and further into smaller chunks. explosives felt a little cartoony due to bricks flying and shattering too extremely.”
    - Some responses advised that the larger explosions and higher velocity projectiles caused the destruction to look cartoony/lego brick like
        - “At large explosions, the bricks do not break but instead gets shoot out, it feels cartoony but somewhat realistic (or acceptable for in-game practices)”
        - “The crumbling of the bricks were fairly realistic, if a bit lego-like in how they broke apart. When shot with higher-velocity projectiles it had disproportionate force in its explosions”

- **Question:** Please describe how you felt about the impact and interactions of projectile 1 (Rock (blue))
    - A number of responses advised that the projectile felt low powered but consistent
        - “Good overall, probably the most consistant.”
        - “Exactly as expected”
        - “Fairly weak on the tougher ones, but largely realistic”
        - “Feels like a heavy solid object, that also breaks apart on impact, fairly accurate and straightforward. Can also bounce off debris or other projectiles.”
    - Some responses thought it was too low powered
        - “pretty lack luster, would expect a bit more of a dent”
        - “The power of the rock is to low”
- **Question:** Please describe how you felt about the impact and interactions of projectile 2 (Impact Bomb (orange))
    - Generally described as strange, inconsistent, or overpowered or some combination of the 3
        - “A little bit strange, explosion strength was tricky to judge at first. First hits on the wooden surface would have little impact, subsequent impacts would completely destroy it. Also the small projectile would occasionally seem to go through a solid surface. ”
        - “The grenade was the only projectile that felt off consistently, the initial impact of the grenade itself felt like it had too much power compared to the explosion.”
        - “The power of the impact bomb is to powerful”
        -
    - A number of responses advised the explosion was cartoony/excessive
        - “felt more like a large bullet than a bomb, often erred more towards cartoonish in how high-energy the explosions were”
        - “Felt amazing with how it impacted immediately. explosive force felt a bit excessive. ”
- **Question:** Please describe how you felt about the impact and interactions of projectile 3 (Timed explosive (red))
    - A number of responses stated this projectile was frustrating to use
        - “Most varied, sometimes hard to gauge.”
        - “took getting used to, and it was a little cartoonish with the wooden walls.”
        - “It being a ball made it functionally annoying to use and its explosive radius was massive which affected nearby walls. the explosive was excessive (could be intended), obliterating all the walls.”
        - irritating to interact with due to the roll effect without much ground friction, but created satisfying impacts. Occasionally the effects were a little too high intensity to be realistic, like with the bricks

        - Additionally as above, a number of responses advised that it had an extremely intense effect that took it into the cartoony realm more frequently.
- **Question:** Please describe how you felt about the impact and interactions of projectile 4 (Cannon Ball (yellow))
    - Most respondents advised that the projectile felt good to use
        - “feels about right, maybe a little too strong”
        - “Working just fine”
        - “Felt amazing having a fast-moving projectile with high force made perfect dents in walls. ”
        - “Very hard hitting and it is nice to play with. Heavy hitting”
    - A number of respondents advised that on occasion the damage impact would be strange, cartoonish or inconsistent.
        - “The most cartoonish. Even though its high velocity its destruction felt sometimes disproportionate”
        - “Also a little confusing, sometimes more destructive than expected, sometimes less. Mostly expected to behave like a faster stronger rock (projectile 1), but damage impact would sometimes be strange.”
- **Question:** Please describe any instances where you encountered performance issues or bugs.
    - Hitting resume on pause menu diables WSAD
    - Impact bomb would phase through solid surfaces on occasion
    - Some slowdown on large destruction
    - Timed explosive would sometimes phase through walls

2.7. Discussion

The response from the survey was varied and in a number of cases there were significant differences between respondents with more than one occasion where two respondents may rate an interaction, projectile or material at complete opposite ends of the scale. This disagreement can be observed in the two box-distribution graphs (Figures 8 and 9) for materials and projectiles that exhibit significant range between the median data values and the extremes. That being said, when taking the average of the responses given as observable in figures 6 and 7, none fell below 3 indicating that on average the materials and projectile interactions were more like what the testers expected than not. It is interesting to observe however the significant diversity in response.

Whilst it was not done in this test, it would be interesting to see trends within a respondents entire response instead of just treating each response as it’s own unique data point. I.e. Were some respondents just more positive across the board and some more negative, or were there circumstances where users who preferred wood fracture

also tended to dislike concrete fractures? A further investigation into these trends may be warranted in order to further understand player perceptions on environmental destruction.

The long answer responses are of less significance in establishing the success or failure of this project to achieve its goals due to their purely qualitative nature. That being said, these long answer responses help shed light on the areas of the projectile material interactions that were affecting different testers' opinion on the destruction one way or another. Therefore these responses can be effectively utilized as feedback to identify what aspects of the different interactions can be or need to be improved.

2.8. Conclusion

The gathered data demonstrates that the hypothesis that the destruction simulations will behave how the players expect them to more often than not, was supported. In all cases the average response/rating to the projectile and material interactions averaged greater than 3 indicating in all cases the simulations behaved more as expected than they behaved cartoony to the majority of testers.

3. Deliverable Evaluation

This project set out to create a number of high simulation fidelity environmental destruction interactions. To this end a number of metrics were created and developed in order to gauge the level of success or failure of the project to achieve this objective.

The first metric setout was related to the ability to create complex interactions between materials resulting in satisfying destruction patterns. This was not directly tested in either of the above studies; they are straight forward features that can be easily observed to be present or not in the application. The suggested interactions included:

1. Overpenetration.
    a. This was achieved. Within the project projectiles can overpenetrate walls carrying a fraction of their momentum through the wall and out the other side. Notably however this did not require much additional input from myself outside of tuning the wall and projectile velocities and resistances. This is because interactions such as the above are handled within the engine's existing physics system.
2. Projectile fragmentation
    a. This was achieved by making the projectiles also utilize chaos fractures in their creation. This allowed me to create projectiles designed to crack and fragment on impact.
3. Projectile deformation

    a. This was not achieved, deformation of meshes in this manner can potentially be handled by chaos flesh/soft bodies which is an area of the chaos physics system this project did not explore. This however would be a worthwhile and interesting expansion on the project especially if able to be applied to things like a metal like material.
4. Spalling
    a. This was achieved in a similar manner to overpenetration in the sense that with correct tailoring of the projectile and wall materials (especially damage propagation factors) spalling can be seen to occur on occasion within the project. This however was largely due to the existing chaos physics dynamics and required little focused development from myself.

A second metric utilized in the assessment of this project was the ability to operationalise a new game engine (UE5.4). This was not directly demonstrated by either of the above tests either, however the fact that a completed built and packaged application was able to be sent out to testers demonstrates clearly that this objective was also achieved.

The two primary metrics utilized to assess this project however were the metrics that the above tests were designed to assess. These metrics were

1. Maintain gameplay performance whilst dynamically simulating environmental destruction.
2. Utilize UE5 Chaos destruction in order to reduce the “Lego-Brick” look of destruction simulations made from small component conglomerates

The first of these was directly assessed through the performance tests conducted above. This metric was separated into 3 levels used to determine the scale of any success or failure of this metric to maintain 30fps, 60fps, or 90fps. As can be seen in figures 1 and 2 above, the application was able to maintain an average fps of greater than 40 in all cases. On initial inspection this may appear to satisfy the first level of this metric, however when we look at the 1% low measurements in all test cases except one this measure dropped below 30fps. These low frame rate sections indicate that the consistent minimum performance of the application was actually between 20fps and 30fps.

This infrequent drop in performance to below 30fps presents itself as stuttering/lag during hardware intensive frames such as at the point of interaction and subsequent destruction simulation. This was actually noted by one of the user testers during their play session “No issues, a little bit of slowdown when doing massive destruction, but nothing too serious.”. Whilst this is a relatively minor and common performance issue, in order for any application to utilize this method of environmental destruction en masse this performance shortfall would have to be addressed. As such it cannot be said that this metric was achieved and further optimization of the physics interaction and destruction within this application would be required.

The second metric was assessed utilizing user testing. Users were asked a number of questions regarding their thoughts on the different destructible walls and projectiles along with being asked to rate the interactions between each wall and each projectile from 1(cartoony) to 5 (behaved as expected). Whilst there was a wide variety of differing responses as demonstrated in figures 8 and 9, the average rating always remained above 3. This indicated that on average users felt that every interaction felt more like they expected it to feel than it did “cartoony”. This indicates that whilst the physics interactions were not considered perfect by the testers they were able to provide destruction that did not for the most part have a cartoony feel.

The user testing did however have some limitations due to the fact that there was a relatively low sample group of only 8 respondents. This therefore limits the accuracy of the returned results and in future additional testing would be useful to achieve more meaningful results. Unfortunately for this project there was not time to do further rounds of testing. Only a small number of testers external to the university were available, limiting the potential respondents to come mainly from the limited testing time frames available at the university. Despite the testing's limitations, I believe that it can be concluded that this metric was satisfied based upon the average user scores/ratings.

Additionally, whilst a number of user comments were gathered the utility of these as an assessment tool is low as they are highly qualitative and as such hard to draw definitive conclusions from. However, these comments do provide valuable insight into the areas that most affected testers ratings allowing for clear avenues and areas in which the application can be improved.

The final metric used to measure this application's success was if the application were able to Incorporate real world physics calculations to improve simulation fidelity of the environmental destruction. This metric was provided early on in the development process when manual physics calculations were expected to be required. However, it became evident that through the project these calculations would be directly handled by the Chaos Physics system and would only be manipulated through different parameters of the fractures and physics material applied to these fracture meshes. Therefore this metric instead utilized real world material properties such as the compressive, shear, and tensile strength of Oak, granite, tungsten, concrete,clay brick, and other materials, in order to manipulate the destruction interactions. Due to the fact that no comparative tests were conducted comparing destruction of fracture only meshes and fracture + physics material meshes it cannot be concluded if this did or did not improve simulation fidelity.

In conclusion, it is difficult to assert if this project can be considered to have successfully satisfied the metrics set out against it. Whilst the majority of metrics were completely or partially completed, other metrics vital to this application's industry relevance such as performance did not achieve their minimum goals. With that in mind however a key metric of reducing the lego brick look of the simulation destruction was achieved given

the average user ratings provided (figure 3) indicated that the simulations were more like expected than not. However, even here there are some issues as the wording of the questions asked could have further been improved by changing the descriptor “Cartoony” to “Lego-brick like”. This project was successful in building my skills as a developer within unreal engine 5 and has significantly improved my ability to create destruction simulations, understand Chaos Physics and collision handling, improve my 3D modeling knowledge, operationalize an unreal engine 5 application, and has developed my skills in designing, applying, and assessing user testing for games.

This project however was not comprehensively successful in achieving its own goals of providing high simulation fidelity and performant environmental destruction and will require significant optimisations and improvements in order to achieve this goal. If this project is developed further additional care should be taken into creating comprehensive testing plans for each of the metrics early on in development in a manner similar to the principle of Test Driven Development. This would help ensure each metric is adequately tested whilst also providing a more clear understanding of what is required of the project for it to be considered undoubtedly successful.

I believe the success of this application is indeterminate, what was achieved and what was not based on the provided metrics is roughly equal and I am unable to provide a clear and accurate assessment of its success. The application was built and provided unique and satisfying destruction, but was unable to do so in a consistently performant manner.