

UNIVERSITÀ DEGLI STUDI DI BOLOGNA



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI INFORMATICA

LAUREA TRIENNALE

in Informatica Per Il Management

A/B TESTING NEI VIDEOGIOCHI: METODI E
APPLICAZIONI PER L'OTTIMIZZAZIONE
DELL'USER EXPERIENCE

Candidato:
Giuseppe Cozza

Relatore:
Prof. Marco Roccetti

a.a. 2024/2025

*Ringrazio di cuore
i miei genitori, mamma e papà, che mi hanno sostenuto sempre con orgoglio,
la mia sorellina, persona cara e speciale,
e ovviamente il piccolo Pandino, la mascotte della nostra famiglia...*

INDICE

1. INTRODUZIONE	5
1.1 DATI SULL'INDUSTRIA DEI VIDEOGIOCHI	5
1.2 COS'È IL VERSIONAMENTO A/B.....	6
1.3 ORIGINI DEL METODO: DAL WEB MARKETING AI VIDEOGIOCHI	8
2. FONDAMENTI DELL'A/B TESTING	10
2.1 METODOLOGIA: GRUPPI DI UTENTI, RACOLTA DATI, METRICHE.....	10
2.2 DIFFERENZE CON ALTRI METODI	11
2.2.1 TEST MULTIVARIATI.....	12
2.2.2 A/B/N TESTING	12
2.2.3 USER RESEARCH	13
3. PSICOLOGIA DEL GIOCATORE E USER EXPERIENCE	15
3.1 ELEMENTI TESTABILI NEL GAMING.....	16
3.1.1 GAMEPLAY E MECCANICHE DI BASE	16
3.1.2 SISTEMA DI PROGRESSIONE	17
3.1.3 MONETIZZAZIONE ED ECONOMIA DEL GIOCO	19
3.1.4 INTERFACCIA UTENTE (UI/UX).....	21
3.1.5 ONBOARDING E PRIMA ESPERIENZA.....	22
3.1.6 GRAFICA E POST-PROCESSING	23
3.1.7 ASPETTI AUDIOVISIVI E ATMOSFERA	25
4. VANTAGGI, RISCHI E IMPLICAZIONI.....	27
4.1 RISCHI DI OVER-OPTIMIZATION	27
4.2 BILANCIAMENTO TRA DATI E CREATIVITÀ.....	27
4.3 COSTI E RISORSE DELL'A/B TESTING	28
5. CASO DI STUDIO SPERIMENTALE	29
5.1 DESIGN DELL'ESPERIMENTO	29
5.1.1 STRUTTURA DELLE VARIANTI SPERIMENTALI	29
5.1.2 ARCHITETTURA TECNICA	32

5.2 METODOLOGIA DI TESTING E RACCOLTA DATI	33
6. RISULTATI E ANALISI.....	34
6.1 PERFORMANCE DEI SITI	34
6.2 PERFORMANCE DI GIOCO.....	38
6.2.1 GRAFICA.....	38
6.2.2 TUTORIAL	39
6.2.3 UI E FONT	41
6.2.4 RETENTION ED ENGAGEMENT	43
6.3 VALIDAZIONE STATISTICA DEI RISULTATI	44
7. CONCLUSIONI E POSSIBILI SVILUPPI FUTURI	48
8. RINGRAZIAMENTI	50
9. BIBLIOGRAFIA	51

1. INTRODUZIONE

Negli ultimi decenni l'industria videoludica è cambiata profondamente portandolo da un settore guidato principalmente dalla creatività a un ecosistema molto più complesso, dove i dati giocano un ruolo importante nel processo decisionale. Questa evoluzione ha portato alla diffusione di varie metodologie scientifiche che ottimizzano l'esperienza del giocatore, tra cui spicca l'A/B testing come uno degli strumenti più efficaci per validare scelte di design e massimizzare l'engagement.

Questo lavoro vuole dimostrare come l'A/B testing sia ormai diventato un elemento chiave nello sviluppo dei videogiochi, non solo dal punto di vista tecnico e metodologico, ma anche per le implicazioni psicologiche e creative. Attraverso un caso di studio sperimentale, verrà mostrato come diverse variazioni di design possano influire in modo significativo sull'esperienza del giocatore. Per farlo è stata usata una combinazione di dati quantitativi (tempo di gioco e retention rate) e dati qualitativi raccolti tramite dei questionari compilati dai giocatori al termine della loro esperienza. Questo approccio misto permette di avere sia una misura oggettiva dei comportamenti, sia una valutazione soggettiva delle percezioni dei giocatori. In questo modo si ottiene una visione più completa dell'impatto delle diverse scelte di design.

1.1 DATI SULL'INDUSTRIA DEI VIDEOGIOCHI

Secondo i dati di Newzoo, «Il mercato globale dei videogiochi è cresciuto esponenzialmente nell'ultimo decennio, con un picco durante il periodo pandemico, e ha ora raggiunto una fase di maturità. Nel 2024, l'industria videoludica ha generato un fatturato di \$184.3 miliardi e si prevede che possa raggiungere un valore complessivo di \$198 miliardi entro il 2027»¹. Questa crescita è dovuta a un aumento della complessità nel panorama competitivo, dove migliaia di nuovi titoli vengono rilasciati quotidianamente sui vari store digitali. In questo contesto iper-competitivo, la capacità di prendere delle decisioni basate sui dati diventa un fattore fondamentale per il successo. Le aziende di sviluppo più innovative tendono a migliorare ogni aspetto del gioco analizzando metriche specifiche per ottimizzare le prestazioni.

¹ Digital Bros, "Il mercato dei videogames", 2024. <https://digitalbros.com/it/gruppo/il-mercato-dei-videogiochi/>

L'importanza dei dati è percepibile anche tramite le evoluzioni delle industrie stesse. Aziende come Riot Games hanno intere divisioni dedicate all'analisi dei dati e alla ricerca sull'user experience. Questi team di data scientist e user experience lavorano a stretto contatto con i designer e gli sviluppatori in modo da garantire che ogni elemento del gioco sia ottimizzato per massimizzare l'engagement e la soddisfazione del giocatore. Attraverso tecniche di analytics avanzate, gli sviluppatori sono in grado di tenere traccia di ogni azione del giocatore riuscendo a identificare pattern comportamentali e prevedere con precisione quando un utente è a rischio di abbandonare il gioco, permettendo interventi mirati per migliorare la retention e l'engagement.

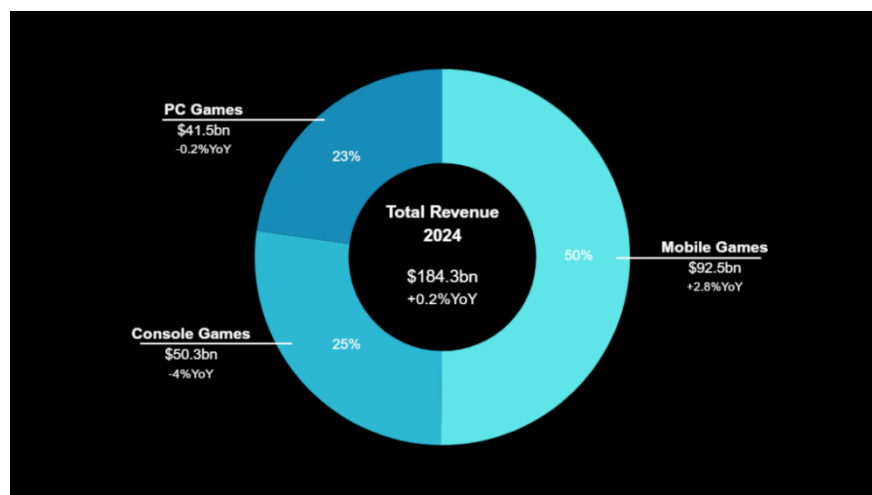


Figura 1 Fatturato dell'industria videoludica nel 2024

1.2 COS'È IL VERSIONAMENTO A/B

L'A/B testing, noto anche come split testing o versionamento A/B, è una metodologia sperimentale che consiste nel confrontare due versioni di un elemento per determinare quale performa meglio rispetto a obiettivi specifici. Nel contesto dei videogiochi, questa tecnica permette agli sviluppatori di testare varie tecniche di design, meccaniche di gioco o elementi dell'interfaccia per capire quale delle due versioni porta più engagement, retention o monetizzazione. Il principio fondamentale risiede nella suddivisione casuale degli utenti in due gruppi distinti: il gruppo di controllo (versione A) e il gruppo sperimentale (versione B). Il gruppo di controllo sperimenta la versione standard del gioco, mentre il gruppo sperimentale è esposto alla variante modificata. Attraverso la raccolta e l'analisi dei dati comportamentali dei giocatori da entrambi i gruppi, è possibile determinare statisticamente quale versione produce i risultati migliori. La forza dell'A/B testing risiede nella sua capacità

di evitare i pregiudizi o idee non oggettive (bias), facendo parlare i dati. Invece di affidarsi all'intuizione o alle preferenze personali dei designer, viene usata l'esperienza vera e propria dei giocatori per capire cosa funziona meglio e cosa meno. I dati raccolti sono empirici, poiché sono basati su fatti osservabili e misurabili.

Per considerare validi i risultati di un A/B test, è necessario che la differenza tra le due versioni sia statisticamente significativa, ovvero che non sia frutto del caso. Generalmente si usa un livello di confidenza del 95%, che indica che c'è solo il 5% di probabilità che i risultati ottenuti siano casuali. Per raggiungere questa affidabilità, è fondamentale avere un numero sufficiente di partecipanti e una durata adeguata del test. Un aspetto distintivo dell'A/B testing nei videogiochi rispetto ad altri settori è la complessità delle metriche da considerare. Mentre nel web marketing l'obiettivo è semplicemente quello di aumentare il click-through rate, nei videogiochi gli sviluppatori devono bilanciare molteplici KPI che possono a loro volta essere tra loro conflittuali: migliorare la retention potrebbe ridurre la monetizzazione nel breve termine, o aumentare la difficoltà potrebbe migliorare l'engagement per giocatori esperti ma scoraggiare i neofiti.

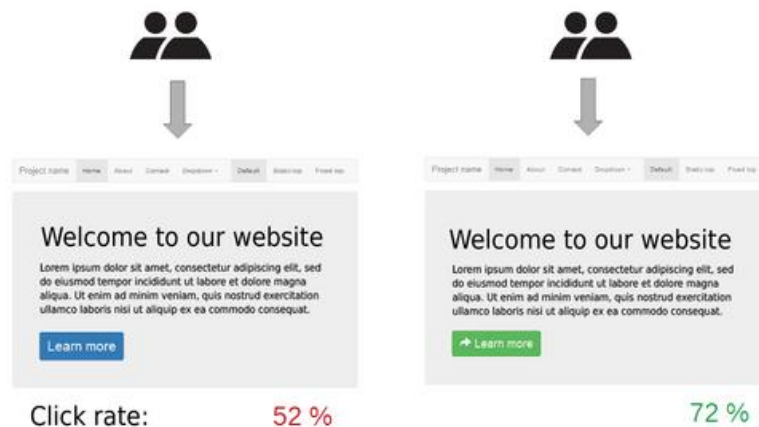


Figura 2 Esempio di A/B testing applicato a un sito web

1.3 ORIGINI DEL METODO: DAL WEB MARKETING AI VIDEOGIOCHI

«Le radici storiche sono dibattute e non esiste un consenso chiaro su chi sia stato il primo inventore»². Tuttavia, è possibile tracciare un'evoluzione che parte dalla ricerca scientifica del XX secolo e arriva fino alle applicazioni moderne nei videogiochi. I primi test nascono negli anni '20 nel campo dell'agricoltura con Ronald Fisher, statistico britannico, che sviluppò diverse tecniche per migliorare la qualità dei raccolti, testando come diversi fertilizzanti influenzassero la crescita. Per farlo, ha distribuito in modo casuale i trattamenti a diversi gruppi di piante, introducendo il concetto di randomizzazione, che permetteva di capire quale fosse il metodo migliore. Questo approccio scientifico si diffuse rapidamente nella ricerca medica e farmaceutica, dove era fondamentale dimostrare con rigore l'efficacia di nuovi trattamenti. «Nel 1938, una grande azienda statunitense di pubblicità, il Comitato Johnson, adottò il metodo di Fisher per testare efficacemente gli annunci pubblicitari»³, segnando il primo utilizzo sistematico di questa metodologia nel settore commerciale. La vera esplosione dell'A/B testing avvenne con l'avvento di Internet negli anni '90. Grazie ad esso, divenne possibile testare rapidamente diverse versioni di pagine web e misurare i risultati in tempo reale, eliminando le limitazioni dei test tradizionali su carta stampata riducendone i costi, tempi e la difficoltà nella raccolta dei dati. Google fu tra i pionieri più noti, conducendo esperimenti su ogni elemento dell'interfaccia utente. L'esempio più famoso fu il test di 41 diverse tonalità di blu per i link pubblicitari, che portò alla scoperta di una specifica sfumatura capace di generare 200 milioni di dollari aggiuntivi di fatturato annuo.

Il primo gioco a implementare sistematicamente l'A/B testing fu Farm Ville di Zynga, lanciato nel giugno 2009. Mark Pincus, il fondatore, utilizzò le metodologie del web marketing nel game design. Sotto la guida del CTO Cadir Lee, sviluppò delle capacità di analisi avanzate dei big data che permisero agli sviluppatori di implementare il modello "Game as Service". Implementarono continuamente delle nuove funzionalità, aggiungendo contenuti basati sui risultati degli esperimenti per mantenere il gioco coinvolgente a lungo termine. Post-lancio, FarmVille evolse costantemente attraverso A/B testing: dalla disposizione degli elementi dell'interfaccia ai tempi di crescita delle colture, dalla frequenza delle notifiche alle meccaniche di monetizzazione, ogni aspetto veniva modificato basandosi sui dati. In appena una settimana, raggiunse 1 milione di utenti attivi giornalieri. Il gioco arrivò a fatturare 82 milioni di utenti attivi al mese, diventando l'applicazione Facebook più

² Wikipedia, "A/B Testing", 2025. https://en.wikipedia.org/wiki/A/B_testing

³ Movylo, "Cos'è l'A/B Test", 2025. <https://www.movylo.com/it/cose-la-b-test/>

popolare del suo tempo. Questo successo dimostrò come i giochi social e basati sui dati potessero trasformarsi in enormi business. L'eredità di questo approccio si vede chiaramente anche nei videogiochi moderni. Aziende come King (con Candy Crush Saga), Supercell (Clash Of Clans) e Riot Games (League Of Legends) portano avanti ogni giorno centinaia di esperimenti, testando vari aspetti dei loro prodotti per migliorare costantemente il coinvolgimento, fidelizzazione e ricavi.

2. FONDAMENTI DELL'A/B TESTING

L'A/B testing è una tecnica di sperimentazione che nasce dal metodo scientifico e dalla statistica inferenziale. Alla base dell'A/B testing c'è il concetto di controllo sperimentale: si modifica un elemento del gioco (la variabile indipendente) per vedere come cambia comportamento o le prestazioni dei giocatori (variabili dipendenti). Per assicurarsi che i risultati siano affidabili, bisogna progettare l'esperimento in modo vigoroso, evitando che altri fattori influenzano i risultati e assicurandosi che i partecipanti siano distribuiti in modo casuale nei diversi gruppi. Nel contesto videoludico però, l'A/B testing presenta alcune particolarità rispetto ad altri settori. I giocatori non sono semplici utenti passivi ma imparano, si adattano sviluppano preferenze e modificano il loro modo di giocare nel tempo. Questo rende necessario adattare il metodo sperimentale, tenendo conto del comportamento in continua evoluzione dei giocatori.

2.1 METODOLOGIA: GRUPPI DI UTENTI, RACOLTA DATI, METRICHE

Fare A/B testing all'interno dei videogiochi richiede un metodo preciso che tenga conto delle particolarità del mondo videoludico. Si parte definendo per bene che cosa si vuole testare e quali risultati andare a misurare. Bisogna, inoltre, tenere sempre presente che i giocatori cambiano nel tempo e che chi è alle prime armi reagisce in modo diverso rispetto a chi ha più esperienza; quindi, anche i test ne devono tenere conto. Le formazioni dei gruppi sperimentali rappresentano un passaggio molto importante e richiede particolarmente attenzione. Anche se la scelta migliore rimane quella di dividere i giocatori in modo casuale (randomizzazione), bisogna fare attenzione a come lo si fa. A volte gli sviluppatori dividono prima i giocatori in gruppi in base a caratteristiche come esperienza, ore di gioco trascorse, soldi spesi e poi gli assegnano casualmente ai test. In questo modo si otterranno dei gruppi molto più equilibrati.

Nei videogiochi, in confronto ad altri ambiti, serve un numero più alto di partecipanti per ottenere risultati affidabili. Questo perché i giocatori si comportano in modo diverso tra di loro: alcuni giocano solo pochi minuti, mentre altri per ore. Questa varietà rende necessario usare dei campioni più ampi per notare davvero le differenze tra le versioni del gioco testate. Ovviamente non bisogna anche trascurare la durata degli esperimenti. In altri settori, come il marketing online, bastano solo pochi giorni per vedere i risultati. Nei videogiochi invece

servono spesso settimane o addirittura mesi, soprattutto per misurare gli effetti a lungo termine, come il numero di giocatori che continuano a giocare. Alcuni cambiamenti, infatti, mostrano i loro effetti solamente dopo che i giocatori ci hanno giocato per un po' di tempo. Anche, le metriche usate nei test A/B sono spesso molto più numerose e complesse rispetto agli altri settori. Le principali misurano quanto il gioco coinvolge il giocatore, ad esempio quanto durano le sessioni (session length), quanto spesso tornano a giocare (session frequency) e quanto velocemente avanzano nel gioco (progression rate). Le metriche di retention includono la 1-day retention (percentuale di utenti che ritornano il giorno successivo), 7-day retention e 30-day retention. Per quanto riguarda la parte dei guadagni, le metriche principali includono la percentuale di utenti che fanno acquisti nel gioco, quanto in media ogni giocatore spende (ARPU), quanto spende chi paga davvero (ARPPU) e quanto guadagna il gioco da ogni utente nel lungo periodo (LTV). Tutte queste metriche sono importanti visualizzarle insieme, altrimenti si rischia di migliorare un dato ma peggiorare l'andamento generale del gioco.

La raccolta dei dati invece, comporta delle sfide tecniche complesse, poiché c'è bisogno di tenere traccia degli eventi e delle interazioni in modo granulare (dal movimento dei personaggi ai click sull'interfaccia, dalle progressioni al comportamento sociale). A supporto di questo processo, esistono dei sistemi avanzati che monitorano le Analytics, come Unity Analytics, che permettono di registrare eventi personalizzati e standard (es. Sessioni di gioco, acquisti in-app) e di esplorarli tramite dashboard e SQL Explorer. Un altro aspetto fondamentale è la privacy: servono consenso esplicito e trasparenza nell'uso dei dati, come richiesto dal GDPR. Unity Analytics, per rispettare le normative sulla privacy, richiede che la raccolta dei dati venga attivata esplicitamente dal programmatore solo dopo aver ottenuto il consenso da parte del giocatore. In pratica, quando si apre un gioco fatto con Unity, il sistema non inizia a registrare nulla finché non vengono accettati esplicitamente (tramite un popup o un bottone) che i dati vengano raccolti. Il responsabile di creare questi popup e di rispettare le leggi della privacy (come il GDPR in Europa) è il programmatore. Unity fornisce solamente gli strumenti tecnici, ma la responsabilità legale e implementativa ricade sullo sviluppatore.

2.2 DIFFERENZE CON ALTRI METODI

Oltre all'A/B test sono presenti numerose altre metodologie di ricerca che al giorno d'oggi vengono ancora impiegate nello sviluppo dei videogiochi. Queste si distinguono per il numero di variabili da considerare, la tipologia di dati raccolti (qualitativi o quantitativi) e la complessità delle analisi necessarie.

2.2.1 TEST MULTIVARIATI

Il test multivariato (MVT) serve a testare più elementi contemporaneamente. È un tipo di esperimento molto simile a quello dell'A/B test, ma invece di andare a confrontare una singola modifica alla volta (ad esempio, due colori diversi per un bottone), vengono confrontate più variabili e varianti contemporaneamente analizzando anche tra loro tutte le possibili combinazioni (colore del bottone, dimensione, posizione, testo, etc.). Questo approccio nei videogiochi è molto potente perché permette di andare a confrontare e ottimizzare interfacce complesse o sistemi di gameplay con molteplici componenti interconnessi tra loro. Tuttavia, rappresenta anche delle grosse sfide poiché la necessità di testare tutte le possibili varianti di combinazioni richiede dei campioni estremamente ampi per raggiungere la significatività statistica. Ad esempio, testare 3 variabili ciascuna con 3 varianti richiederebbe 27 diverse condizioni sperimentali, aumentando drasticamente i requisiti di sample size. Per questo motivo, questo testing viene principalmente fatto per giochi con user base molto ampie.

2.2.2 A/B/N TESTING

L'A/B/N testing è un'estensione dell'A/B testing, dove vengono testate simultaneamente tre o più varianti (A, B, C, D, ecc.). Questa metodologia permette di avere un quadro più ampio di soluzioni in un singolo esperimento, in modo tale da ottimizzare il tempo necessario per trovare la variante più performante. Prendiamo l'esempio concreto di Clash Royale. Il gioco sta attualmente testando diversi sistemi di ricompense sui giocatori. Alcuni vedono più monete d'oro, altri ricevono più bauli, altri ancora ottengono bonus speciali. Ogni giocatore vede solo la propria versione, ma Supercell raccoglie i dati da tutte le varianti insieme per capire quale funziona meglio. Questo li porta ad avere un vantaggio temporale, perché se dovessero testare ogni sistema di ricompense in modo separato, ci vorrebbero dei mesi, mentre così proprio grazie all'A/B/N testing, ottengono i risultati una volta sola e possono decidere rapidamente quale versione implementare definitivamente nel gioco. «Stiamo testando nuove Ricompense di Battaglia! A partire da oggi, alcuni giocatori vedranno configurazioni leggermente diverse rispetto ad altri.»⁴

⁴ Fragster, "Clash Royale Battle Rewards Update", 2025. <https://www.fragster.com/clash-royale-battle-rewards-update-testing-chests-gold-more/>



Figura 3 Clash Royale Battle Rewards A/B/N testing

Ovviamente ci sono anche delle problematiche come, per esempio, il fatto che servono molti più giocatori per ottenere dei buoni risultati. Compagnie come la Supercell può permetterselo perché possiede milioni di giocatori, ma un gioco piccolo no. Proprio per questo motivo solo le grandi compagnie di videogiochi con enormi basi di utenti, come Clash Royale o Candy Crush Saga, può permettersi campioni così ampi. Un'altra problematica è legata al fatto che i risultati ricevuti sono molto più complicati. Quando vengono confrontate tante versioni insieme, è facile confondere i risultati casuali con quelli veri. Proprio per questo motivo, gli sviluppatori devono utilizzare dei calcoli statistici precisi per evitare di prendere delle decisioni sbagliate.

2.2.3 USER RESEARCH

La User Research è una metodologia che studia come i giocatori si comportano davvero mentre giocano. Invece di testare numeri e statistiche come fa l'A/B testing, qui i ricercatori osservano direttamente i giocatori e parlano con loro per capire che cosa provano. Uno degli strumenti più comuni della User Research sono i playtest, ossia delle sessioni di gioco controllate in cui un gruppo di persone prova il gioco (o una sua parte), mentre i ricercatori osservano e raccolgono i dati. Solitamente vengono selezionati un gruppo di giocatori (tra le 5–20 persone) in uno studio o da casa e li fanno giocare mentre loro li osservano. Come spiega il team di Ubisoft, la loro «missione è aiutare i designer a creare l'esperienza che

vogliono offrire»⁵. Durante i test, i ricercatori possono vedere tutto quello che fa il giocatore: dove cliccano, dove si bloccano, cosa li confonde, ecc. «Sul nostro schermo, possiamo vedere quale tasto il giocatore preme in tempo reale».⁶ Poi li fanno delle domande come per esempio: “Cosa hai pensato quando hai visto questo menu?” oppure “Perché hai scelto quest’arma?”. Ubisoft per il suo titolo Rainbow Six Siege, porta i giocatori nei loro studi di sviluppo per testare nuovi contenuti prima del rilascio. Fanno giocare sia i giocatori normali che i professionisti, filmando le loro sessioni di gioco e parlando con loro per capire cosa funziona e cosa no. «Tiene conto di molte dimensioni, come la comprensione del divertimento e del coinvolgimento del giocatore. Siamo uno strumento che potenzia i designer. Il laboratorio è diviso in diversi ruoli per raggiungere questo obiettivo. Gli User Researcher, gli analisti che definiranno e condurranno i test, costituiscono l'80% del team.»⁷

⁵ Ubisoft Paris, “Playtest Day”, 2021. <https://paris.ubisoft.com/en/blog/our-jobs-en/playtest-day/>

⁶ Ubisoft Paris, “Playtest Day”, 2021. <https://paris.ubisoft.com/en/blog/our-jobs-en/playtest-day/>

⁷ Ibid.

3. PSICOLOGIA DEL GIOCATORE E USER EXPERIENCE

Nel mondo dei videogiochi, la psicologia del giocatore è un elemento cruciale nella progettazione dell'esperienza utente (UX). I giochi non sono dei semplici prodotti digitali, ma sono ambienti interattivi che devono stimolare emozioni, motivazioni e comportamenti complessi. Proprio per questo motivo, la progettazione dell'esperienza utente in questo ambito, deve tener conto di variabili psicologiche come il senso di competenza, la ricompensa, la progressione e la socialità. «All'interno di un prodotto videoludico è una componente essenziale per permettere all'utente un'esperienza immersiva ottimizzata.»⁸ La psicologia applicata ai videogiochi ci dice che il vero “campo da gioco” non è solo lo schermo o i personaggi, ma la mente del giocatore. In pratica, come pensa e percepisce il gioco determina le sue scelte e azioni. Celia Hodent, esperta di UX per videogiochi, dice che «l'esperienza dell'utente di un videogioco avviene nella mente del giocatore».⁹ Questo significa che, quando uno sviluppatore deve progettare un gioco, deve tenere conto dei limiti della mente umana, come quante cose riesce a ricordare contemporaneamente, quando riesce a concentrarsi e come elabora le informazioni, altrimenti il gioco potrebbe diventare confusionario o frustrante. «L'UX Design si focalizza sulla psicologia del giocatore, cerca di prevedere il loro comportamento e processo di ragionamento per un'ottimizzazione di tutte le componenti di gioco.»¹⁰

La motivazione rappresenta il motore principale che spinge le persone a giocare. Psicologi come Deci e Ryan hanno sviluppato una teoria chiamata “Self-Determination Theory” che spiega cosa davvero motiva le persone. Secondo questa teoria, tutti noi abbiamo tre bisogni psicologici fondamentali:

1. **Autonomia:** Avere la libertà di poter scegliere. Nei videogiochi questo significa poter decidere come giocare, quali strategie usare e come personalizzare il nostro

⁸ Csainformazione, “UX Design nei videogiochi: cos'è e perché è importante”, 2022.
<https://csainformazione.it/ux-design-videogiochi/>

⁹ Celia Hodent, “Cognitive Psychology Applied to User Experience in Video Games”, 2016.
<https://celiahodent.com/cognitive-psychology-applied-to-user-experience-in-video-games/>

¹⁰ Csainformazione, “UX Design nei videogiochi”

personaggio. Se un gioco costringesse il giocatore a seguire un unico percorso e ad usare un'unica strategia potrebbe risultare frustrante.

2. **Competenza:** Vogliamo sentirci bravi in quello che facciamo. Un gioco deve farci sentire che stiamo progredendo e che le nostre abilità crescono. Se un gioco è troppo facile ci annoiamo, se è troppo difficile ci scoraggiamo.
3. **Relazione sociale:** Vogliamo essere interconnessi. Proprio per questo motivo i giochi multiplayer, le classifiche e i sistemi di chat sono così popolari.

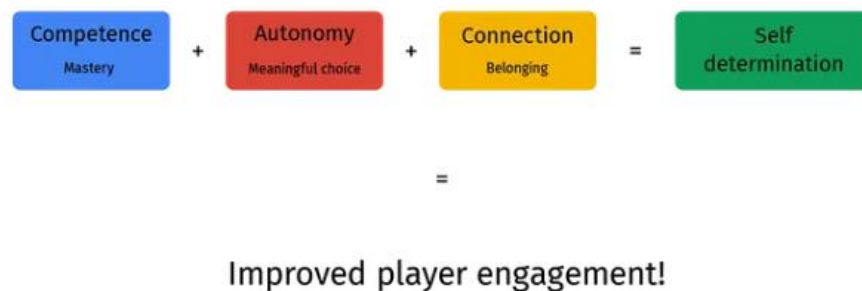


Figura 4 Self-Determination Theory

Quando un gioco soddisfa tutti questi bisogni, i giocatori sono naturalmente motivati a continuare a giocare. L'A/B testing serve proprio a identificare quale versione del gioco risponde meglio a queste esigenze.

3.1 ELEMENTI TESTABILI NEL GAMING

Attraverso l'A/B testing è possibile testare praticamente ogni aspetto dell'esperienza di gioco. Questa metodologia dà la possibilità agli sviluppatori di sperimentare con qualsiasi elemento, dalle meccaniche base fino ai dettagli più sottili dell'interfaccia utente e persino aspetti economici come i prezzi e acquisti in-game. Questa ottimizzazione è possibile realizzarla basandosi su dati reali piuttosto che su supposizioni, rendendo il processo di sviluppo più preciso e orientato alle reali preferenze dei giocatori.

3.1.1 GAMEPLAY E MECCANICHE DI BASE

Le meccaniche di gioco sono spesso oggetto di test per ottimizzare il bilanciamento e il coinvolgimento. Questi elementi determinano come i giocatori interagiscono con il mondo di gioco e la loro ottimizzazione può decretare il successo o meno di un titolo.

La sensazione del controllo è cruciale per una buona esperienza di gioco e, proprio per questo motivo, Riot Games ha recentemente testato l'introduzione di controlli WASD per il suo gioco League Of Legends, confrontandoli con il sistema tradizionale del point-and-click. L'azienda ha implementato una metodologia di A/B testing su più fasi specificatamente progettata per gestire il rischio di una modifica così significativa. La prima fase ha coinvolto dei test interni, nuovi giocatori e pro-player, per identificare i problemi tecnici di base. La seconda fase prevede un testing controllato con un gruppo selezionato (Public Beta Environment) per 1-2 mesi monitorando metriche specifiche come bilanciamento dei campioni e interazioni dei colpi. La terza fase invece introduce i controlli WASD nelle modalità non ranked in modo da sperimentare senza una pressione competitiva, permettendo alla Riot di raccogliere dei dati su una scala molto più ampia. Durante queste implementazioni la Riot si assicurerà che «ascolterà in maniera continua i feedback dei giocatori, per assicurarsi che non ci siano bug e un vantaggio di uno schema di controllo rispetto all'altro gestendone il bilanciamento¹¹».

3.1.2 SISTEMA DI PROGRESSIONE

La progressione rappresenta uno degli aspetti più complessi e cruciali dell'A/B testing, poiché è ciò che motiva i giocatori a continuare nel lungo periodo e influenza direttamente sia la retention che la monetizzazione. Questi sistemi devono essere bilanciati con precisione estrema per avere un equilibrio tra sfida e gratificazione, un processo che richiede test continui e aggiustamenti basati su dati empirici. Supercell ha sviluppato uno degli approcci più sofisticati in questo ambito, andando a segmentare la propria base di utenti. Come rivelato da Lasse Lauhento, «l'azienda divide il pubblico in tre categorie: principianti, fascia media, e gruppi élite. Ognuno di questi gruppi gioca in modo diverso e ha tassi di retention diversi». ¹² Questa suddivisione in categorie, permette alla Supercell di sperimentare sistemi di progressione differenziati per ogni gruppo, scoprendo che esiste un “punto dolce” critico nella velocità di avanzamento. Una progressione troppo veloce porta al burnout dei giocatori, che completano velocemente tutti i contenuti disponibili e abbandonano il gioco, mentre una progressione troppo lenta causa una frustrazione e un abbandono precoce.

¹¹ Riot Games, “WASD Controls Are On the Way”, 2025. <https://www.leagueoflegends.com/en-us/news/dev/dev-wasd-controls-are-on-the-way/>

¹² PocketGamer, “1, 7 and 30-day retention rates are key”, 2012. <https://www.pocketgamer.biz/1-7-and-30-day-retention-rates-are-key-says-tablet-first-supercell/>

Il core loop di Clash Of Clans rappresenta un esempio di come l'A/B testing possa ottimizzare sistemi di progressione complessi. Supercell ha costruito il gioco attorno a tre azioni fondamentali che si autoalimentano: raccolta delle risorse, costruzione e addestramento, e combattimento. «Una sessione molto veloce di 30 secondi potrebbe riguardare solo la raccolta di risorse, mentre una sessione completa con attacco e preparazione di nuove truppe richiede meno di 5 minuti»¹³. Questo è il risultato di estensivi A/B test sui tempi di ogni fase del loop. Supercell ha sperimentato diverse durate per la produzione delle risorse, tempi di costruzione degli edifici e l'addestramento delle truppe, scoprendo che questo sistema che inizialmente risulta veloce all'inizio e che col progredire del tempo rallenta, crea degli obiettivi potenti a medio e lungo termine per i giocatori. Questo bilanciamento tra gratificazione immediata (raccolta delle risorse) e obiettivi a lungo termine (upgrade costosi), ha creato quello che i game designer chiamano “dopamine-seeking reward loop”, che è in grado di mantenere i giocatori coinvolti attraverso ricompense sia immediate che dilazionate nel tempo.



Figura 5 Clash of Clans core loop

Anche la monetizzazione integrata nei sistemi di progressione richiede particolarmente attenzione nell' A/B testing, poiché deve bilanciare la generazione di ricavi con la soddisfazione del giocatore. Supercell nel loro gioco Clash Of Clans ha condotto A/B test estensivi su questo meccanismo psicologico, testando diverse quantità di gemme iniziali gratuite, diversi rapporti di conversione gemme-risorse e diversi timing per le prime offerte

¹³ GameAnalytics, “Mid-Core Success”, 2023. <https://www.gameanalytics.com/blog/mid-core-success-part-1-core-loops>

di acquisto. Grazie a questi esperimenti, hanno dimostrato l'efficacia di questo approccio. «Il gioco inizialmente fornisce delle gemme gratuite che vengono spese per oro ed elisir, creando l'abitudine mentale all'acquisto. Quando le gemme gratuite finiscono, è più facile convincere il giocatore a spendere soldi veri per comprarne altre, perché ormai ne conosce l'utilità e ne percepisce il valore»¹⁴. Questo meccanismo ha contribuito a generare oltre \$6,5 miliardi in ricavi lifetime, dimostrando come l'integrazione strategica della monetizzazione nel sistema di progressione possa essere un circolo virtuoso dove il pagamento non viene vissuto come “obbligatorio per progredire”, ma come un modo per accelerare i progressi e ottenere prima ciò che si desidera.

L'A/B testing della difficoltà rappresenta un altro elemento fondamentale nei sistemi di progressione moderni. King implementa esperimenti continui su Candy Crush Saga attraverso un sistema automatizzato che «monitora costantemente tutti i livelli quando vengono rilasciati, controllando quanti giocatori superano un livello specifico e quanto tempo impiegano per superarlo. Se è troppo facile o troppo difficile, “modificano” un po' i livelli.»¹⁵ Vengono divisi automaticamente i giocatori in gruppi di controllo e sperimentali, testando versioni alternative dello stesso livello con diversi parametri di difficoltà. Questo sistema di monitoraggio continuo, consente degli aggiustamenti della difficoltà in tempo reale, ma la vera innovazione risiede nell'utilizzare dei modelli predittivi per anticipare i potenziali problemi di bilanciamento, permettendo interventi che mantengono l'esperienza di progressione ottimale senza nessuna interruzione percettibile dal giocatore.

3.1.3 MONETIZZAZIONE ED ECONOMIA DEL GIOCO

La monetizzazione rappresenta uno degli aspetti più cruciali e direttamente misurabili dell'A/B testing nei videogiochi. Per questo motivo, chi fa test deve essere molto preciso: ogni elemento dell'economia del gioco va testata con esperimenti controllati cercando di aumentare i guadagni ed evitando di peggiorare l'esperienza del giocatore. Il trucco sta nel trovare un equilibrio tra fare profitto e fare divertire i giocatori, cosa che si ottiene con tanti test concreti. Uno degli aspetti più testati è il prezzo degli acquisti in-app. Gli sviluppatori, infatti, conducono A/B test continui su diversi punti di prezzo e «sperimentano con diverse strategie di pricing, inclusi prezzi fissi, abbonamenti, o pricing dinamico basato sul

¹⁴ GameDeveloper, “Game monetization design: Analysis of Clash of Clans”, 2014.

<https://www.gamedeveloper.com/business/game-monetization-design-analysis-of-clash-of-clans>

¹⁵ Mobile Gamer, “How King defines a ‘good’ Candy Crush Saga level”, 2024. <https://mobilegamer.biz/how-king-defines-a-good-candy-crush-saga-level-and-why-it-constantly-prunes-the-bad-ones/>

comportamento dei giocatori o condizioni di mercato»¹⁶. Questi test non si limitano a confrontare soltanto dei prezzi singoli, ma esplorano delle strategie più complesse come i prezzi dinamici che cambiano in base al comportamento del giocatore, il loro storico delle spese e la loro propensione a fare altri acquisti, e persino il loro tempo trascorso nel gioco.

Le strategie di bundling rappresentano un altro aspetto critico dove l'A/B testing dimostra il suo valore. Gli sviluppatori testano costantemente «diverse combinazioni di oggetti nei bundle per vedere quale risuona meglio. Provano anche ad aggiungere oggetti esclusivi a tempo limitato per aumentare il valore percepito»¹⁷. Supercell, per esempio, ha perfezionato questa pratica scoprendo che i pacchetti che combinano le risorse immediate (come oro e gemme) con oggetti che aiutano a progredire nel gioco, generano il 40% in più di acquisti rispetto ai pacchetti standard. L'azienda testa anche la "psicologia dello sconto", sperimentando con diverse percentuali di sconto (15%, 25%, 40%). Anche l'integrazione della pubblicità all'interno dei videogiochi tramite A/B testing richiede una maggiore attenzione per non infastidire i giocatori e compromettere la loro esperienza utente. Bisogna difatti «sperimentare con diverse versioni delle funzionalità in-game, pricing e posizionamenti pubblicitari per vedere quale produce i migliori risultati»¹⁸. Gli sviluppatori testano non solo la frequenza degli annunci, ma anche il timing ottimale per la loro visualizzazione. Ad esempio, uno studio condotto sui giochi casual ha rivelato che un annuncio che viene mostrato immediatamente dopo una vittoria genera il 25% in più di engagement rispetto a quelli mostrati dopo una sconfitta, poiché i giocatori sono in uno stato emotivo più positivo. Inoltre, l'A/B testing ha dimostrato che i rewarded video ads (video pubblicitari al termine della quale offrono una ricompensa), possono effettivamente aiutare anche a riconoscere funzionalità premium che altrimenti non avrebbero provato, aumentando la probabilità di acquisti futuri.

Infine, abbiamo la segmentazione degli utenti che permette di personalizzare le strategie di monetizzazione, ma che richiede però A/B testing multivariati su larga scala. Gli sviluppatori non usano più un'unica soluzione per tutti, ma implementano strategie diverse per tipi diversi

¹⁶ Segwise, "Effective Strategies to Boost In-App Purchases in Mobile Games", 2024. <https://segwise.ai/blog/effective-strategies-to-boost-in-app-purchases-in-mobile-games-clxbr489d0010d2agh26cvdy0>

¹⁷ Segwise, "Best Practices for Pricing and Bundling in Mobile Games", 2024. <https://segwise.ai/blog/best-practices-for-pricing-and-bundling-in-mobile-games-clxo9g5vs001cagmd5b99bm71>

¹⁸ Chartboost, "How to successfully implement in-app purchases and in-app advertising", 2023. <https://www.chartboost.com/resources/guides/balancing-monetization-how-to-successfully-implement-in-app-purchases-and-in-app-advertising/>

di giocatori. I “whales” (giocatori che spendono molto) vedono offerte costose e pacchetti premium, mentre i giocatori occasionali ricevono principalmente micro-transazioni a basso costo e annunci premiati. Questi tipi di test sono complessi, perché richiedono di gestire tante variabili insieme: tipo di giocatore, prezzi personalizzati, momenti delle offerte e tipi di annunci, spesso testandone dozzine di combinazioni su centinaia di migliaia di giocatori contemporaneamente.

3.1.4 INTERFACCIA UTENTE (UI/UX)

L’interfaccia utente (UI) rappresenta il punto di contatto diretto tra il giocatore e l’esperienza di gioco, rendendo l’A/B testing in questo ambito particolarmente fondamentale. Un’interfaccia ben progettata dovrebbe essere “invisibile” all’utente, permettendo un’interazione fluida e intuitiva, mentre una mal progettata può compromettere irreparabilmente l’engagement anche del gioco più ben sviluppato dal punto di vista meccanico. Epic Games ha sviluppato una delle metodologie di A/B testing più sofisticate per l’ottimizzazione dell’UI/UX, utilizzando un approccio scientifico che combina test empirici con principi di psicologia cognitiva applicata al game design. Il processo di testing dell’interfaccia di Fortnite illustra perfettamente la complessità dell’A/B testing UI/UX nei videogiochi moderni. Come spiegato da Celia Hodent, ex Director di UX di Epic Games, «non cerano pratiche UX e neanche strumenti specifici»¹⁹, ma l’azienda ha sviluppato una metodologia che si basa sulla comprensione delle limitazioni cognitive umane. Un esempio concreto di questo approccio emerso durante lo sviluppo di Fortnite, quando l’azienda ha scoperto che «il simbolo originale che era stato scelto per rappresentare le trappole nel gioco non veniva percepito come tale da tutti i giocatori che testarono il gioco. Alcuni giocatori pensavano che sembrassero munizioni, o alberi»²⁰. Questo ha portato a un ripensamento completo del design iconografico, dimostrando come anche elementi apparentemente minori dell’interfaccia richiedano validazione empirica mediante A/B testing.

Il principio di Epic Games è confrontare il “modello mentale” dei designer con quello dei giocatori: spesso i designer capiscono intuitivamente qualcosa che i giocatori non percepiscono nello stesso modo. Per questo motivo, l’azienda implementa test sistematici su ogni elemento dell’interfaccia, dalla disposizione dei menu alla comunicazione visiva degli obiettivi di gioco. L’ottimizzazione per diverse piattaforme rappresenta una sfida particolare

¹⁹ Celia Hodent, “Developing UX Practices at Epic Games”, 2015. <https://celiahodent.com/ux-practices-epic-games/>

²⁰ Medium, “Understanding the success of Fortnite”, 2020. <https://medium.com/ironsource-levelup/understanding-the-success-of-fortnite-a-ux-user-experience-psychology-perspective-208cee587dc6>

nell'A/B testing UI/UX, specialmente per giochi cross-platform come Fortnite. L'interfaccia deve essere «progettata per essere adattiva, garantendo un'esperienza coerente sia che i giocatori siano su smartphone, tablet, o persino console. Gli elementi dell'UI si adattano automaticamente alla dimensione dello schermo, e i layout dei controlli sono ottimizzati per il touch sui dispositivi mobile pur rimanendo intuitivi su schermi più grandi o quando si usa un controller.»²¹ Questo richiede A/B testing multiplatforma dove le stesse funzionalità vengono testate simultaneamente su dispositivi diversi per assicurare che l'esperienza rimanga ottimale indipendente dalla piattaforma di gioco. Epic Games testa frequentemente elementi come la dimensione dei pulsanti touch, il posizionamento degli elementi HUD, e la leggibilità dei testi su schermi di diverse dimensioni, utilizzando metriche specifiche come il tempo di reazione ai comandi e la precisione dell'input per ogni variante testata. Anche il carico cognitivo gioca un ruolo importante: se il giocatore deve ricordare troppe cose o fare troppi passaggi per compiere un'azione, l'esperienza sarà peggiore.

3.1.5 ONBOARDING E PRIMA ESPERIENZA

L'A/B testing dell'esperienza di prima utilizzazione (FTUE – First Time User Experience) è diventato così sofisticato che gli sviluppatori testano ogni microsecondo dell'interazione iniziale, riconoscendo che «quasi il 20% dei giocatori non completa nemmeno la prima quest nel tutorial»²². Questa statistica drammatica ha spinto l'industria a sviluppare strategie A/B testing estremamente granulari per l'onboarding, dove ogni elemento viene testato andando a tracciare ogni singola azione del giocatore dal momento del primo avvio dell'app. Anche la struttura e la durata del tutorial sono diventate oggetto di A/B testing intensivi, con risultati che spesso contraddicono l'intuizione degli sviluppatori. Alcuni studi sull'industria di giochi per dispositivi mobile ha rivelato che i tutorial interattivi durante un gameplay mostrano un +45% di completion rate rispetto ai tutorial passivi, mentre tutorial brevi (3-5 minuti) generano un 60% di retention rate rispetto a tutorial lunghi (10-15 minuti). Leo Kihlman, Game Director di Nitro Games, ha sviluppato un approccio estremo: far provare il gioco senza tutorial a utenti completamente estranei al genere, per capire se il gameplay fosse abbastanza intuitivo da spiegarsi da solo.

Il timing delle ricompense durante la fase di onboarding ha mostrato effetti significativi: la distribuzione progressiva delle ricompense, invece di una gratificazione finale unica,

²¹ Kreonit, “5 must-have mobile game UI features”, 2024. <https://kreonit.com/testing-and-qa/mobile-game-ui/>

²² Udonis, “First-Time User Experience (FTUE) nei giochi”, 2024. <https://www.blog.udonis.co/mobile-marketing/mobile-games/first-time-user-experience>

mantiene l'utente motivato nei momenti più critici dell'esperienza iniziale. Questo principio è stato applicato da sviluppatori come Wooga nel loro gioco June's Journey, dove hanno implementato un sistema di microricompense che mantengono i giocatori motivati in ogni step del tutorial. L'A/B testing ha anche dimostrato che la presenza di «barre di progresso e checklist è un incentivo visivo che incoraggia le persone a continuare e completare il processo di tutorial/onboarding»²³, con variazioni che testano vari stili di visualizzazioni di progresso, dal semplice contatore percentuale a sistemi di achievement gamificati.

Anche l'economia durante la fase di tutorial richiede un bilanciamento particolare delicato e testato mediante A/B testing dei “tap e sink”. Come sottolineato da Udonis, «alcune altre metriche FTUE che dovrebbero tracciare sono le metriche dell'economia di gioco come tap e sink. Un tap è da dove proviene una risorsa del gioco mobile... i sink sono tutti i luoghi dove i giocatori possono spendere risorse»²⁴. L'A/B testing in questo contesto determina la quantità ottimale di risorse gratuite da fornire durante l'onboarding: troppe risorse creano un'aspettativa irrealistica per il gameplay futuro, mentre troppe poche potrebbero frustrare i giocatori prima che comprendano bene il valore del gioco. Gli sviluppatori testano sistematicamente «quanti crediti gratuiti (valuta in-game) offrire per garantire l'equilibrio ideale tra utenti felici (e sperabilmente, che ritornano) e aumento della redditività»²⁵, utilizzando la segmentazione per determinare i valori ottimali per diversi tipi di giocatori e testando l'impatto di queste decisioni sulla retention e monetizzazione a lungo termine.

3.1.6 GRAFICA E POST-PROCESSING

La grafica rappresenta il primo impatto visivo che un giocatore ha con un videogioco e l'A/B testing in questo ambito si è rivelato fondamentale per determinare quali fattori e scelte estetiche massimizzano l'engagement e la retention. L'obiettivo è testare ogni aspetto visivo, dalla palette dei colori agli effetti post-processing. L'impatto della qualità grafica sull'esperienza di gioco è stato documentato attraverso studi sulla Quality Of Experience (QOE). Una ricerca che ha comparato giochi con grafica high-quality e low-quality ha dimostrato che «gli utenti riportano costantemente di sentirsi più connessi quando interagiscono con giochi che offrono grafica di alta qualità. Se la qualità grafica ricevuta è

²³ Udonis, “First-Time User Experience (FTUE) nei giochi”, 2024. <https://www.blog.udonis.co/mobile-marketing/mobile-games/first-time-user-experience>

²⁴ Ibid.

²⁵ Vwo, “What is Mobile App A/B Testing?”, 2025. <https://vwo.com/mobile-app-ab-testing/>

bassa, la valutazione dell'utente per quel particolare gioco è bassa»²⁶. Le scelte estetiche non solo influenzano le aspettative dei giocatori, ma anche la loro connessione con il gioco. Gli effetti grafici avanzati, come quelli utilizzati nel post-processing, sono un altro aspetto importante nel design di un videogioco. Sebbene il loro obiettivo sia migliorare l'aspetto visivo, l'effetto che producono può variare molto tra i giocatori. Per questo motivo, è fondamentale testare questi effetti per capire come influenzano l'esperienza.

La scelta dello stile grafico richiede un A/B testing particolarmente attento poiché influisce direttamente sulla percezione del gioco da parte degli utenti. CrazyLabs, uno dei principali publisher di giochi mobile con oltre 6 miliardi di download, sottolinea come gli sviluppatori non sempre rappresentino il target di giocatori a cui si rivolgono. «Gli sviluppatori di giochi hyper-casual non sono necessariamente giocatori hyper-casual. Questo significa che non sei il tuo target di riferimento, quindi non fare supposizioni - testa tutto»²⁷.

Un caso emblematico riguarda la scelta tra grafica high-end e stile semplificato: «potresti volere grafica di alta qualità per il tuo gioco, ma i giocatori hyper-casual potrebbero interpretare la tua bellissima grafica 3D diversamente, e sospettare che il tuo gioco sia destinato a un pubblico di giocatori più hardcore»²⁸. Questo principio è dimostrato attraverso test che confrontano stili artistici opposti come nel caso di “Hyper Hill Climb” dove la versione con grafica cartoon stilizzata (ambiente semplificato con colori piatti) viene testata contro una versione con rendering realistico (terreno dettagliato con texture complesse e illuminazione avanzata).

²⁶ MDPI, “Quality of Experience That Matters in Gaming Graphics”, 2024. <https://www.mdpi.com/2079-9292/13/15/2998>

²⁷ CrazyLabs, “5 Tips for Mobile Games A/B Testing”, 2023. <https://www.crazylabs.com/blog/5-tips-for-mobile-games-a-b-testing/>

²⁸ Ibid.



Figura 6 Hyper Hill Climb, test sulla grafica di gioco

La ricerca sui KPI dei giochi mobile conferma che «metriche di esperienza utente come tempi di caricamento, frequenza degli errori e crash rate sono spesso utilizzate come indicatori predittivi per determinare altri KPI come engagement e monetizzazione»²⁹. Questo significa che la scelta stilistica ha un impatto diretto su questi parametri: una grafica complessa può causare problemi di performance sui dispositivi entry-level, riducendo l'accessibilità e influenzando negativamente la retention.

3.1.7 ASPETTI AUDIOVISIVI E ATMOSFERA

Il suono e l'atmosfera non servono solo a “decorare” un videogioco, ma contribuiscono in modo decisivo a come il giocatore vive l'esperienza. La musica, gli effetti sonori, e l'ambiente grafico, possono cambiare le emozioni di chi gioca e influenzare la sua permanenza nel gioco. Proprio per questo motivo anche l'audio deve essere testato con attenzione tramite A/B testing. A volte, infatti, le differenze non sono subito percepibili in modo consapevole, ma incidono comunque sul comportamento dei giocatori. Uno studio condotto su 20 partecipanti ha dimostrato che «l'uso di effetti audio sembra migliorare

²⁹ MetricFire, “The Most Important KPIs for Monitoring Mobile Games”, 2025.

<https://www.metricfire.com/blog/the-most-important-kpis-for-monitoring-mobile-games/>

l'immersione dei giocatori nel videogioco»³⁰, confrontando direttamente esperienze con e senza elementi sonori. Ricerche più approfondite hanno rivelato che anche diversi componenti audio hanno impatti differenziati: «la musica di sottofondo porta a maggiore immersione, mentre i suoni ambientali creano un ambiente acustico vibrante e realistico che attira le persone»³¹.

Un esperimento condotto su League Of Legends ha implementato una metodologia rigorosa dove i partecipanti hanno giocato alla modalità di allenamento per giocatore singolo in quattro diverse condizioni sonore (tutte le combinazioni di suoni ambientali e suoni del personaggio attivati o disattivati) e hanno fornito feedback su vari aspetti dell'esperienza di gioco dopo ogni sessione. Lo studio ha misurato tre dimensioni chiave: l'immersione (intesa come senso di appartenenza al mondo di gioco virtuale), l'identificazione dell'avatar (connessione emotiva con il personaggio controllato), e il divertimento (piacevolezza complessiva dell'esperienza). I risultati dello studio hanno evidenziato che l'uso degli effetti sonori, sia ambientali che legati al personaggio, ha migliorato l'immersione e il coinvolgimento del giocatore, aumentando la sensazione di “essere parte” del mondo di gioco.

³⁰ SpringerNature, “Game audio — an investigation into the effect of audio on player immersion”, 2013. <https://link.springer.com/article/10.1007/BF03392342>

³¹ Mdpi, “Exploring the Impact of Ambient and Character Sounds”, 2024. <https://www.mdpi.com/2076-3417/14/2/583>

4. VANTAGGI, RISCHI E IMPLICAZIONI

L'A/B testing nei videogiochi non è solo una tecnica per migliorare le vendite, ma ha cambiato completamente il modo in cui si fanno i giochi. Come ogni strumento potente, porta dei vantaggi ma anche dei rischi che bisogna conoscere per usarlo bene. L'adozione di questo metodo ha spostato l'attenzione degli sviluppatori dalle decisioni basate sull'intuizione a scelte guidate dai dati, migliorando in modo concreto la performance dei prodotti. Allo stesso tempo, ha imposto una riflessione sulla creatività nel game design, richiedendo di trovare un equilibrio tra innovazione artistica e ottimizzazione basata su dati empirici, senza rischiare che i giochi diventino troppo simili tra loro.

4.1 RISCHI DI OVER-OPTIMIZATION

Uno dei principali rischi associati all'A/B testing è l'over-optimization, ovvero la tendenza a modificare continuamente gli elementi di gioco per massimizzare singole metriche, come la retention o il tasso di conversione, trascurando però la qualità complessiva dell'esperienza. Questo fenomeno può portare a giochi che, pur essendo performanti dal punto di vista tecnico, risultano privi di originalità. Un esempio concreto è l'uso eccessivo di strategie di monetizzazione, come la visualizzazione aggressiva degli annunci o la spinta verso acquisti in-app, che possono far crescere i ricavi nel lungo periodo, ma che a lungo andare possono allontanare i giocatori più affezionati e danneggiare la reputazione di gioco. Come sottolineato da Jens-Fabian Goetzmann, «affidarsi esclusivamente a tecniche di ottimizzazione come l'A/B testing significa rischiare di cadere in una mentalità di solo miglioramento incrementale, che non aiuta a fare passi audaci verso una visione innovativa»³².

4.2 BILANCIAMENTO TRA DATI E CREATIVITÀ

Il vero valore dell'A/B testing nel game design sta nella possibilità di integrare dati e creatività all'interno di un processo continuo e dinamico. Se da un lato i dati forniscono le indicazioni su cosa funziona meglio e cosa no, dall'altro sono le intuizioni creative a dare vita a esperienze di gioco memorabili e significative. Un esempio efficace di questa integrazione è Apex Legends, che ha introdotto innovazioni come il sistema di ping grazie

³² Jefago, "You Can't A/B Test Your Way to Greatness", 2020. <https://www.jefago.com/product-management/ab-testing-does-not-create-great-products/>

alla quale è possibile segnalare un punto o un oggetto sulla mappa senza usare la chat vocale. Nonostante inizialmente non fosse previsto, il successo di questa feature è stato in parte dovuto alla capacità del team di ascoltare i feedback dei giocatori e adattare il gioco di conseguenza. Tuttavia, trovare il giusto equilibrio tra dati e creatività non è sempre semplice. Come evidenzia il blogger e game designer Sande Chen, «i dati forniscono informazioni preziose sul comportamento dei giocatori, ma non possono sostituire l'innovazione artistica o la visione del designer»³³. Pertanto, è fondamentale che i designer mantengano una visione coerente e chiara del gioco, utilizzando i dati come uno strumento per affinare e perfezionare la loro creatività, senza lasciarsi guidare esclusivamente da essi. È per questo che alcuni studi parlano oggi di approccio data-informed piuttosto che data-driven: i dati devono informare le decisioni, non determinarle in maniera assoluta.

4.3 COSTI E RISORSE DELL'A/B TESTING

Un altro rischio legato all'A/B testing riguarda i costi economici e organizzativi che questa pratica comporta. Per condurre un test efficace, è necessario sviluppare delle versioni alternative degli elementi di gioco da confrontare, predisporre strumenti di raccolta e analisi dei dati e gestire l'esperimento su un campione di utenti sufficientemente ampio da garantire la significatività statistica. Tutto ciò richiede tempo, risorse umane e un investimento economico che può risultare oneroso, soprattutto per le aziende di piccole dimensioni o gli studi indipendenti. Inoltre, non è sempre detto che i test possano portare dei benefici concreti. A volte possono risultare inconcludenti o non generare cambiamenti importanti, con il rischio di aver investito energie senza nessun guadagno reale. Per questo motivo è fondamentale valutare con attenzione se l'A/B testing convenga davvero oppure no.

³³ GameDeveloper, “Balancing Data and Design”, 2017.
<https://www.gamedeveloper.com/business/balancing-data-and-design>

5. CASO DI STUDIO SPERIMENTALE

Per analizzare in modo pratico l'applicazione dell'A/B testing nei videogiochi, è stato realizzato un gioco di sopravvivenza in 2D, sviluppato in due varianti. L'obiettivo principale è dimostrare come differenti scelte di game design possano influenzare significativamente l'esperienza del giocatore e le relative metriche di performance, fornendo un caso pratico delle metodologie teoriche precedentemente analizzate. Dal punto di vista metodologico, lo studio non presenta un A/B testing "puro", ma piuttosto uno split testing non randomizzato. Infatti, i partecipanti hanno scelto autonomamente quale versione provare rispetto all'approccio classico dell'A/B testing. Nonostante questo limite, la ricerca si basa comunque sui principi fondamentali della metodologia, con l'obiettivo di valutare l'impatto di diverse filosofie di design sull'engagement e sulla preferenza dei giocatori. L'analisi raccoglie sia dati quantitativi che dati qualitativi, tramite la raccolta di questionari, offrendo una valutazione più completa delle due varianti di design.

5.1 DESIGN DELL'ESPERIMENTO

L'architettura consiste nella creazione di due versioni (A e B) di un gioco di sopravvivenza in 2D, che differiscono tra loro in molteplici campi, mettendo a confronto due filosofie di game design differenti. Anche le pagine da cui sarà possibile scaricare il gioco hanno dei tratti stilistici distinti tra loro, grazie alla quale ci sarà la possibilità di capire quanto la coerenza grafica e uno stile curato possa avere una buona percezione, o meno, iniziale degli utenti.

5.1.1 STRUTTURA DELLE VARIANTI SPERIMENTALI

La versione A implementa un sistema di gameplay basato su onde progressive, dove i nemici vengono generati in gruppi con pause strategiche tra un'ondata e la successiva. Ogni azione del giocatore è accompagnata da effetti visivi e sonori, insieme a un effetto di camera shake che rende l'esperienza più intensa. L'interfaccia mantiene uno stile coerente con la grafica pixel art, grazie a una palette di colori uniforme e a un font in linea con lo stile. Anche l'audio è curato, con suoni per le interazioni dell'utente e una colonna sonora pensata per seguire il ritmo delle ondate. È presente anche un tutorial integrato che guida i nuovi giocatori attraverso le meccaniche fondamentali.

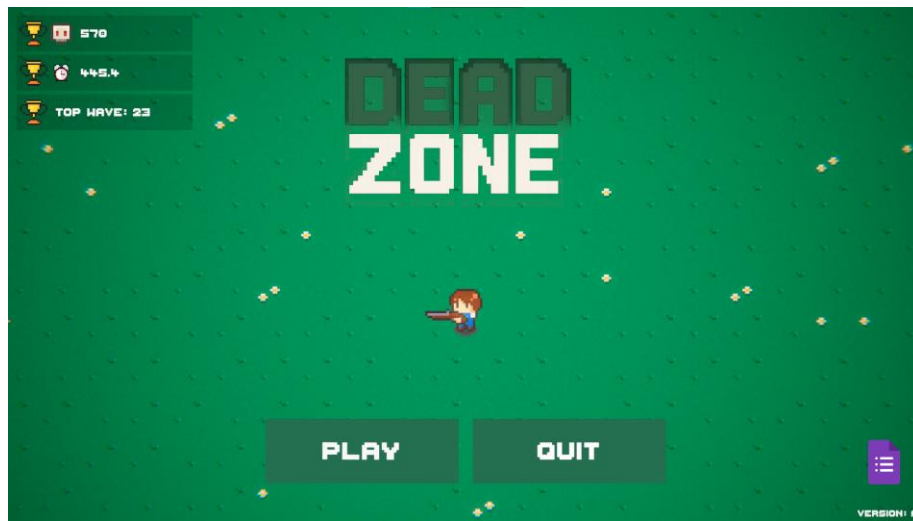


Figura 7 Versione di gioco A

La pagina di itch.io, da cui è possibile scaricare il gioco, riflette questo approccio di alta qualità attraverso un layout professionale e in linea con l'intero stile del design. Viene utilizzata una palette cromatica armoniosa in linea con i vari screenshot del gameplay e con i colori dell'immagine di sfondo, includendo una tipografia consistente e una chiara disposizione delle informazioni sulla trama e sui comandi del gioco.



Figura 8 Game Store versione A

La versione B, al contrario, adotta un approccio di generazione continua di nemici, eliminando qualsiasi pausa naturale nel flusso di gioco. In questa variante sono stati ridotti elementi di feedback considerati “non essenziali” come: il camera shake, gli effetti particellari per le eliminazioni dei nemici e, inoltre, il sistema di UI opera senza feedback

sonoro. Il font e i colori dell'interfaccia differiscono significativamente dalla versione principale poiché non seguono lo stile pixel art del gioco, creando una discontinuità stilistica intenzionale. Un'altra differenza principale con la versione A, è che in questa versione non è presente una fase di tutorial iniziale.

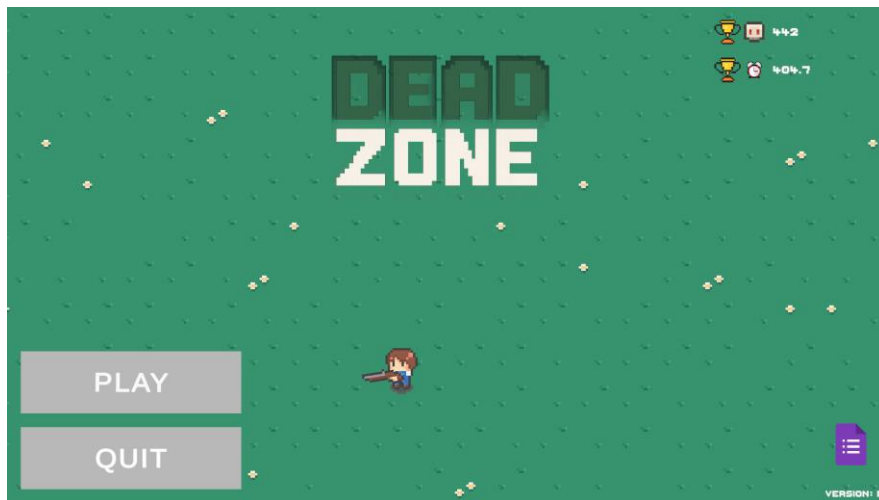


Figura 9 Versione di gioco B

Per quanto riguarda la corrispondente landing page su itch.io, essa adopera deliberatamente un approccio meno curato e professionale. Il sito utilizza una tipografia inconsistente con l'estetica del gioco, una palette cromatica con colori piatti, e una quantità significativa ridotta di screenshot e informazioni descrittive. Anche la miniatura del gioco differisce dalla versione A, utilizzando un'immagine meno accattivante. Questo approccio testa se un design minimalista e meno curato possa comunque attrarre un segmento specifico di giocatori o se invece riduca l'appeal complessivo del prodotto.

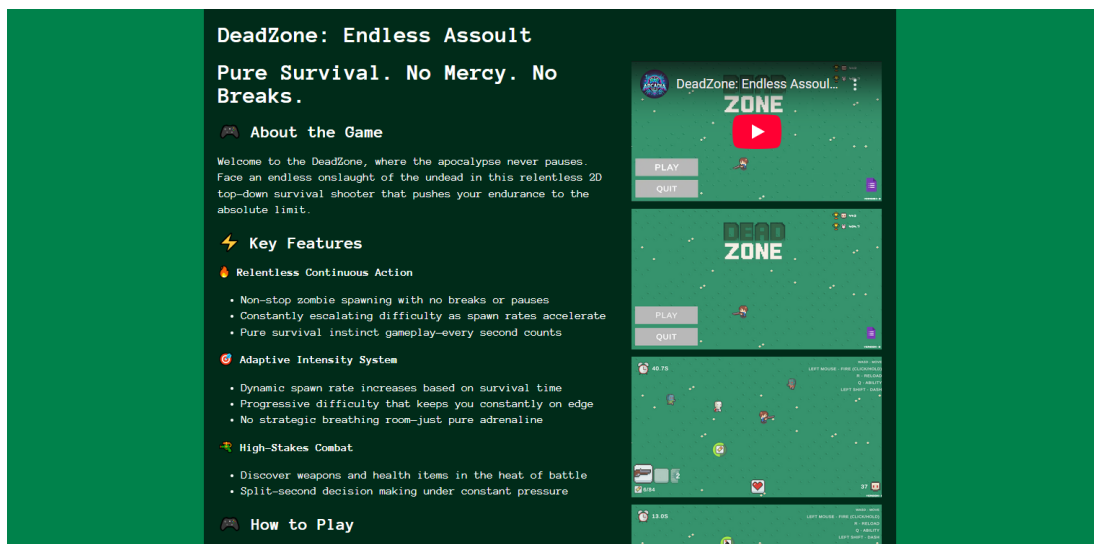


Figura 10 Game Store versione B

La differenziazione sistematica tra le due landing pages permette di analizzare l'impatto della prima impressione fino alla decisione di download, fornendo indicazioni su quanto continuo coerenza grafica e cura della presentazione nella percezione iniziale degli utenti.

5.1.2 ARCHITETTURA TECNICA

Il gioco è stato sviluppato utilizzando Unity 2022.3.52f1 come motore principale, mentre per i disegni pixel art dei personaggi e dei nemici è stato utilizzato Aseprite. L'architettura del codice è stata progettata per essere modulare e facilmente estendibile, rispettando i principi SOLID e permettendo di gestire le due versioni di gioco senza duplicare il lavoro. Il cuore di questo sistema è il VersionManager, uno script che gestisce quale versione del gioco viene caricata. Questo componente utilizza gli Scriptable Object di Unity per memorizzare le configurazioni di ciascuna versione (colori, font, impostazioni audio, effetti particellari, etc.), permettendo di cambiare rapidamente tra le due varianti senza dover ricompilare il progetto. La comunicazione tra i vari componenti avviene tramite l'Observer Pattern mediante un sistema di eventi. Quando un nemico viene ucciso, per esempio, viene lanciato un evento che aggiorna contemporaneamente il contatore delle uccisioni, le statistiche, l'interfaccia utente e i dati di Analytics, senza che questi sistemi debbano conoscersi direttamente. Un altro pattern utilizzato è lo Strategy Pattern, applicato nel codice sulla generazione dei nemici in modo da alternare facilmente le due tipologie di generazione (continuo o ondate).

Per quanto riguarda la raccolta dei dati è stato integrato Unity Analytics, che registra in automatico le azioni dei giocatori: dal tempo di sopravvivenza, ai nemici eliminati, le sessioni di gioco e i vari pattern di movimento. Il sistema rispetta la privacy degli utenti richiedendo il consenso esplicito prima di iniziare qualsiasi tracciamento, in conformità con le normative del GDPR. I dati vengono raccolti solo dopo che il giocatore ha accettato i termini di privacy attraverso un pop up dedicato, che appare all'apertura del gioco per la prima volta.

5.2 METODOLOGIA DI TESTING E RACCOLTA DATI

La metodologia segue due approcci complementari: il tracciamento automatico dei comportamenti di gioco attraverso le Unity Analytics e la raccolta di feedback diretto tramite dei questionari mirati. Il sistema di Analytics registra automaticamente tutte le azioni del giocatore, dal tempo di sopravvivenza al numero di nemici eliminati, fino ai pattern di movimento e alle interazioni con l'interfaccia. Questi dati vengono raccolti solo dopo che il giocatore ha accettato esplicitamente i termini e condizioni per la privacy. Il vero punto di forza della metodologia risiede nel sistema di questionari a quattro livelli che copre l'intera esperienza dell'utente. Ogni versione del gioco ha due questionari dedicati: uno per la sua pagina di itch.io e uno per chi effettivamente gioca. Quando un giocatore chiude il gioco, il sistema apre in automatico il link al questionario corretto in base alla versione giocata. Per la versione A, le domande si focalizzano sulla percezione del sistema a onde: se i giocatori trovano soddisfacenti le pause tra le ondate, se sentono di progredire nel gioco, e se il bilanciamento della difficoltà risulta appropriato. Per la versione B, l'attenzione si sposta sulla gestione del gameplay continuo: se l'assenza di pause risulta stressante o coinvolgente, come i giocatori percepiscono il ritmo costante dell'azione e cosa li motiva a riprovare in un ambiente di sfida continua. Tutti questi questionari utilizzano una combinazione di domande a risposta multipla con scale da 1 a 10 e campi aperti per commenti liberi. Le domande sono formulate in modo neutrale per evitare di influenzare le risposte. Tale raccolta dati, si è estesa per circa due settimane raccogliendo un campione di 71 giocatori. I dati sono stati analizzati sia individualmente che in modo comparativo, cercando pattern comuni e differenze significative tra le esperienze delle due versioni.

6. RISULTATI E ANALISI

L'esperimento è stato condotto dal 27 agosto al 20 Settembre 2025. In questo periodo, della durata di 24 giorni, sono stati raccolti dati sia dalle pagine store di itch.io sia dalle build dei due giochi. La fase di raccolta ha coinvolto due livelli di analisi: le performance delle landing pages (pagine di store) e la performance legata all'esperienza di gioco vera e propria. In ciascuna delle due versioni di gioco, con il relativo sito, sono stati presi in esame vari aspetti, con l'obiettivo di capire quale di essi prevalesse sugli altri. Per validare la significatività statistica delle differenze osservate, verrà utilizzato spesso il test di Kolmogorov-Smirnov a due campioni indipendenti, utilizzando un livello di significatività $\alpha = 0.05$ (5%).

6.1 PERFORMANCE DEI SITI

Per quanto riguarda le pagine store, la versione A ha totalizzato 335 visualizzazioni con 39 downloads, mentre la versione B ha ottenuto 298 visualizzazioni con 32 downloads. I questionari relativi alle pagine store hanno raccolto 17 risposte per la versione A e 14 risposte per la versione B. Tra le due pagine, emerge una differenza marcata nel CTR (click-through rate), 6.91% per la versione A e 6.3% per la versione B. Tali differenze possono essere dovute alle scelte di design e nella cura della presentazione, poiché possono influenzare significativamente la percezione degli utenti e le metriche di engagement. Difatti, l'aspetto più rivelatore emerge dall'analisi qualitativa dei questionari.

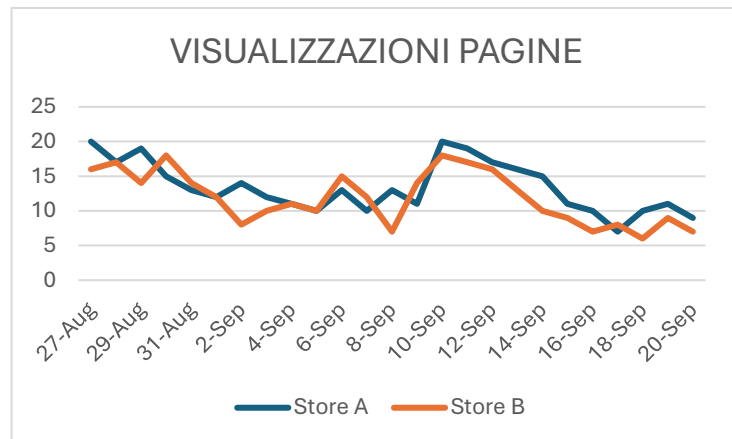


Grafico 1 Visualizzazioni totali delle pagine web delle due build

I colori e il layout rappresentano il primo impatto visivo che determina se un visitatore proseguirà nell'esplorazione. La versione A ha concentrato tutte le sue valutazioni nelle fasce alte di gradimento, creando un'esperienza visiva armoniosa e che comunica qualità e cura. Al contrario, la versione B, con la sua palette più piatta e la disposizione meno curata di elementi, ha raccolto per lo più valutazioni negative, suggerendo una percezione di scarsa qualità. Alcuni utenti hanno descritto la versione B come "troppo monotona", con un

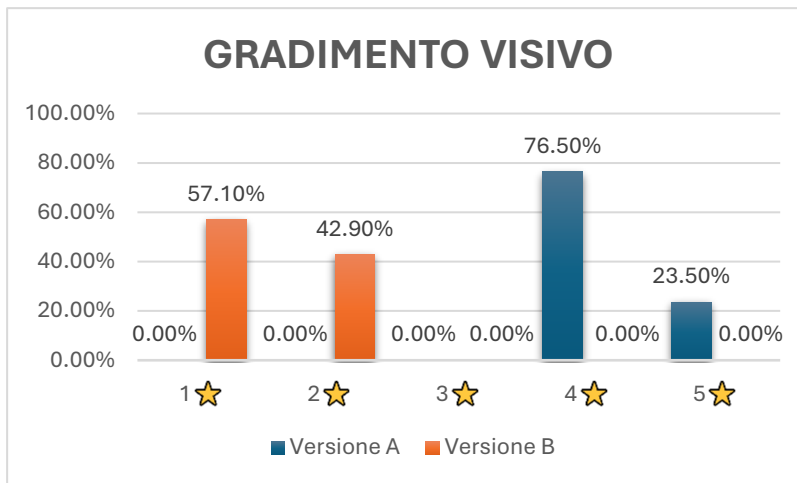


Grafico 2 Gradimento visivo delle due pagine web

visuali coinvolgenti e la scelta di uno sfondo uniforme, possano compromettere l'appeal di una pagina web.

partecipante che ha specificatamente commentato: “Per migliorare il sito, inserirei uno sfondo che non sia solamente verde, ma preferibilmente un'immagine del gioco o che lo rappresenti, altrimenti è troppo monotono.” Questo feedback evidenzia come la mancanza di elementi

Per verificare la significatività di questa divergenza è stato applicato il test di Kolmogorov-Smirnov a due campioni indipendenti. In questo caso, l'ipotesi nulla assume che la percezione visiva della pagina store sia analoga tra le due versioni e che eventuali differenze siano attribuibili al caso. L'ipotesi alternativa, invece, considera che le scelte di design abbiano effettivamente influenzato il gradimento degli utenti, producendo distribuzioni di valutazioni differenti.

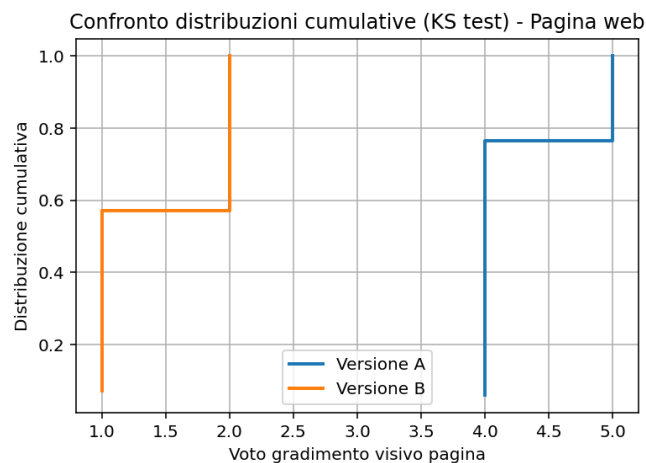


Grafico 3 Grafico distribuzioni cumulative - Grafica Siti

I risultati del test confermano la seconda ipotesi: è stata ottenuta una statistica $KS = 1.0$ e un $p\text{-value} < 0.001$. Questo porta a rigettare l'ipotesi nulla e a concludere che le due distribuzioni sono significativamente diverse, concludendo che la versione A della pagina store è percepita in modo molto più positivo rispetto alla versione B. Anche il font gioca un ruolo importante nella percezione di qualità visiva della pagina, difatti, la versione A, che adotta un font coerente con l'estetica pixel art del gioco, ha ottenuto un consenso unanime da parte degli utenti, dimostrando come la coerenza stilistica tra contenuto e presentazione sia percepita come segno di professionalità e attenzione ai dettagli. La versione B, invece, utilizzando una tipografia scollegata dallo stile del gioco, ha portato a che più di 2/3 degli utenti esprimessero un giudizio negativo.

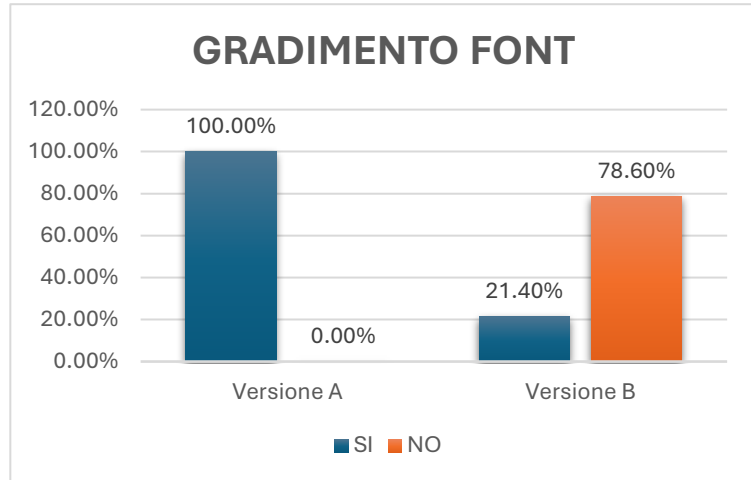


Grafico 4 Gradimento font siti web

Il confronto diretto tra le due versioni evidenzia chiaramente la preferenza degli utenti: tra chi ha esplorato entrambe le pagine (circa il 39% del campione totale), la scelta è stata inequivocabilmente quella della versione A.

VISITA PAGINE

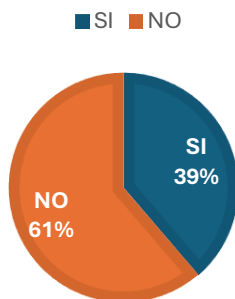


Grafico 5 Percentuale di utenti che hanno visitato anche l'altro sito

PREFERENZA VERSIONE

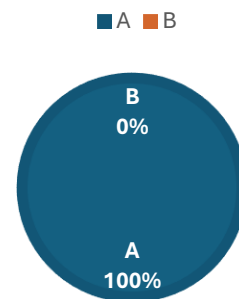


Grafico 6 Preferenza degli utenti tra le due versioni visitate

Complessivamente i risultati della versione A vedono una netta maggioranza di voti positivi in fascia alta (tra le 8-10 stelle), a differenza della versione B la cui maggior parte degli utenti ricadono tra le 4 e le 5 stelle complessive.

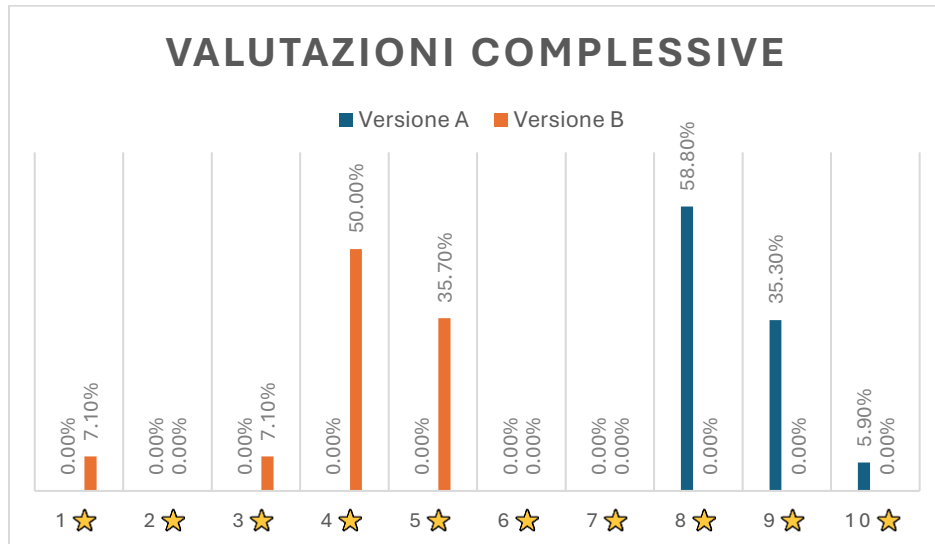


Grafico 7 Valutazioni complessive dei siti web

Anche qui per validare statisticamente questa evidente divergenza nelle valutazioni complessive, è stato utilizzato il test di Kolmogorov-Smirnov a due campioni indipendenti. L'ipotesi nulla assume che le due valutazioni delle pagine web siano equivalenti e che le differenze osservate siano attribuibili alla casualità statistica. L'Ipotesi alternativa, al contrario, sostiene che le diverse scelte di design hanno effettivamente influenzato la percezione complessiva degli utenti, generando distribuzioni di giudizi significativamente differenti.

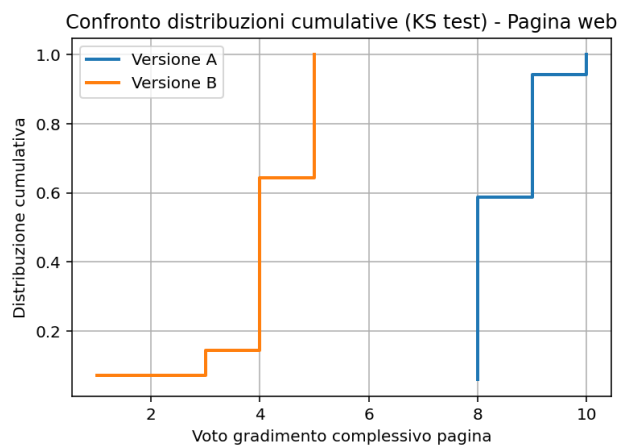


Grafico 8 Grafico distribuzioni cumulative - voti totali pagine store

Il test ha restituito dei risultati inequivocabili: una statistica KS pari a 1.0 e un p-value ampiamente al di sotto della soglia di significatività (< 0.001). Questi risultati portano a rigettare l'ipotesi nulla, confermando che le due distribuzioni sono statisticamente e sostanzialmente diverse, dimostrando come la combinazione degli elementi curati, dalla coerenza tipografica alla ricchezza visiva, dalla professionalità del layout alla scelta della palette dei colori, contribuiscono a mantenere l'interesse dell'utente traducendosi in metriche di performance superiori.

6.2 PERFORMANCE DI GIOCO

L'analisi del game design rivela dinamiche che evidenziano come diversi fattori posso influenzare sia la percezione qualitativa sia i comportamenti dei giocatori. La versione A, con i suoi 39 downloads, ha ottenuto 14 risposte al questionario, mentre la versione B, con i suoi 32 downloads, ne ha raccolte 11.

6.2.1 GRAFICA

Partiamo analizzando i vari dati raccolti dai Google Forms, in particolare la grafica, la quale rappresenta uno degli aspetti più impattanti tra le due versioni. La versione A, con il suo rating grafico medio di 8.29/10 stelle, possiede effetti particellari visivamente più attraenti

che arricchiscono l'esperienza visiva, concentrando la maggior parte delle valutazioni tra le 4 e le 5 stelle. La versione B, al contrario, con un rating grafico medio di 5.0/10 stelle, mostra una distribuzione più frammentata delle valutazioni, accumulandone la maggior parte nelle fasce intermedie e basse.

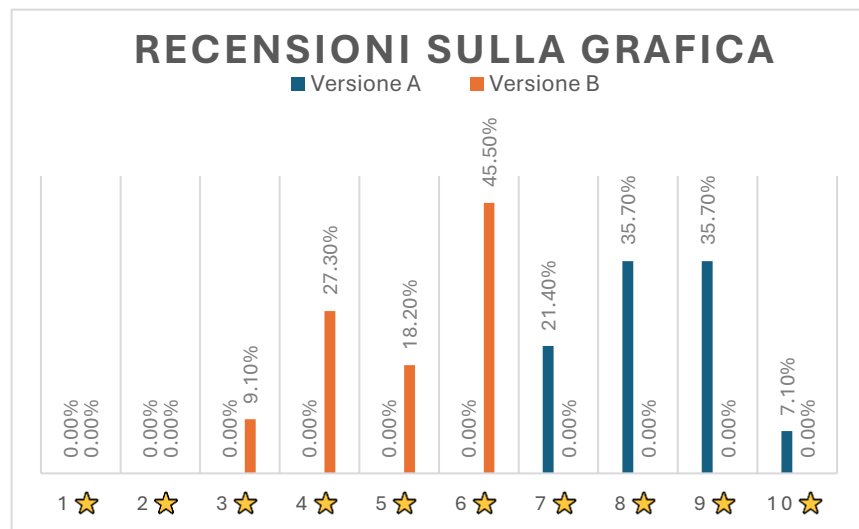


Grafico 9 Recensioni grafica delle due build

Per verificare se le differenze osservate fossero statisticamente significative è stato applicato il test di Kolmogorov-Smirnov a due campioni indipendenti. È stata definita l'ipotesi nulla per cui si suppone che non esistano differenze significative tra le due versioni, al contrario per l'ipotesi alternativa secondo la quale invece esistono.

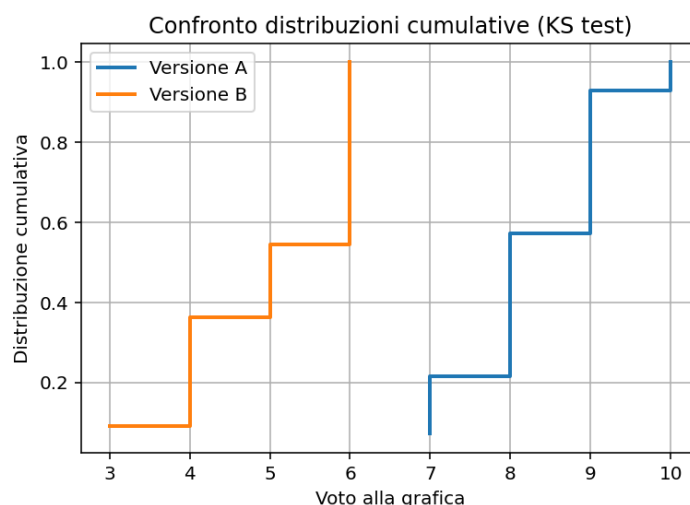


Grafico 10 Grafico distribuzioni cumulative - Grafiche gioco

Il grafico delle distribuzioni cumulative mostra chiaramente come le due curve non si sovrappongono: la Versione A presenta una distribuzione sistematicamente spostata verso i punteggi alti (7-10), mentre la versione B si concentra sui valori bassi-medi (3-6). Il test di Kolmogorov-Smirnov conferma questa differenza, restituendo un p-value < 0.001 e una statistica KS pari a 1.0, che consente di rigettare l'ipotesi nulla di uguaglianza delle distribuzioni. Possiamo dunque concludere che la grafica della versione A è percepita nettamente meglio rispetto alla versione B, non solo in termini di media dei voti, ma anche nella forma complessiva delle distribuzioni di giudizio fornite dagli utenti.

6.2.2 TUTORIAL

L'analisi della fase di apprendimento iniziale mette in evidenza due approcci diversi al design dell'onboarding. Nella versione A il giocatore viene guidato da un tutorial integrato, che semplifica la comprensione dei comandi e fornisce un supporto strutturato fin dai primi momenti di gioco. Questo ha portato a un 64.30% degli utenti che valuta l'esperienza con il massimo punteggio (5 stelle). Di contro, la versione B propone un modello basato sul "learning by doing", cioè lascia che il giocatore impari sperimentando direttamente. In questo caso le valutazioni sembrano distribuite in modo molto più equilibrato, suggerendo

che l'assenza di tutorial non per forza rappresenta un ostacolo per tutti. Al contrario, per alcuni giocatori può persino costituire un elemento positivo, in quanto favorisce un'esperienza più esplorativa e autonoma.

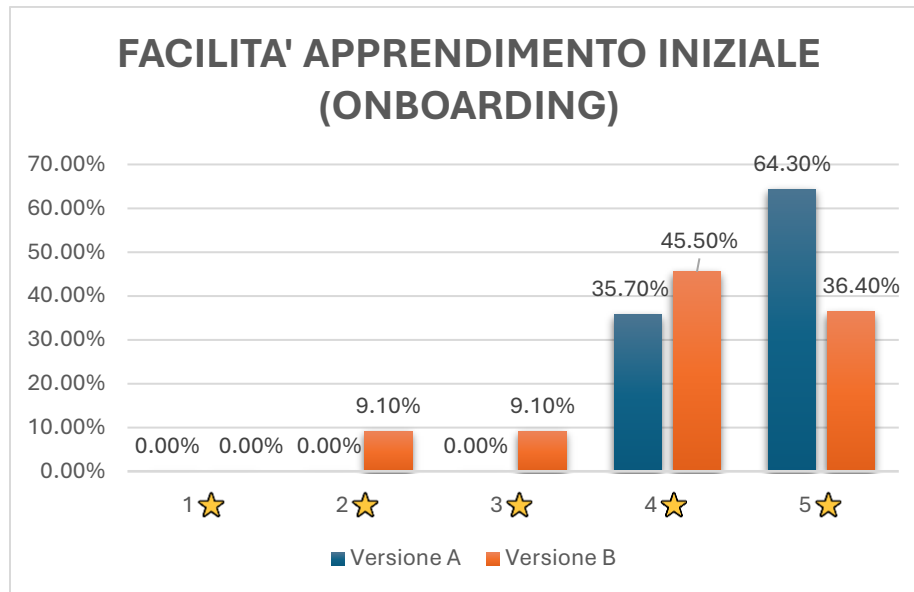


Grafico 11 Facilità apprendimento iniziale dei comandi

Anche in questo contesto, è stato applicato Kolmogorov-Smirnov per verificare la significatività statistica. L'ipotesi nulla assume che l'esperienza di apprendimento iniziale venga percepita nello stesso modo indipendentemente dalla presenza o meno del tutorial, e quindi che le due distribuzioni di voti risultino sovrapponibili. L'ipotesi alternativa, al contrario, sostiene che la presenza del tutorial influenza effettivamente la percezione dei giocatori, producendo una distribuzione dei voti diversa rispetto a quella della versione basata esclusivamente sull'apprendimento tramite sperimentazione.

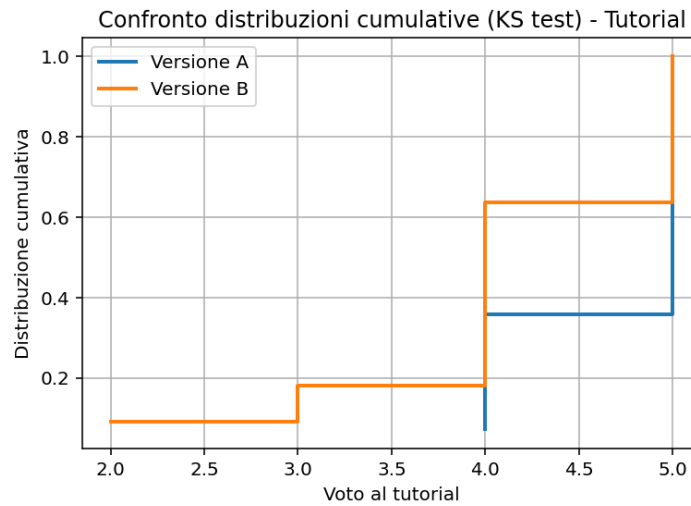


Grafico 12 Grafico distribuzioni cumulative - Tutorial

Dal grafico delle distribuzioni cumulative si può osservare chiaramente come la versione A si concentri esclusivamente sui voti alti (4-5), mentre la versione B include anche valutazioni più basse. Tuttavia, il test ha restituito una statistica KS pari a 0.279 e un p-value = 0.619, superiore alla soglia $\alpha = 0.05$. Questo porta a non rigettare l'ipotesi nulla, indicando che, dal punto di vista statistico, non è presente un'evidenza sufficiente per affermare che le due distribuzioni siano significativamente diverse. In sintesi, sebbene i dati ottenuti dal sondaggio mostrino una tendenza che suggerisce una percezione più stabile e positiva con il tutorial integrato (versione A), tale differenza non raggiunge la significatività statistica, probabilmente anche a causa della dimensione ridotta del campione.

6.2.3 UI E FONT

I punteggi ottenuti dai sondaggi sulla componente visiva dell'interfaccia utente (UI) mettono in evidenza differenze nette tra le due versioni di gioco. La versione A ha raccolto giudizi fortemente positivi per cui il 71.40% degli utenti assegna 4 stelle e il restante 28.60% assegna il massimo punteggio di 5 stelle. La versione B, invece, ottiene valutazioni più basse in cui il 36.40% degli utenti ha recensito con 1 stella e il restante 63.60% con 2 stelle, senza alcun voto superiore. Anche per quanto riguarda il font utilizzato, si può notare un gradimento unanime per la versione A, al contrario di quello usato nella versione B, che è stato valutato in modo nettamente negativo.

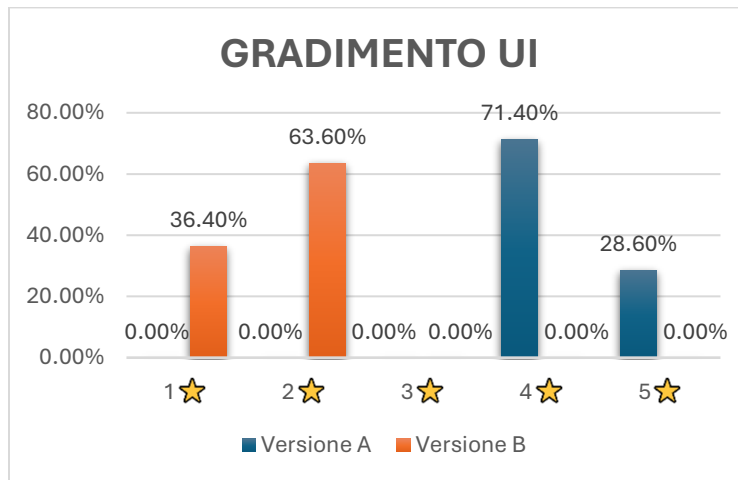


Grafico 13 Recensioni sulla UI delle versioni di gioco

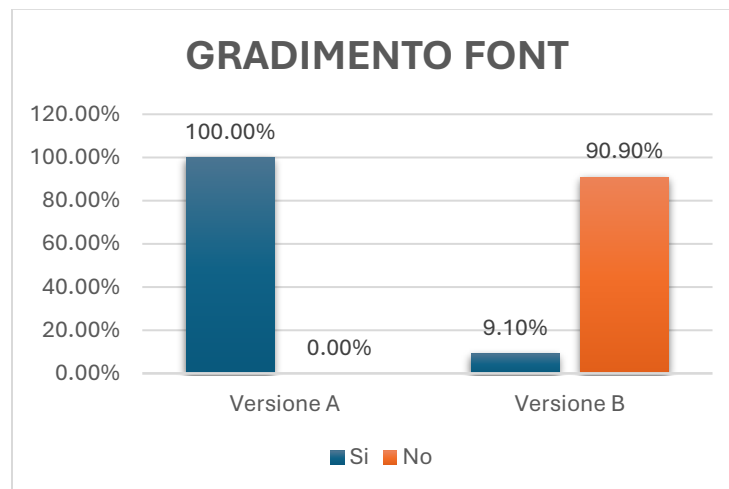


Grafico 14 Gradimento stile tipografico

Applicando il test di Kolmogorov-Smirnov a due campioni indipendenti, è stato cercato di verificare se tale dipendenza fosse statisticamente significativa. In questo contesto, l'ipotesi nulla presuppone che il gradimento della UI non cambi al variare della versione, per cui le distribuzioni dei voti dovrebbero risultare simili. L'ipotesi alternativa, invece, considera che le due versioni del gioco abbiano influenzato in modo diverso la percezione degli utenti, producendo distribuzioni di giudizi tra di loro divergenti.

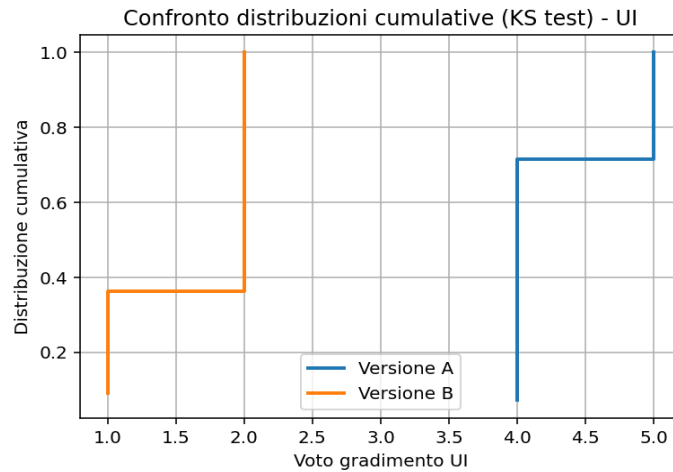


Grafico 15 Grafico distribuzioni cumulative - Gradimento UI

I risultati confermano in maniera equivocabile questa seconda ipotesi: il test ha restituito una statistica KS pari a 1.0 e un p-value < 0.001, valore bel al di sotto della soglia di 0.05. Tali risultati portano quindi a rigettare l'ipotesi nulla e a concludere che le due distribuzioni sono significativamente diverse. In altre parole, il gradimento della UI risulta fortemente a favore della versione A.

6.2.4 RETENTION ED ENGAGEMENT

Dal punto di vista delle metriche di engagement, la versione A, ha ottenuto un tempo medio di sessione superiore rispetto a B (14.5 minuti contro i 9 minuti) causato, possibilmente, da tempi di attesa tra le varie ondate maggiori e da una curva di difficoltà più graduale.

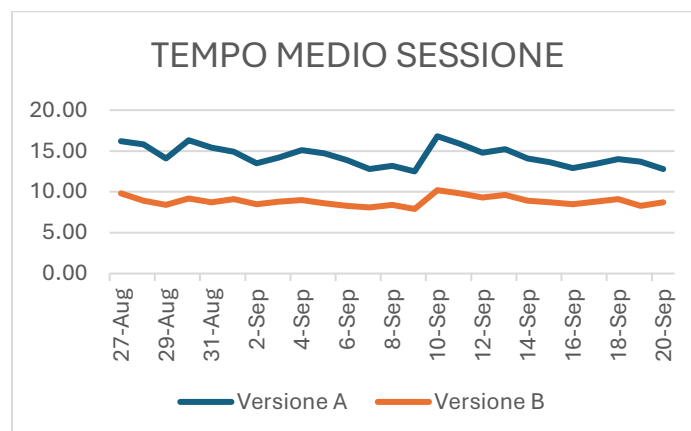


Grafico 16 Tempi medi di gioco delle due versioni

Un aspetto interessante emerge anche dall'analisi della retention a lungo termine. Sebbene la versione B abbia un punteggio medio complessivo del gioco più basso rispetto alla

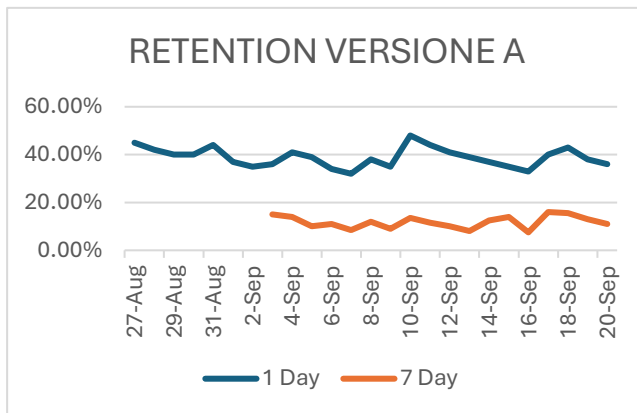


Grafico 17 Game Retention versione A

versione A (6.18 contro i 7.64), essa ottiene una retention leggermente superiore dopo 7 giorni (13% contro 12%). Questo dato, apparentemente controintuitivo, suggerisce che il gameplay con la generazione di nemici in modo continuativo, pur essendo meno curato dal punto di vista estetico e dell'esperienza utente, riesca comunque a stimolare un coinvolgimento più duraturo. Ciò potrebbe dipendere dalla natura stessa del loop del gioco continuo,

che non offre interruzioni e rende quindi più difficile per il giocatore interrompere la sessione.

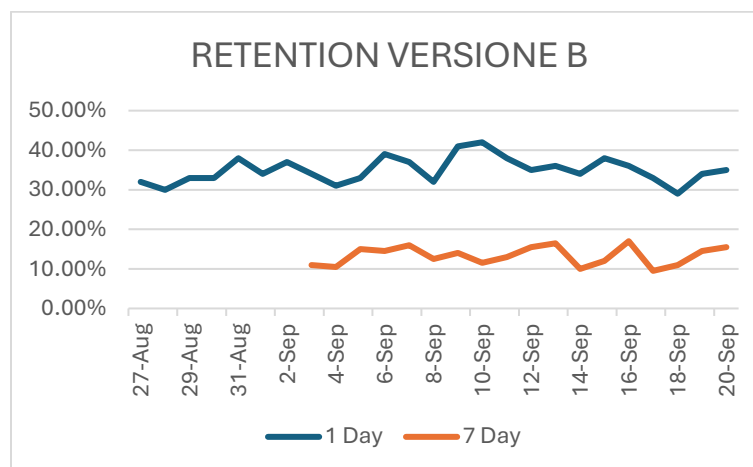


Grafico 18 Game Retention versione B

6.3 VALIDAZIONE STATISTICA DEI RISULTATI

Un aspetto fondamentale di questo esperimento è stato quello di verificare se le differenze osservate tra le due versioni fossero statisticamente significative e non attribuibili al caso. Per fare ciò, è stato applicato spesso il test di Kolmogorov-Smirnov a due campioni indipendenti, che ha permesso di stabilire, con un margine di confidenza prestabilito, se rigettare o meno l'ipotesi nulla. È stato adottato un livello di significatività $\alpha = 0.05$,

corrispondente ad una probabilità del 5%, di commettere un errore. Il principio alla base del test prevede due ipotesi contrapposte: l'ipotesi nulla (H_0) assume che le due distribuzioni siano identiche e che eventuali differenze osservate siano dovute al caso, mentre, l'ipotesi alternativa (H_1), sostiene invece che le distribuzioni sono effettivamente diverse. Per decidere quale ipotesi accettare, si confronta il valore del p-value ottenuto con il livello di significatività prestabilito e, se il valore è al di sotto del livello, si rigetta l'ipotesi nulla concludendo che tali differenze sono statisticamente significative, ovvero che c'è meno del 5% di probabilità che siano frutto del caso.

La scelta del test di Kolmogorov-Smirnov è motivata da diverse caratteristiche che lo rendono adatto ai dati raccolti in questo studio. Innanzitutto, i dati provenienti dai questionari sono di tipo ordinale, ossia valutazioni su scale (da 1 a 5 o da 1 a 10 stelle), e il test KS non richiede che questi dati seguano una distribuzione normale, rendendolo ideale per lavorare con feedback soggettivi dei giocatori. In secondo luogo, la ridotta dimensione del campione (11-17 risposte per versione) richiede un test statistico che mantenga una buona affidabilità anche con pochi dati, caratteristica di Kolmogorov-Smirnov. Un altro vantaggio di questo test è che confronta tutti i dati raccolti e non solo le medie. Nel game design è importante, infatti, vedere come sono distribuite le opinioni dei giocatori, e non limitarsi a dire quale versione ha preso il voto medio più alto. Infine, i due gruppi di utenti sono completamente indipendenti (ogni utente ha valutato solo una versione), condizione ideale per applicare il test KS a due campioni indipendenti.

Il test è stato applicato a tutti gli aspetti principali dell'esperienza utente: la percezione visiva delle pagine store, la qualità grafica percepita nel gioco, il gradimento dell'interfaccia utente (UI), la facilità di apprendimento iniziale e le valutazioni complessive. Per ogni confronto sono state formulate delle ipotesi specifiche relative alla variabile presa in esame.

Variabile testata	Statistica KS	p-value	Significativo ($\alpha = 0.05$)?	Interpretazione
Gradimento visivo pagine store	1.0	< 0.001	Si	La pagina store della versione A è percepita nettamente superiore dal punto di vista estetico

Valutazione complessiva pagine store	1.0	< 0.001	Si	La versione A risulta preferita in modo significativo
Qualità grafica di gioco	1.0	< 0.001	Si	Gli effetti particellari e la cura visiva della versione A sono apprezzati
Gradimento interfaccia utente (UI) e font	1.0	< 0.001	Si	La coerenza stilistica della versione A genera maggiore soddisfazione
Facilità di apprendimento iniziale (tutorial)	0.279	0.619	No	La differenza non raggiunge significatività statistica

Tabella 1 Sintesi dei risultati del test di Kolmogorov-Smirnov applicato alle diverse variabili sperimentali

Come si può osservare dalla tabella, su cinque confronti effettuati, quattro hanno mostrato differenze statisticamente significative ($p < 0.001$) a favore della versione A, con valori di statistica KS pari a 1.0 che indicano una separazione completa tra le due distribuzioni. Questo risultato, suggerisce che le scelte di design implementate nella versione A (palette cromatica, font coerente con lo stile pixel art, effetti particellati maggiori e curati e l'interfaccia professionale) hanno avuto un impatto sostanziale e consistente sulla percezione degli utenti. L'unica variabile che non ha raggiunto la significatività statistica è stata la facilità di apprendimento iniziale, relativa alla presenza o meno del tutorial. Con una statistica KS di 0.279 e un p-value di 0.619, non è possibile concludere con sufficiente confidenza che il tutorial integrato della versione A offra un vantaggio percepibile rispetto all'approccio "learning by doing" della versione B. Tale risultato può essere interpretato in due modi: da un lato potrebbe indicare che per un gioco con meccaniche relativamente semplici, come questo, l'apprendimento in gioco sia altrettanto efficace di un tutorial guidato, dall'altro, la mancata significatività statistica potrebbe essere dovuta alla dimensione limitata del campione (14 risposte per la versione A, 11 risposte per la versione B), che non fornisce sufficiente potenza statistica per rilevare le differenze più sottili. Inoltre, è importante notare che tutti gli altri test effettuati hanno prodotto un p-value estremamente

basso (< 0.001), indicando che la probabilità che le differenze osservate siano casuali è inferiore allo 0.1%. Questo livello di confidenza, considerato molto robusto, può farci concludere che le scelte di design della versione A hanno effettivamente migliorato l'esperienza utente in modo misurabile, suggerendo che l'attenzione ai dettagli e la cura nella presentazione non sono aspetti secondari, ma elementi fondamentali che influenzano profondamente la percezione di qualità di un prodotto videoludico.

7. CONCLUSIONI E POSSIBILI SVILUPPI FUTURI

Il lavoro presentato ha dimostrato come l'A/B testing sia una metodologia efficace per ottimizzare l'esperienza dell'utente nei videogiochi, confermandone il ruolo centrale nello sviluppo moderno. L'esperimento ha mostrato chiaramente come le scelte di design influenzino l'esperienza di gioco, anche nei dettagli apparentemente minori, come il font, gli effetti sonori e la cura grafica. L'impiego di test statistici, in particolare Kolmogorov-Smirnov, ha permesso di verificare in maniera rigorosa le differenze tra le due versioni sperimentali. I risultati principali possono essere così sintetizzati:

- La versione A ha ricevuto valutazioni significativamente più alte in termini di qualità grafica e cura complessiva ($p < 0.001$), risultando percepita più professionale e rifinita.
- La versione B, nonostante recensioni mediamente inferiori, ha avuto una retention a 7 giorni leggermente superiore (13% vs 12 %)

I risultati suggeriscono che si potrebbe migliorare la versione A integrando alcuni elementi vincenti della versione B attraverso ulteriori cicli di A/B testing. Per esempio, si potrebbero introdurre delle modalità temporanee chiamate "Wave Rush" grazie alla quale è possibile unire la qualità percepita dalla versione A con l'intensità della versione B. Anche i sistemi di unlock e achievement potrebbero rappresentare un'altra opportunità di miglioramento della retention. L'integrazione di un sistema economico nel gioco, come delle monete ottenibili all'eliminazione dei nemici, permetterebbe di testare diverse strategie di monetizzazione interna e progressione del giocatore. Questo approccio seguirebbe il modello già validato da Supercell nel loro titolo Clash of Clans. L'A/B testing su questo sistema economico aprirebbe nuove dimensioni sperimentali: si potrebbero confrontare diverse configurazioni di drop rate delle monete, testare l'impatto di diversi prezzi per le abilità acquistabili e sperimentare con bundle opzionali per ciascuna versione.

Naturalmente lo studio presenta anche delle limitazioni, che aprono la strada a sviluppo futuri. Il campione relativamente ridotto suggerisce la necessità di test su una scala molto più ampia che però, richiede del tempo. Allo stesso modo, anche la durata limitata (24 giorni) non ha permesso di analizzare la retention a lungo termine oltre i 30 giorni. Un miglioramento metodologico fondamentale per futuri esperimenti consiste nell'implementazione di un sistema di randomizzazione dinamica che garantisca l'equilibrio

tra i due gruppi sperimentali. Attualmente, la distribuzione dei partecipanti tra le due versioni avviene in modo non casuale, ma l'integrazione di un algoritmo di bilanciamento automatico potrebbe ottimizzare significativamente la validità statistica dei risultati.

L'A/B testing non è soltanto uno strumento di ottimizzazione, ma un vero e proprio ponte tra l'arte del game design e la scienza del comportamento umano. Grazie a questa metodologia, le idee possono trasformarsi in ipotesi concrete da verificare, permettendo di confermare o smentire le scelte di design con dati reali. In questo modo si riducono i rischi legati a decisioni basate solo sulle opinioni e si velocizzano i cicli di sviluppo.

8. RINGRAZIAMENTI

Arrivato alla fine di questo percorso, è giunto il momento di ringraziare chi ha contribuito a renderlo possibile.

Un pensiero speciale va alla mia famiglia, che ha creduto in me fin dall'inizio e non ha mai smesso di sostenermi. A mamma e papà, che mi hanno dato la libertà di scegliere la mia strada e che, con pazienza e fiducia, mi hanno accompagnato in ogni fase di questo cammino. A mia sorella, per tutte le volte in cui bussava alla mia porta e si sentiva rispondere “non posso, devo studiare”, senza mai perdere la pazienza.

Desidero anche ringraziare il Prof. Marco Rocchetti per aver accolto la mia proposta di tesi e per avermi guidato con competenza e attenzione in questo lavoro di ricerca sull'A/B testing nei videogiochi.

Un ringraziamento affettuoso ai miei amici di sempre e i miei compagni di corso, con i quali ho condiviso gioie, fatiche e ricordi che resteranno indelebili. Ognuno di voi ha lasciato qualcosa di speciale in questo viaggio e, anche se non posso nominarvi tutti (altrimenti questa sezione diventerebbe lunga quanto la tesi stessa), sappiate che ogni momento trascorso insieme è stato prezioso.

Infine, come non posso non ringraziare Pandino, la piccola mascotte di famiglia, che con la sua presenza rende più felici le nostre giornate.

A tutti voi, grazie di cuore.

9. BIBLIOGRAFIA

- “Il mercato dei videogames”. DigitalBros.com. *Ultimo accesso 6 agosto 2025*. <https://digitalbros.com/it/gruppo/il-mercato-dei-videogiochi/>
- “A/B Testing: funzionalità, esempi, pro e contro”. Qualtrics.com. *Ultimo accesso 6 agosto 2025*. <https://www.qualtrics.com/it/experience-management/ricerca/ab-testing/>
- “A/B Testing”. Wikipedia.org. *Ultimo accesso 7 settembre 2025*. https://en.wikipedia.org/wiki/A/B_testing
- “So what is A/B Testing?”. Mcmillanphillips.com. *Ultimo accesso 7 settembre 2025*. <https://mcmillanphillips.com/blog/a-brief-history-on-a-b-testing-and-why-you-need-it-today/>
- “Cos’è l’A/B test?”. Movylo.com. *Ultimo accesso 7 settembre 2025*. <https://www.movylo.com/it/cose-la-b-test/>
- “The A/B Test: inside the technology that’s changing the rules of business”. Wired.com. *Ultimo accesso 7 settembre 2025*. <https://www.wired.com/2012/04/ff-abtesting/>
- “The Google 41 shades of blue test”. Insight4print.ceo. *Ultimo accesso 7 settembre 2025*. <https://www.insights4print.ceo/2022/02/the-google-41-shades-of-blue-test-is-fundamentally-flawed/>
- “Zynga and the launch of Farmville”. Stanford.edu. *Ultimo accesso 7 settembre 2025*. <https://www.gsb.stanford.edu/faculty-research/case-studies/zynga-launch-farmville>
- “Clash Royale Battle rewards update: Testing Chests, Gold & More”. Fragster.com. *Ultimo accesso 30 agosto 2025*. <https://www.fragster.com/clash-royale-battle-rewards-update-testing-chests-gold-more/>
- “Che cos’è il test A/B?”. Oracle.com. *Ultimo accesso 7 agosto 2025*. <https://www.oracle.com/it/cx/marketing/what-is-ab-testing/>
- “Glossario: Test A/B e Test Multivariati”. Tagmanageritalia.it. *Ultimo accesso 31 agosto 2025*. <https://www.tagmanageritalia.it/glossario-test-ab-test-multivariati/>
- “A/B/n Testing”. Optimizely.com. *Ultimo accesso 31 agosto 2025*. <https://www.optimizely.com/optimization-glossary/abn-testing/>
- “What is A/B/n Testing?”. Harness.io. *Ultimo accesso 31 agosto 2025*. <https://www.harness.io/harness-devops-academy/a-b-n-testing>
- “Multivariate vs A/B Testing: Which is right for you?”. Statsig.com. *Ultimo accesso 31 agosto 2025*. <https://www.statsig.com/perspectives/multivariate-vs-ab-testing-which-is-right-for-you>

- “What is A/B Testing in mobile games?”. Medium.com. *Ultimo accesso 31 agosto 2025*. <https://medium.com/@alidrsn/what-is-a-b-testing-in-mobile-games-af9338170ff7>
- “A/B Testing: How to optimize in-app ad performance for mobile games”. Chartboost.com. *Ultimo accesso 31 agosto 2025*. <https://www.chartboost.com/resources/guides/a-b-testing-how-to-optimize-in-app-ad-performance-for-mobile-games-chartboost-academy/>
- “Playtest day”. Ubisoft.com. *Ultimo accesso 30 agosto 2025*. <https://paris.ubisoft.com/en/blog/our-jobs-en/playtest-day>
- “UX Design nei videogiochi: cos’è e perché è importante”. Csainformazione.it. *Ultimo accesso 30 agosto 2025*. <https://csainformazione.it/ux-design-videogiochi/>
- “Brains, UX & Games”. Celiahodent.com. *Ultimo accesso 30 agosto 2025*. <https://celiahodent.com/cognitive-psychology-applied-to-user-experience-in-video-games/>
- “Self-determination theory”. Wikipedia.org. *Ultimo accesso 30 agosto 2025*. https://en.wikipedia.org/wiki/Self-determination_theory
- “WASD controls are on the way”. Leagueoflegends.com. *Ultimo accesso 31 agosto 2025*. <https://www.leagueoflegends.com/en-us/news/dev/dev-wasd-controls-are-on-the-way/>
- “1,7 and 30-day retention rates are key”. Pocketgamer.biz. *Ultimo accesso 31 agosto 2025*. <https://www.pocketgamer.biz/1-7-and-30-day-retention-rates-are-key-says-tablet-first-supercell/>
- “Game monetization design: Analysis of Clash of Clans”. Gamedeveloper.com. *Ultimo accesso 31 agosto 2025*. <https://www.gamedeveloper.com/business/game-monetization-design-analysis-of-clash-of-clans>
- “UX Research for Product vs Games”. Odettejansen.com. *Ultimo accesso 1 settembre 2025*. <https://odettejansen.medium.com/ux-research-for-products-vs-games-a-comparative-exploration-7264c98a8e6>
- “User Interface success and failures through a video game: Apex Legends”. Medium.com. *Ultimo accesso 1 settembre 2025*. <https://william-pinaud.medium.com/a-case-study-in-user-experience-user-interface-success-and-failures-through-a-video-game-apex-4d37c7f08a4b>
- “Gaming companies run thousands of experiments a year”. Geteppo.com. *Ultimo accesso 1 settembre 2025*. <https://www.geteppo.com/blog/gaming-companies-run-thousands-of-experiments-a-year>
- “Self-determination theory for gaming experiences”. Medium.com. *Ultimo accesso 1 settembre 2025*. <https://medium.com/@lupigd/self-determination-theory-for-gaming-experiences-f8df37c031ec>

- “Mid-Core success Part 1: Core Loops”. Gameanalytics.com. *Ultimo accesso 31 agosto 2025*. <https://www.gameanalytics.com/blog/mid-core-success-part-1-core-loops>
- “How King defines a good Candy Crush Saga level”. Mobilegamer.biz. *Ultimo accesso 31 agosto 2025*. <https://mobilegamer.biz/how-king-defines-a-good-candy-crush-saga-level-and-why-it-constantly-prunes-the-bad-ones/>
- “5 tips for mobile games a/b testing”. Crazylabs.com. *Ultimo accesso 2 settembre 2025*. <https://www.crazylabs.com/blog/5-tips-for-mobile-games-a-b-testing/>
- “The most Important KPIs for Monitoring Mobile Games”. Metricfire.com. *Ultimo accesso 2 settembre 2025*. <https://www.metricfire.com/blog/the-most-important-kpis-for-monitoring-mobile-games/>
- “Effective Strategies to Boost in App Purchases in Mobile Games”. Segwise.ai. *Ultimo accesso 2 settembre 2025*. <https://segwise.ai/blog/effective-strategies-to-boost-in-app-purchases-in-mobile-games-clxbr489d0010d2agh26cvdy0>
- “Best Practices for Pricing and Bundling”. Segwise.ai. *Ultimo accesso 2 settembre 2025*. <https://segwise.ai/blog/best-practices-for-pricing-and-bundling-in-mobile-games-clxo9g5vs001cagmd5b99bm71>
- “Developing UX Practices at Epic Games”. Celiahodent.com. *Ultimo accesso 3 settembre 2025*. <https://celiahodent.com/ux-practices-epic-games/>
- “Understanding the success of Fornite: A UX and psychology perspective”. Medium.com. *Ultimo accesso 3 settembre 2025*. <https://medium.com/ironsource-levelup/understanding-the-success-of-fornite-a-ux-user-experience-psychology-perspective-208cee587dc6>
- “Iterative testing allows Nitro Games to increase their tutorial completion rate to 90%”. PlaytestCloud.com. *Ultimo accesso 3 settembre 2025*. <https://start.playtestcloud.com/blog/nitro-games-case-study>
- “First-Time-User Experience (FTUE) in Mobile Games”. Udonis.co. *Ultimo accesso 3 settembre 2025*. <https://www.blog.udonis.co/mobile-marketing/mobile-games/first-time-user-experience>
- “Game audio – an investigation into the effect of audio on player immersion”. Springer.com. *Ultimo accesso 3 settembre 2025*. <https://link.springer.com/article/10.1007/BF03392342>
- “Interview with Zynga’s Mark Pincus”. Medium.com. *Ultimo accesso 4 settembre 2025*. <https://medium.com/the-lindberg-interviews/interview-with-zyngas-mark-pincus-f35b8a36f76>
- “Why does my game look different from someone else’s?”. Supercell.com. *Ultimo accesso 5 settembre 2025*. <https://support.supercell.com/clash-royale/en/articles/why-does-my-game-look-different-from-someone-elses.html>

- “How King Games keeps you crushing candy for years”. Campaignlive.com. *Ultimo accesso 5 settembre 2025*. <https://www.campaignlive.co.uk/article/king-games-keeps-crushing-candy-years/1440143>
- “Riot Games Data Analyst Interview”. Interviewquery.com. *Ultimo accesso 5 settembre 2025*. <https://www.interviewquery.com/interview-guides/riot-games-data-analyst>
- “Balancing Data and Design”. Gamedeveloper.com. *Ultimo accesso 3 settembre 2025*. <https://www.gamedeveloper.com/business/balancing-data-and-design>
- “You can’t A/B test your way to Greatness”. Jefago.com. *Ultimo accesso 8 settembre 2025*. <https://www.jefago.com/product-management/ab-testing-does-not-create-great-products/>
- “The power of A/B testing in Game Development: Enhancing Player Experience”. Meliorgames.com. *Ultimo accesso 10 settembre 2025*. <https://meliorgames.com/game-development/the-power-of-a-b-testing-in-game-development-enhancing-player-experience/>
- “A/B Testing in Mobile Games”. Superscale.com. *Ultimo accesso 11 settembre 2025*. <https://superscale.com/a-b-testing-in-mobile-games/>

