



HEURISTISK ANVÄNDBARHETSUTVÄRDERING

Komparativ undersökning av 10 Usability
Heuristics och Net Heuristic List

HEURISTIC USABILITY EVALUATION

Comparative investigation of 10 Usability
Heuristics and Net Heuristic List

Examensarbete inom huvudområdet
Informationsteknologi
Grundnivå 15 högskolepoäng
Vårtermin 2026

Gabriel Malm

Handledare: Martin Hagvall
Examinator: Jana Rambusch

Sammanfattning

Spel innebär ofta en konstant interaktion mellan spel och spelare. Ett sätt att utvärdera hur väl detta fungerar kallas heuristisk användbarhetsutvärdering (HU). HU kombinerar traditionell användbarhetsanalys med fokus på spelupplevelsen. Två framstående listor som vardera innehåller 10 punkter att utvärdera är Nielsens 10 Usability Heuristics (Nielsens 10) och jämförs med Net Heuristics List (NHL). Den första är framtagen för all sorts programvara, den andra är specifikt utvecklad för spel. I denna studie undersöks vilken lista som bäst stödjer spelutvecklare som saknar uttalad GUX-kompetens när de vill utvärdera sina spel inom användbarhet. Genom ett test av ett befintligt spel och en enkät riktad till målgruppen jämförs listorna. Resultaten från båda metoderna indikerar slutsatsen att NHL upplevs som den mest användbara listan, även om vissa avvikelser och intressanta observationer framkommer.

Nyckelord: heuristik, användbarhetsutvärdering, spelutvecklare

Innehållsförteckning

1	Introduktion	1
2	Bakgrund	2
2.1	Användbarhet i spel	2
2.2	Heuristisk användbarhetsutvärdering	3
3	Problemformulering	5
3.1	Metodbeskrivning	5
3.2	Förväntat resultat	7
3.3	Forskningsetik	8
4	Undersökning/Genomförande	9
4.1	Enkäten	9
4.2	Testet	10
5	Resultat och Analys	13
5.1	Resultat	13
5.2	Analys	17
6	Sammanfattning och diskussion	22
6.1	Sammanfattning	22
6.2	Diskussion	22
6.3	Samhälleliga och etiska aspekter	24
6.4	Framtida arbete	25
	Referenser	26

1 Introduktion

Games User Research (GUR) (Syddansk Universitet 2026), handlar om att försöka förbättra design inom spel genom en rad olika metoder fokuserade på spelares beteende och upplevelser. Forskning om spelare och deras interaktion med spelmedier har blivit ett brett ämne i takt med datorspelens ökade popularitet. En avgränsning inom forskningen handlar om att utvärdera ett spels användbarhet under utvecklingsfasen för att identifiera förbättringsmöjligheter och åtgärda dessa. En populär metod inom forskningen heter heuristisk användbarhetsutvärdering, från engelskans 'Usability Heuristic Evaluation', härfter HU. HU av digitala spel har utvecklats i skärningspunkten mellan traditionell användbarhetsanalys och orientering kring spelupplevelsen. Det moderna startskottet kom 1982 när Thomas Malone (Malone 1982) kom ut med en lista av heuristiker framtagna för utbildningsspel. Därefter har ämnet vuxit och idag finns det flera etablerade listor vars författare förfäktar att just de punkterna de har med är ett bra rättesnöre att förhålla sig till, framför allt i den tidiga spelutvecklingsfasen. I tillägg till HU, finns användbarhets design, från engelskans User Experience (UX) vilket inom spel blir Games User Experience (GUX). Det är ett område som växer och hanterar just frågor kring användbarhet. Inom UX och GUX talas det om UX-mognad, vilket betyder hur väl ett företag är bestyckat för att förstå, prioritera och systematiskt implementera bra UX design.

Två av de mest populära HU listorna är 10 Usability Heuristics, härfter Nielsens 10, från Jakob Nielsen och Rolf Molich (Nielsen 1994, uppdaterad 2024) och Net Heuristics List, härfter NHL, från Pinelle, Wong och Stach (2008). Dessa listor utgör denna undersöknings grundmaterial. Nielsens 10 är specifikt framtagen för att användas med vilken programvara som helst. Den är IT-agnostisk. NHL är specifikt framtagen för användning mot spel. Spel är visserligen programvara, men är av natur oftast väsentligt annorlunda än till exempel traditionella kommersiella applikationer. Både i syfte och utformande. Krav på användbarhet, det vill säga interaktionen mellan spelare och programmet är hög när det kommer till spel. Inte minst eftersom antal publicerade spel det senaste decenniet har vuxit explosionsartat. För att öka chansen att lyckas är det viktigt att lägga energi på spelets användbarhet. För spelutvecklare som saknar särskild kunskap inom GUX, utforskas vilken av Nielsens 10 och NHL som bäst lämpar sig som hjälpande verktyg i spelutvecklingsfasen för att identifiera och förbättra användbarhetsaspekter.

I undersökningen genomförs ett test och en enkät. Testet tar sig an en djupdykning i jämförelse mellan Nielsens 10 och NHL. Detta görs med hjälp av ett konkret fall som är en utvärdering av ett befintligt spel som funnits på spelmarknaden sedan 2013. Enkäten riktar sig specifikt mot spelutvecklare utan särskild kompetens inom UX som också får jämföra alla de listpunkter som finns med inom Nielsens 10 samt NHL och välja vilka som bäst lämpar sig för ändamålet. Analysen triangulerar därefter datan som testet och enkäten uppvisar och diskuterar fynden utifrån bakgrunden och frågeställningen.

2 Bakgrund

2.1 Användbarhet i spel

För att lyckas med en digital produkt förklarar Buis, Ashby och Kouwenberg (2023), att en av de viktigaste aspekterna är användarens upplevelse (från engelskans 'User Experience', UX). Ett program har kärnfunktioner, som i spel menas med: spelmekaniker, narrativ, ljudambians med mera. Men UX handlar mer konkret om spelarens upplevelse i interaktion med programmet, i det här fallet spelet. Då används begreppet användbarhet. Eftersom det har blivit en så pass viktig faktor att få rätt, har UX blivit en helt egen specialiserad kategori med både egna utbildningar och forskning som resultat. Inom spelbranschen heter det Games User Experience (GUX).

UX-mognad

Chapman och Plewes (2014) pratar om UX-mognad, som i korthet innebär ett mått på den formella kompetensen som innehas inom området som rör användarens upplevelse och interaktion med programvaran. En högre UX-mognad innebär att företaget har en högre kunskap kring att strukturerat och metodiskt tillämpa de principer som innebär bra användbarhet. En lägre UX-mognad behöver inte nödvändigtvis betyda att ett spel är sämre, men att den formella kompetensen bland spelutvecklarna kring fundamentala UX principer antagligen är lägre. Gibbons med flera är UX-konsulter hos Nielsen Group och har identifierat sex olika steg av UX-mognad för att hjälpa företag hitta sin nivå och förstå var eventuella kunskapsluckor befinner sig (Gibbons, Moran, Pernice & Whitenton 2024). Framöver refereras detta som Nielsens skala då stegen är utgivna under namnet Nielsen.

De sex stegen i Nielsens skala är:

Nivå 1: Frånvarande

Nivå 2: Begränsad

Nivå 3: Framväxande

Nivå 4: Strukturerad

Nivå 5: Integrerad

Nivå 6: Användar-driven

Förenklat kan man se nivå 1 som total kunskapslucka. På den här nivån vet en spelutvecklare eller företag inte ens om ämnet och har definitivt ingen fokus på förbättring. Vid nivå 2 börjar upptäckten och en yttlig ansträngning görs – ibland. Den här svaga ansträngningen lämnar mycket att önska. Nivå 3 är någonstans startskottet för en medveten insats. Det finns en tanke och metodologi, men den kan vara inkonsekvent genomförd såväl som ineffektiv. Nivåerna 4-6 innehåller en tilltagande struktur och en systematisk medvetenhet såväl som strukturerade försök att förbättra UX-ansträngningar.

Utvärdering i spelutveckling

Enligt statistik från en av dataspelvärldens mest populära försäljningsplattformar, Steam (SteamDB 2026), har det senaste decenniet visat en kraftig ökning av antal publicerade spel. Från 2015 där 2 822 spel släpptes till 2025 som såg 20 013 lanseringar har denna återförsäljare allena sett en nästan 8-fald ökning i volym. Pressen blir därigenom svårare för en enskild spelutvecklare. Kvaliteten behöver vara hög, inte bara spelmekaniskt och eventuellt narrativ

men även i hur känslan är för spelaren att interagera med mediet. För att öka chansen att slå igenom bruset, är speltestning ovärderligt enligt Agaté (2026). Ständig utvärdering i spelutvecklingsprocessen ökar chansen att en användare som provar ett spel faktiskt vill fortsätta spela, trots den massiva konkurrens som råder på marknaden.

Utvärdering i spelutvecklingsfasen har därmed blivit en viktig komponent och akademien har lagt stora ansträngningar på att kartlägga och identifiera nycklar för framgång. Paraplybegreppet Games User Research (GUR) har blivit centralt inom spelforskningen. GUR har som syfte att försöka förbättra spelbarhet och övergripande designkvalitet genom att förstå spelarnas beteenden, upplevelser och interaktioner (Desurvire & El-Nasr, 2013). Flera metoder och verktyg har utvecklats för att ge bättre insikter hur spelare erfar spelen och hur detta kan förbättras. En sådan, och fokus i detta arbete, är heuristisk användbarhetsutvärdering.

2.2 Heuristisk användbarhetsutvärdering

Heuristisk användbarhetsutvärdering, HU, kommer från engelskans 'Heuristic Usability Evaluation'. Inom dataspelsvärlden handlar ämnet om att mäta hur bra användbarhet ett spel uppvisar, sett utifrån spelarens perspektiv av interaktion med mediet. Det ska vara enkelt för spelaren att interagera med spelet även om spelet i sig är utmanande (Strååt, Rutz och Johansson 2014). Det är numera en väl dokumenterad metod inom GUR som har ambitionen att snabbt och effektivt bedöma huruvida de valda aspekterna är väl utformade enligt sitt syfte – eller om det finns aspekter som behöver fortsatt bearbetning. Olika forskare har tagit fram så kallade heuristiklistor, som är deras samling av vad som kan och enligt författarna till och med bör utvärderas för att säkerställa att ett spels användbarhet håller en hög nivå. Ett par exempel på det som listorna undersöker är hur väl ett spel låter spelaren pausa eller skippa delar, hur intuitivt kontrollstyrning är eller hur tydliga målen och spelets instruktioner är. Tanken är att HU ska tjäna som ett verktyg, inte minst i den tidiga spelutvecklingsfasen (Nielsen 1994, uppdaterad 2024). Detta sker genom att spelutvecklare, ofta med hjälp av speltestare, jämför sitt spel mot parametrarna som återfinns i en eller flera heuristiklistor och betygsätter därefter.

En av fördelarna med heuristisk användbarhetsutvärdering är att det är snabbt och effektivt att ge en, i breda drag, återkoppling på spelets kvalitet utifrån de valda användbarhetsaspekterna (Omar och Jaafar 2010). Kostnaden för att utvärdera är i princip endast tid. Metoden är flexibel, vilket gör att den därmed passar de flesta spelutvecklare, även nystartade spelföretag.

Spelföretag som har begränsade resurser och som därigenom har en låg UX-mognad har nytta av heuristisk användbarhetsutvärdering eftersom den i sammanhanget av GUX är en av de enklare metoderna att förklara och förstå anser Wilson (2014). Heuristisk användbarhetsutvärdering blir således ett stöd för att finputs ett spel och en väsentlig del av utvecklingsfasen. Större och mer etablerade studios kan ha tillräckliga resurser för att tillsätta spetskompetens inom GUX, men även de kan tillskansa sig nyttan av metoden som en del av deras utvärderingsprocess.

Wilson (2014) identifierar samtidigt flera brister med HU. Eftersom det är människor som utvärderar blir både tolkningar såväl som förståelse för olika begrepp högst personligt. Vad är bra UX design i menyn? Fråga två personer och få två olika svar. Vilka brister har spelet inom användbarhet? Fråga olika forskare eller spelutvecklare och återigen få olika svar. Dessutom är det inte alltid att de som utvärderar spelet är den tänkta konsumenten och kan sakna

kunskap kring etablerad praxis inom olika genrer eller erfarenhet att spela spel generellt.

Olika heuristiklistor

Ämnets moderna ursprung kommer från 1980-talet när Malone (1982) började belysa ämnet i den akademiska litteraturen. Hans ambition var att utveckla riktlinjer för underhållande gränssnitt.

Ungefär 20 år senare, utvecklade Desurvire, H., Caplan, M., & Toth, J. A. (2004) Heuristics for Evaluating Playability (HEP) vilket är en ofta citerad studie. HEP är ett sätt att mäta hur bra ett spel svarar upp mot användbarhet i själva utformandet. HEP blev till en slags checklista mot vilken olika delar av ett spel kunde mätas. Listan vidareutvecklades senare till Heuristics of Playability, kortfattat PLAY (Desurvire, H., & Wiberg, C. 2009) där fokus flyttades närmare spelarens upplevelse, snarare än ren funktionalitet.

Därtill finns Jakob Nielsen och Rolf Molich som tillsammans tog fram grunden för 10 Usability Heuristics, härafter kallat Niensens 10 (Nielsen 1994, uppdaterad 2024) som är en av de mest välkända och ofta citerade listorna. Likt de tidigare nämnda listorna, har Niensens 10 inspiration med start från Malone (1982). Den är dock framtagen för mjukvara i stort och utger sig inte heller vara specifikt för spel. Listan innehåller 10 punkter som en spelutvecklare bör vara uppmärksam på, för att lyckas med interaktionen mellan mjukvara och slutanvändare.

Pinelle, Wong och Stach (2008) utgick bland annat från Niensens 10 men även andras tidigare arbeten och valde att sätta ihop en summerad heuristiklista, utformad specifikt för spel. För att välja ut vilka punkter som skulle vara med, började de med att analysera befintliga spel. Efter att ha undersökt och identifierat problem inom användbarhet genom en grundlig analys av 108 spelrecensioner, sammansatte de Net Heuristics List, NHL (Pinelle, Wong och Stach 2008). Deras lista, innehåller likt listan från Nielsen & Molich 10 punkter som ska täcka in användbarhet men då specifikt i spel. Därför återfinns både likheter men olikheter mellan listorna. Även NHL är en av de listor som ofta citeras inom den akademiska litteraturen.

Heuristiken i sammanhang

Både Niensens 10 och NHL är som nämnt välkända och citeras i andra arbeten. De används ofta båda två i ett bakgrundssammanhang när heuristisk användbarhetsutvärdering undersöks.

Till exempel undrar Strååt, Rutz och Johansson (2014) om det går det att mäta ett spels kvalitet, utifrån perspektivet användbarhet, med hjälp av den heuristiska metoden. De skrev en artikel där de först granskade den akademiska litteraturen såväl som att själva utforska ämnet i en förenklad studie. Efter en genomgång av både Niensens 10, HEP, PLAY och NHL utgår de från den sistnämnda och utvärderade tio olika spel (Strååt et al 2014). Deras slutsats var att heuristisk användbarhetsutvärdering är ett värdefullt verktyg för spelutvecklare. De ansåg även att NHL i det stora hela var bra, trots att de avslutade med några egna tillägg till listan.

Liao och Shen (2012) tog en annan vinkel och undersökte specifikt inlärningsspel. De tar framför allt avstamp ur Niensens 10 och nämner andra listor övergripande. Även dessa författare bekräftade att de finner stöd i att HU har en stor fördel eftersom det är en snabb, billig såväl som lättillgänglig metod för spelutvecklare att utvärdera sitt spels användbarhet. Den kritik de ändå för fram läggs dels mot att Niensens 10 är för allmänna och inte tillräckliga för kategorin utbildningsspel.

3 Problemformulering

Agnostiskt eller spelspecifikt

NHL, som har vissa rötter ur Nielsens 10, är specifikt framtagen för och från spel medan Nielsens 10 är erkänt IT-agnostisk. Båda används flitigt inom den akademiska litteraturen, och båda har uppvisad relevans inom forskningen av spel.

Båda listor innehåller tio punkter att förhålla sig till, vilket gör att de lämpar sig för närmre jämförelse. Listorna är korta och koncisa och det finns flera gemensamma nämnare. Båda adresserar liknande aspekter. Båda utgår dessutom från idén att slutanvändarens upplevelse ska stå i fokus och att interaktionen mellan användare och mjukvara ska vara smidig. Båda är väl undersökta av akademiker och forskare. Med olika frågeställningar och sammanhang har dessa forskare nått fram till olika slutsatser om vad som är mest relevant innehåll när heuristisk användbarhetsutvärdering ska användas som metod under spelutvecklingsfasen.

Listorna skiljer sig också åt dels genom att Nielsens 10 har lite bredare formuleringar medan NHL är mer spelorienterad i sitt språk. Nielsens 10 har även med två punkter kring fel och felmeddelanden, medan NHL som inte har dessa två punkter, djupdyker lite mer kring spelarens möjlighet till personliga inställningar.

Det akademiska underlaget påvisar att båda listor är relevanta i spelsammanhang och båda har en uttalad fokus på slutanvändarens upplevelse. Men är den ena bättre än den andra? Eftersom HU anses vara lättillgänglig även för spelutvecklare med låg UX-mognad blir det även intressant att undersöka huruvida en av de utvalda listorna lämpar sig mer naturligt än en annan. Föredrar spelutvecklare det förenklade breda språket från Nielsens 10, eller det mer spelorienterade ordvalet som återfinns i NHL? Även om inte NHL ska tolkas som en definitiv jämförelse för alla heuristiklistor som riktar in sig specifikt mot spel, passar den väl i jämförelse med Nielsens 10. De delar arv, citeras ofta och en jämförelse blir både rimlig och träffande.

Frågeställningen

Ifall spelutvecklaren ställs inför bägge listor och ska välja vad som är mest relevant utifrån deras upplevelse och perspektiv, vad blir resultatet? Är det givet att en lista fokuserad på spel är mer träffsäker för att belysa användbarhetsproblem än en lista som är generell för all typ av mjukvara? Spel är trots allt en typ av mjukvara. Skulle en sådan undersökning resultera i blandade resultat? Denna undersökning ämnar gå direkt till källan, spelutvecklaren utan särskilt uttalad GUX-kompetens, och deras upplevelse för svar till frågan: vilken av Nielsens 10 och NHL är mest lämplig för utvärdering av användbarhet i spel?

3.1 Metodbeskrivning

Mål

Målet är att identifiera rimliga metoder för insamling av data från spelutvecklare utan särskild GUX-kompetens som får utvärdera både Nielsens 10 och NHL. Det insamlade underlagets syfte är att identifiera och förstå trender samt dra slutsatser som är relevanta och av akademiskt värde.

Metodval

För att nå målet med empiriska data och en förståelse av denna, kommer en blandad kvantitativ och kvalitativ metod med fokus på enkätsvar och ett test användas. Att genomföra

två undersökningar möjliggör att enkäten kan jämföras med testet och vice versa för att styrka trovärdigheten i de resultat som presenteras. Både enkäten såväl som testet har vardera en kvantitativ såväl som en kvalitativ del.

Enkäten

En enkät ger ett effektivt och smidigt sätt att samla in data på (Gül, Yavuz. 2023). Gül summerar även att genom kvantitativ datainsamling kan trender eller åsikter noteras som är svåra att fånga upp genom observationsmetoder. Eftersom tidsramen för undersökningen dessutom är snäv, är det ett plus att enkätundersökningar utgör en relativt enkel och snabb insamlingsteknik som dessutom går att skala upp utan större svårigheter. Det är en vanlig metod i liknande sammanhang vilket betyder att valet ger akademisk relevans och att den är generellt accepterad att använda i det valda sammanhanget (Richardson, M., Carter, F., Levordashka, A., Voinescu, A., Gilchrist, I.D. & Stanton Fraser, D. 2025). Förhoppning med enkäten är minst 10 svar, givet att enkäten riktar sig mot en nischad målgrupp.

Enkätens frågor är till huvuddel kvantifierade. Första huvudfrågan består av förvalda svarsalternativ som innehåller respektive listors gemensamma punkter. Sex av tio punkter från listorna korrelerar med varandra. Spelutvecklaren som deltar får under relevanta rubriker kryssa i det alternativ, ett från Nielsens 10 eller ett från NHL, som de anser vara bäst lämpad för att utvärdera spel. I huvudfråga två, får deltagaren välja ut fyra alternativ som de anser kompletterar de första sex från huvudfråga ett. Detta görs genom att övriga åtta listpunkter, fyra från vardera lista, presenteras. I bägge huvudfrågor kommer deltagaren inte få någon information om vilket alternativ som kommer från vilken lista. Data kommer objektivt kunna sammanställas och tolkas för att belysa empiriskt vad deltagarna tycker. Detta utan risk för bias och feltolkning av forskaren.

Borg och Westerlund (2014) belyser att kvalitativa metoder framför allt innehåller språkligt beskrivande. Enkäten avslutas med en frivillig öppen frågeställning med fritextsvar. Detta möjliggör att deltagaren kan motivera, tycka till eller på annat sätt förmedla åsikter som inte ryms inom de kvantifierbara frågorna. Denna kvalitativa aspekt av enkäten har som syfte att ge utrymme för en djupare förståelse av de svar som inkommit.

Testet

I tillägg till enkäten kommer ett test genomföras av undertecknad och en till person. Syftet är att gå igenom ett spel, Dota 2 (Valve 2013) och jämföra både Nielsens 10 och NHL för att identifiera ifall den ena eller den andra är mer relevant eller tydligare pekar på de aspekter som ska utvärderas. En kollega, som tidigare varit med undertecknad och skrivit en litteraturanlys inom ämnet kommer göra samma test och resultaten kommer jämföras. Eftersom båda tidigare har undersökt ämnet djupare anses båda vara kvalificerade för uppgiften.

Testet utgår från enkätens två huvudfrågor för att få in kvantifierbara svar kring preferens mellan Nielsens 10 och NHL när Dota 2 undersöks. Före genomgången av heuristiklistorna, sker ett samtal mellan de två testdeltagarna för att hitta samförståelse kring begrepp och koncept. Efter testet sker en diskussion för att analysera svar och fynd. Denna kvalitativa del blir ett djupdykande komplement till den tidigare kvantitativa utvärderingen.

Testet som genomförs inklusive diskussionstillfällen före och efter, ger möjligheten till ett större djup som inslag i analysen av hur väl de olika listorna uppmuntrar användbarhetsutvärdering med konkret utgång från Dota 2.

Urval

Enkäten kommer spridas genom en kombination av kriterieurval, bekvämlighetsurval och snöbollsurval.

Kriterieurval beror på ett krav för deltagande som är att deltagaren antingen är spelutvecklare, aspirerande spelutvecklare alternativt student inom spelutveckling. Samtidigt är problemformuleringen inriktad mot spelutvecklare utan särskild GUX-kompetens. Med detta avses en UX-mognad enligt Nielsens skala mellan ett och tre. På den här nivån måste det inte, men kan finnas en kännedom kring ämnet och även en ansats till att omsätta den i praktiken. Samtidigt saknas det en tydlig struktur som kännetecknar de högre nivåerna av UX-mognad. Detta kriterium säkerställer att det erhållna empiriska underlaget har högre trovärdighet i analysen och till den ställda frågan.

Kombinationen bekvämlighetsurval och snöbollsurval beror på att författaren först kommer nå ut till de som är mest lättillgängliga inom den sociala närkretsen. Därefter för att få en större spridning kommer enkäten, om möjligt, delas med hjälp av deltagarna som kan länka vidare till andra som passar kriterieurvalet.

Urvalet kring testdeltagare följer samma mönster, fast begränsat till kombinationen kriterieurval och bekvämlighetsurval. Det kommer vara den medkollega som samskrivit en litteraturundersökning inom ämnet HU och således både är insatt i ämnet såväl som tillgänglig och villig att bistå undersökningen. Testdeltagarna anses ligga på en trea i Nielsens skala, då de saknar konkret erfarenhet från att metodiskt omsätta kunskap i praktiken.

Metodkritik och möjlighet

Begränsningarna med studiens enkät är att de ger ett relativt litet svängrum för deltagaren att motivera och djupdyka kring frågeställningarna. Den fritextfråga som finns är frivillig. Det finns en risk att den kvalitativa aspekten blir lidande för enkäten, då deltagaren inte leds fram genom flera frågor som vid en mer djupgående fritextenkät eller ett intervjutillfälle. Samtidigt säkerställer testet som innehåller diskussion mellan testdeltagarna att en viss djupdykning ändå kommer ske i ena undersökningen och bidrar med en kvalitativ nyans till helheten.

Eftersom enkätdeltagare kan ha en varierad förståelse för koncept som återspeglas av svarsalternativen finns även en risk att svaren blir skeva i förhållande till varandra. Detta mitigeras genom att enkäten initialt har en beskrivning av HU, syftet med HU som metod såväl som syftet med undersökningen i stort.

Ett av de största problemen med HU enligt Wilson (2014) är subjektivitet som kan ge olika svar på grund av olika egna tolkningar när djupare undersökningar, likt testet, genomförs. Eventuella osäkerheter antas undanröjas till viss del eftersom testets genomförande sker av två individer som tidigare jobbat ihop inom ämnet och bollat fram en gemensam förståelse för begrepp. Dessutom kommer de inför testningen ha en genomgång av de olika kategorierna för att säkerställa konsensus kring olika begrepps syfte och betydelse.

En stor möjlighet är att det finns två undersökningar som kan triangulera resultaten i den slutliga komparativa analysen. Med enkäten och testet tillsammans kan de kvalitativa motiveringarna belysa aspekter som den kvantitativa data endast summativt redogör.

3.2 Förväntat resultat

Förväntningen är att NHL, som spelspecifik heuristiklista, har högre träffsäkerhet än Nielsens 10. När spelutvecklaren får säga sitt, och med sina begränsade kunskaper kring heuristik och användbarhetsutvärdering i åtanke borde formuleringar som träffar mer naturligt rätt till spel

väljas oftare. Samtidigt finns det skillnader mellan listorna där deltagarna mycket väl kan föredra olika vinklar än exakt vad NHL förespråkar.

Förhoppningen är i slutänden att få ett dataunderlag som ger ett tydligt svar på frågeställningen med en hög reliabilitet som går att upprepa med samma resultat (Eliasson 2019). Antingen ger den tonvikt åt ena eller andra listan och oavsett vilket, visar något nytt fynd.

3.3 Forskningsetik

I den utsträckning möjligt och applicerbart följer detta arbete god forskningsetik enligt den europeiska ALLEA-kodexen (Vetenskapsrådet, 2024).

Tillförlitlighet eftersträvas i fråga om forskningens kvalitet. Både den kvantitativa såväl som kvalitativa data som presenteras är representativ och slutsatser är dragna med varsamhet.

Ärlighet uppvisas mot både deltagare såväl som i arbetet överlag. Deltagare är anonyma och trots det, är de välinformerade kring rollen som enkäten har i undersökningen. Möjlighet finns att ställa frågor, inkomma med invändningar eller andra kommentarer kring enskilda aspekter eller arbetet i stort.

Respekt för deltagare och kollega som utför enkät och test. Deras tid och engagemang värdesätts och deras inslag tas med i noggrann beaktning för resultatens utvärdering. Dessutom, för den kollega som deltar i undersökningens test, finns möjligheten tydligt förklarad att närsomhelst kunna avsäga sitt deltagande och få sitt bidrag struket från rapporten.

Ansvar att ta arbetet från idé till publicerat resultat, med stöd av handledare och examinator vilar hos undertecknad som gör ett ärligt försök att nå till godkänd akademisk kvalitet och med en bibehållen forskningsetik.

I övrigt har detta arbete en ambition att tillföra forskningen ny kunskap och verktyg för fortsatt utveckling inom området heuristisk användbarhetsutvärdering.

4 Undersökning/Genomförande

Undersökningen består av två delar, en enkät och ett test. Enkäten genomförs med hjälp av målgruppen, spelutvecklare utan särskild spetskompetens inom GUX. Testet genomförs av två personer som också passar in på beskrivningen spelutvecklare utan spetskompetens inom GUX, men med skillnad att de gjort tidigare arbete i ämnet. Detta medför att de således har införskaffat en något djupare kunskap inom ämnet och insikter kring det rådande läget inom den akademiska litteraturen.

4.1 Enkäten

Enkäten riktar sig mot spelutvecklare utan särskild GUX-kompetens, vilket här definieras som en UX-mognad 1-3 enligt Niensens skala. Det finns alltså en medvetenhet, men det saknas en djupare kunskap och framför allt tillämpning.

Enkät – planering

För att säkerställa att rätt målgrupp svarar på enkäten planeras två inledande frågor för att urskilja kriterieurvalet.

För att forma en enkät som jämför Niensens 10 och NHL sida vid sida, granskas båda listorna noggrant. En kartläggning över de likheter som båda listor tar upp noteras. En extra kontroll utförs för att säkerställa att de listpunkter som identifieras från vardera lista faktiskt till rimlig grad stämmer överens. Detta borde resultera i några punkter från varje lista där en direkt jämförelse blir relevant. Det bör även resultera i några punkter som inte kan jämföras direkt utan som tillägg i slutet.

Därefter blir det naturligt att listorna undersöks på två sätt. Ett som jämför likheter, ett som identifierar olikheter – och vilka av dessa som upplevs mest relevanta i frågeställningen. En enkät kan då med fördel delas upp i två huvudfrågor. Den ena som jämför likartade formuleringar, den andra utgör en utvärdering av vilka övriga delar som anses bäst passa in på en heuristisk användbarhetsutvärdering.

För att få med en kvalitativ aspekt av enkäten planerades en avslutande öppen textfråga, där medverkande får kommentera eller motivera tidigare svar. Alternativt kan de själva bidra med en förändring som de anser inte ryms inom ramen för de tidigare enkätfrågorna. Detta ska inte vara en tvingande fråga utan ger en frivillig möjlighet ifall deltagaren önskar tycka till på ett djupare plan.

När enkätsvaren samlats in kan en kvantitativ översikt kring trender ställas upp samtidigt som de fritextsvar som kommit in granskas och sorteras ifall flera hamnar inom samma kategori. Data kan därefter analyseras för att leta svar på problemformuleringen. Enkätens huvudfrågor lämpar sig för en summativ analys, alltså att räkna ihop hur många svar som hamnade på vilken av de två undersökta listorna. Antal svar ger en överblick över problemformuleringens svarsriktning. Beroende på hur fördelningen mellan listorna är, kan mer eller mindre definitiva svar noteras. Den avslutande frågan ger möjlighet för djupdykning samt reflektion och eventuella motiveringar kan kopplas till de motiveringar och reflektioner som presenteras.

Enkäten kommer skrivas på engelska för att ha möjlighet att nå en bredare målgrupp som inte är begränsad av svenska som modersmål.

Enkät – genomförande och datainsamling

Enkäten fick först två frågor för att säkerställa att deltagaren tillhörde rätt målgrupp. En fråga kring ifall de var spelutvecklare eller student. En fråga kring ifall de känner till HU sedan tidigare.

Efter en rannsakan av Nielsens 10 och NHL kunde listorna sorteras i en jämförande och särskiljande kategori. Enkäten fick som resultat två huvudfrågor. Huvudfråga ett använde följande kategorier, där båda listor har gemensamma formuleringar.

- Presenteras relevant spelinformation till användaren?
- Hur tydliga är spelinformationens estetik?
- Är spelet konsekvent och som förväntat?
- Erbjuder spelet flexibilitet för spelarinställningar?
- Behöver spelaren komma ihåg allt eller finns viktig information lättillgänglig?
- Ger spelet hjälp och utbildning?

Både Nielsens 10 såväl som NHL har en korresponderande punkt som matchar ovan rubriker och deltagaren får välja ut den formulering som de anser bäst hjälper utvärdera ämnet.

Huvudfråga två listar därefter övriga punkter, fyra från respektive lista, med uppmaningen att deltagaren väljer ut de delar som de anser bäst kompletterar svaren från första huvudfrågan. Tillsammans ger detta en bild över vilka tio punkter som enligt enkätdeltagarna bäst stöttar HU inom spelutveckling.

Med en färdig enkät utifrån heuristiklistorna, avslutades enkäten med en fritextfråga som gav möjlighet att motivera tidigare svar. Enkäten är bifogad i Appendix A – Enkät- och Testfrågor. Därefter söktes kandidater upp och enkäten genomfördes enligt följande:

1. Första urvalet, bekvämlighetsurvalet, söktes upp genom direkt kontakt. Detta skedde primärt via privata meddelanden. Kriteriet för de potentiella deltagare som uppsöktes var att de antingen skulle vara spelutvecklare utan särskild GUX-kompetens eller studenter inom området. Detta för att de ska passa in målgruppen som kan tänkas använda HU samtidigt som de inte redan ska ha hög kunskap inom området.
2. Enkäten besvarades av de medverkande genom flerval och fritextsvar.
3. Därefter fick deltagare, om de ville, skicka vidare enkäten till andra. Detta endast efter en kort beskrivning av kriterieurvalet från undertecknad. Det skedde några gånger och snöbollsurvalet var fastställt.
4. Enkätens svar ställdes upp mer överskådligt, analyserades och fynden presenteras i avsnitt 5 under delresultat - enkät.

Enkäten var öppen för deltagande i en arbetsvecka, fem dagar.

4.2 Testet

Testets syfte handlar om att konkret jämföra både Nielsens 10 såväl som NHL mot ett spel, Dota 2. Spelet används för att utvärdera vilken lista som bäst svarar upp mot heuristisk användbarhetsutvärdering och bidrar med både kvantitativ såväl som kvalitativ datainsamling. Även testet riktar sig mot spelutvecklare utan särskild GUX-kompetens, vilket även här definieras som en UX-mognad 1-3 enligt Nielsens skala.

Test – planering

Urvalsstrategin utvärderas då det inom räckvidden av ett bekvämlighetsurval finns personer med varierande grad av kompetens inom GUX. Både de med formell utbildning, som kan

anses ligga i de högsta skikten av Nielsens skala såväl som andra med lägre eller även ingen kunskap och erfarenhet. I flera av de tidigare undersökta artiklarna genomfördes undersökningar med hjälp av individer som anses ha en hög UX-mognad och det föranledde en övervägning att även i detta test göra samma, till skillnad från målgruppen i enkäten. Här blir ändå slutsatsen att det bästa är att hålla testet så enhetligt med enkäten som möjligt, för att den planerade trianguleringen inte skulle bli för ojämn. Samma motivering gör att testdeltagare helt utan GUX-känedom väljs bort och kriterieurvalet förankras.

Eftersom planeringen av enkäten redan är i gång anses det lämpligt att återanvända enkätens underlag även inom ramen för testet. Att utgå från enkäten anses bidra till att det blir enklare att jämföra svar i efterhand. Både mellan testdeltagarna såväl som mellan enkät och test som helhet. Testdeltagarna får på så vis var för sig utvärdera båda listor och identifiera vilken av listorna som är bäst lämpad för att vägleda en spelutvecklare med HU när Dota 2 undersöks.

Då en av testets svagheter anses vara subjektivitet när den akademiska litteraturen undersöktes, blir det viktigt att båda testdeltagare har samsyn kring begrepp och förståelse för koncept. Därför planeras en åtgärd att inleda genomförandet med en gemensam diskussion där ämnena från båda listor undersöks och en konsensus kring definitioner och betydelse kan läggas som en gemensam grund. För att ytterligare säkerställa en gemensam förståelse kring de faktiska svaren med tillhörande tankar och motiveringar planeras att ett uppföljande samtal genomförs efter testet.

Svaren ska noteras skriftligen med tillhörande motivering vid varje listpunkt och vid testets genomförande. Detta säkerställer att svaren är färska och att dessa kan ligga till grund för den efterföljande diskussionen där ytterligare reflektionsutrymme ges.

Med hänvisning till tidsåtgången som förväntas är det även rimligt att testdeltagaren får en kompensation i form av en biobiljett för sin medverkan. Tidsåtgången för testet inklusive inledande och avslutande diskussioner uppskattas landa på ca 2,5 - 3 timmar.

Test – genomförande och datainsamling

I genomförandet av testet var huvudsyftet att utvärdera Nielsens 10 och NHL mot varandra, snarare än att hitta varenda styrka eller svaghet i Dota 2. Testet genomfördes enligt följande. Testdeltagarna ligger i det högre spannet på Nielsens skala 1-3.

Urval:

1. Utöver undertecknad, identifierades en deltagare som var villig och tillgänglig via privat meddelande.

Testet gick till så här:

1. Författaren plus den andra testdeltagaren hade ett uppstartsmöte för att gå igenom testet. Där diskuterades samtliga huvudmoment, men även begrepp inom HU och specifikt de som utvärderas i testet. Konsensus etablerades kring syfte med olika underkategorier.
2. Dota 2 spelades igenom två gånger i 'turbo' läge, vilket är en officiell och vid undersökningens tillfälle den mest populära spelversionen. Turbo innebär att spelet går lite snabbare, men i övrigt är reglerna samma och spelmekanikerna oförändrade.
3. Båda testdeltagaren genomförde den diskuterade enkäten på varsitt håll och noterade sina svar skriftligen. Till skillnad från enkäten så antecknades en svarsmotivering vid varje punkt som utvärderades. Detta utgick från de aspekter som ansågs vara brister eller styrkor i Dota 2 med motivering kring varför.

4. Svaren delades till författaren, som sorterade och jämförde resultaten.
5. En uppföljande diskussion ägde rum ett par dagar senare för att ge möjlighet för motivering, förtydligande vid behov och att dela fynden.
6. Nyckelresultaten delas under i avsnitt 5 under delresultat - test.

5 Resultat och Analys

I det här arbetet analyseras 10 Usability Heuristics, Niensens 10 (Nielsen 1994, uppdaterad 2024) och Net Heuristics List, NHL (Pinelle, Wong och Stach 2008). Målet med både enkät och test var att jämföra vilken av de två listorna som bäst lämpar sig som metod för spelutvecklare utan särskild kompetens inom GUX när det gäller användbarhetsutvärdering av spel. Spelutvecklingsfasen är föränderlig i sin natur, och det är viktigt att hitta rätt verktyg som kan hjälpa spelutvecklare att identifiera huruvida deras produkt återspeglar bra design utifrån användbarhet för slutanvändaren. Analysen som följer tar direkt avstamp från resultaten som både enkät och test gett upphov till.

5.1 Resultat

Delresultat - enkät

Enkäten resulterade i totalt 14 svarande spelutvecklare som alla träffade in på Niensens skala 1-3. 14 deltagare som jämförde de två listorna, producerade totalt 140 svar i enkäten. Varje lista har 10 punkter så bland de 20 svarsalternativ som fanns kunde varje spelutvecklare välja 10. 10 svar x 14 enkätmedtagare = 140 svar totalt. Bild 1 visar hur svaren fördelades mellan Niensens 10 och NHL.

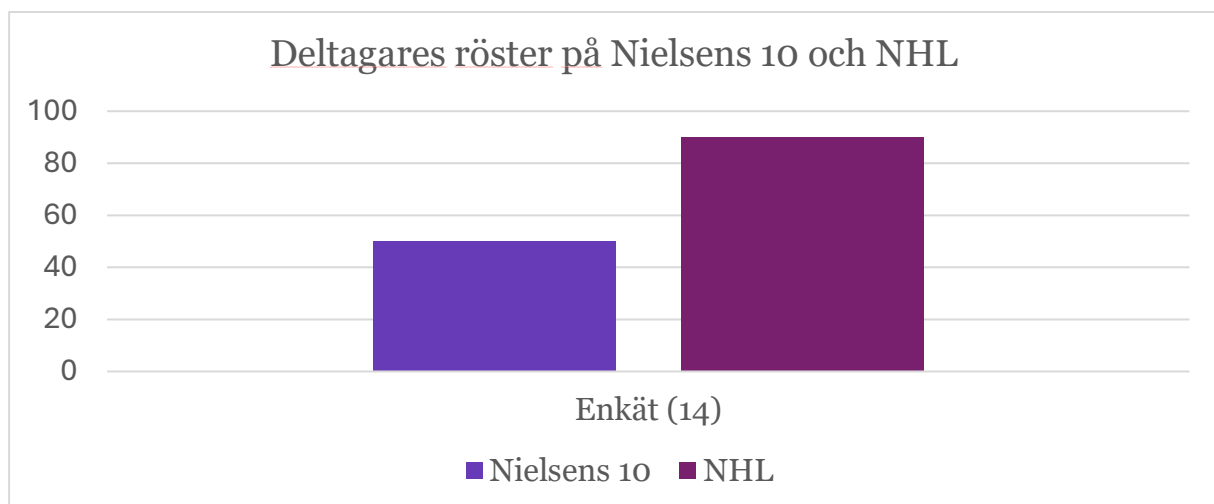


Bild 1: enkätresultatens fördelning, antal svar på Niensens 10 och NHL.

I huvudfråga ett, där både Niensens 10 och NHL överlappar och de två listorna skulle jämföras sida vid sida återfanns följande resultat. I fyra av de sex kategorier som presenterades ansågs NHL som den av de två listorna som bäst lämpade sig för utvärdering av användbarhet bland de spelutvecklare som genomförde undersökningen. I en av sex frågor föredrar enkätmedtagarna Niensens 10 över NHL. Och i en av sex frågor blev resultatet lika många röster på båda listorna.

Enkätfrågan där Niensens 10 ansågs bäst formulerad för att hjälpa utvärdera användbarhet var: Is the relevant game and/or system information presented to the user? 8 deltagare svarade att Niensens 10 beskrivning av problematiken bäst hjälpte utvärdera frågan och NHL fick 6 röster.

Enkätfrågan där det var jämnt mellan de två listorna var: Does the player have to remember everything or is important information readily available? Här ansåg alltså 7 deltagare att Niensens 10 var den bästa vägledningen för svaret och 7 deltagare att NHL formulering var

mest optimal för att leda fram till essensen i utvärderingsfrågan.

I övriga enkätfrågor och även vad gäller enkäten i stort var en majoritet av enkättagarna av åsikten att NHL hade formuleringar som bäst lämpade sig för att utvärdera användbarhet inom spel än Nielsens 10.

När det gäller huvudfråga två, som bestod av åtta listpunkter, fyra från vardera Nielsens 10 såväl som NHL att välja mellan, valdes NHL fyra kvarstående punkter med en tydlig majoritet. Här noteras att trots den tydliga majoriteten fanns även flera röster på Nielsens 10. Resultaten från huvudfråga två gestaltas genom Bild 2 nedan.

Which four (4) of the following aspects would you find most helpful when evaluating the usability of a game?

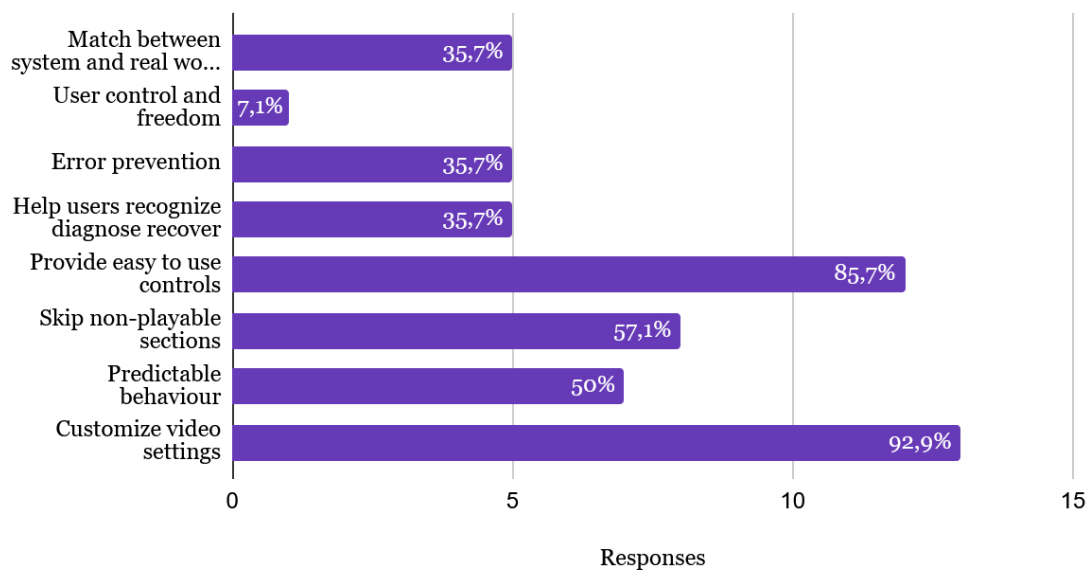


Bild 2: visar spridningen på enkätens huvudfråga två och vad deltagarna valde bland de olika listpunkterna. De första fyra kommer från Nielsens 10, de sista fyra från NHL.

Enkäten avslutades med en frivillig möjlighet att lämna fritextsvar kring motiveringar eller på annat sätt tycka till, vilket 9 av deltagarna gjorde. Tre av de nio fritextsvaren avslöjade inget som bidrog till resultattolkning och förkastades. Innehållsmässigt kunde de summeras med den fiktiva kommentaren,

I am satisfied with my answer. (Anonym enkättagare ett).

Övriga motiveringar diskuteras under en efterföljande analys för att belysa de underliggande motiveringarna till enkätens svarsresultat som presenterats ovan.

Delresultat - test

Testet genomfördes av två testdeltagare. De två testdeltagarna producerade 20 svar efter enkätens utformning och jämförelse av Nielsens 10 och NHL. 10 svar per deltagare x 2 deltagare = 20 svar totalt. Bild 3 visar hur svaren fördelades mellan Nielsens 10 och NHL.

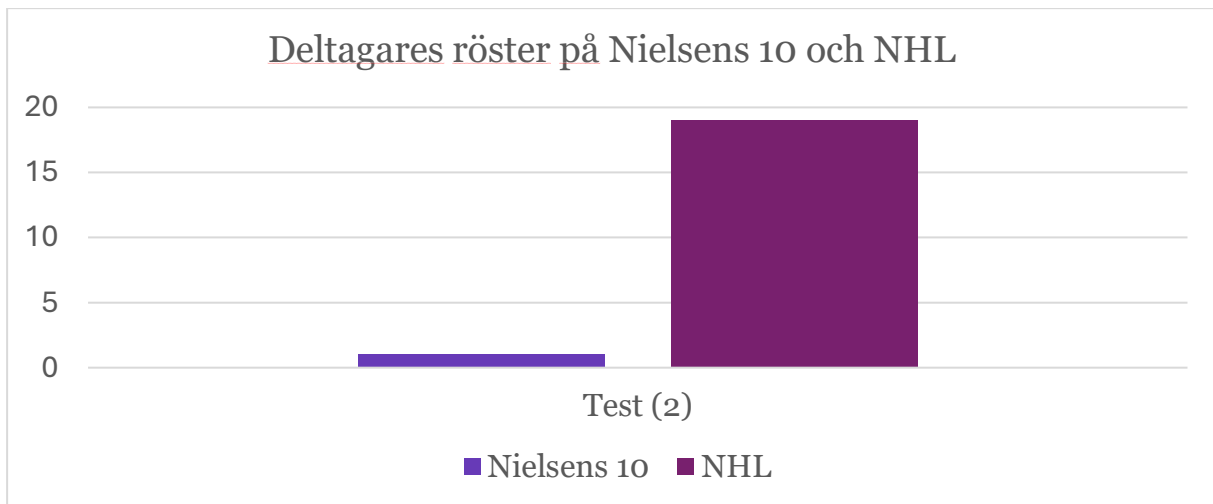


Bild 3: testresultatens fördelning, antal svar på Niensens 10 och NHL.

Båda testare var överens på nästan alla punkter att Net Heuristic List var den listan som var bättre lämpad att sporra fram en utvärdering kring heuristisk användbarhet i spel, när Dota 2 undersöktes.

I huvudfråga ett valde båda testdeltagare samma svarsalternativ på samtliga av de jämförande frågorna. NHL var enhälligt den lista som ansågs bäst lämpad av de två för HU av Dota 2.

Huvudfråga två visade liknande resultat till fördel för NHL. Det var så pass entydigt att alla punkter förutom en fick samma svar. Tre av de fyra listpunkter som skulle väljas ut, av åtta totalt, stämde överens med varandra. Dessutom var båda testdeltagare i väsentliga drag överens om motiveringar varför de valt som de gjorde, men här fanns vissa nyansskillnader i bedömningar kring hur väl eller inte Dota 2 uppfyller kraven för bra användbarhet.

Bild 4 ger en överblick av hur testdeltagarna svarade på huvudfråga två, som handlade om att sortera bland övriga listpunkter och välja ut de som ansågs mest relevanta för HU och i det här fallet även specifikt när Dota 2 utvärderades.

Which four (4) of the following aspects would you find most helpful when evaluating the usability of a game?

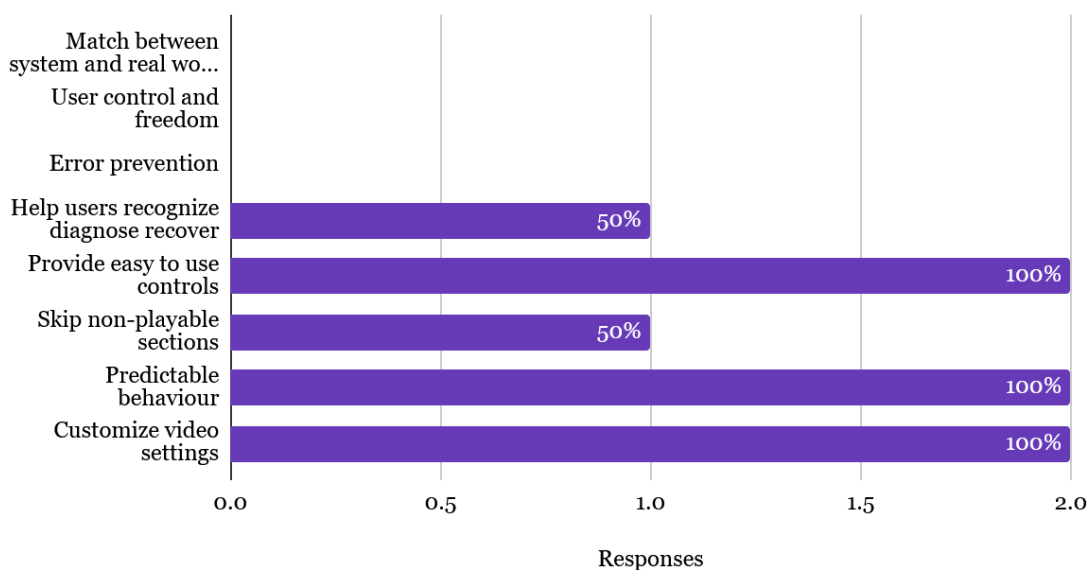


Bild 4: visar spridningen på huvudfråga två enligt testdeltagarnas svar. De första fyra alternativen kommer från Nielsens 10, de sista fyra från NHL.

Testet som genomfördes skiljde sig mot enkäten då varje fråga gav utrymme för motivering och lite mer analytisk redogörelse kring svar och varför. Överlag hade båda testdeltagare liknande motiveringar och åsikter, men i framförallt två avseenden återfanns skillnader. De två punkterna som fick olika röster åt var som följer. Från Nielsens 10, nummer nio:

Help Users Recognize, Diagnose and Recover from Errors. Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution. (Nielsen 1994, uppdaterad 2024).

Och från NHL, nummer fem:

Allow users to skip non-playable and frequently repeated content. Many games include lengthy audio and video sequences, or other types of non-interactive content. Games should allow users to skip non-playable content so that it does not interfere with gameplay. (Pinelle, Wong och Stach 2008).

I en efterföljande analys av testet undersöks både likheter och olikheter som återfanns i motiveringar. Dessa analyseras i förhållande till ämnet och Dota 2 och både den skriftliga återkopplingen såväl som den muntliga diskussionen analyseras.

Övergripande resultat

Nielsens 10 ansågs vara starkt formulerad på första jämförande frågan, Is the relevant game and/or system information presented to the user? Här svarade en majoritet att de tyckte Nielsens 10 var mest lämplig för HU i ämnet. Även på frågan, Does the player have to remember everything or is important information readily available?, var Nielsens 10 stark eftersom resultatet visade lika många röster på båda svarsalternativ. När det gäller att presentera information på ett tydligt sätt, tyckte enkätdeltagarna genom sina svar att Nielsens 10 var tydlig att förstå och vägleda en spelutvecklare som vill tillämpa HU.

När det gäller personliga inställningar, flexibilitet och ämnet träning och hjälp uppvisade inte Nielsens 10 samma styrka utan NHL fick störst stöd bland enkätdeltagarna. Frågan, Is the game consistent and as expected? This can be in regards to input, game mechanics, story, enemy behavior etc., var en tydlig fördel för NHL som fick 10 av 14 röster. Även frågorna, Does the game offer flexibility for the users preferences? och Does the game provide training and help? visade att NHL i större utsträckning hade starka formuleringar för att fånga upp sina respektive kärnfrågor inom HU.

Även om vardera lista hade sina styrkor och omvänt även svagheter, berättar resultaten att på en helhet ansåg både de svarande enkät- samt testdeltagarna till en övergripande fördel för NHL. Listan ansågs vara den starkare kandidaten när dessa två listor jämförs av spelutvecklare med en GUX-kompetens 1-3 enligt Nielsens skala. Samtidigt vid närmre granskning uppvisas en variation i svaren som ändå påtalar att resultaten gömmer en komplexitet i frågan.

Som bakgrunden belyser, finns det inom ämnet ännu flera listor som alla bidrar med olika nyanser på testning inom användbarhet. Denna jämförelse undersökte två av dessa listor. Att svaren som framkom visar att det inte finns ett entydigt svar, berättar att frågan är mer komplex än att titulera en enskild lista som vinnare. Det finns aspekter från vardera lista som enligt deltagarna i undersökningen är styrkor, och det finns aspekter som är svagheter. Något som därför blir intressant att undersöka närmre är de motiveringar som ligger bakom svarsresultaten.

5.2 Analys

Resultatanalys enkät

Det var positivt att antalet enkätsvar överträffade förväntan. 14 svar istället för de 10 förmodade. Enkätens svar sorterades först kvantitativt. Hur många svar på vilken lista vid varje fråga för att identifiera summativa trender. Därefter undersöktes fritextsvaren och om möjligt sorterades även de så att lika svar kunde grupperas. Med tanke på en låg svarsfrekvens på fritextsvaret fanns det inte mycket som kunde sorteras till mer användbar kategorisering.

Enkätens svarsresultat visade tydligt att NHL, en lista som specifikt tagits fram för användbarhetsutvärdering av spel, överlag var bättre lämpad som HU-verktyg enligt de deltagare som svarat. Detta resultat är från en undersökning som var gjord av spelutvecklare som inte innehar specifik GUX kompetens och för samma målgrupp. Att enkätdeltagarna motsvarar den tilltänkta målgruppen styrker fyndens relevans och att dessa i något mått är representativa för målgruppen.

De inledande frågorna visade att samtliga deltagare förutom två, ansåg sig vara spelutvecklare. De övriga två titulerade sig Gamer men kryssade trots det i ja på frågan ifall de tidigare studerat heuristik och var bekanta med ämnet. Sex av de som kryssat i spelutvecklare svarade nej på frågan om kunskap om HU. Det fanns inga detaljerade riktlinjer kring de olika svarsalternativen när det gäller bakgrundsfrågorna, så det anses troligt att svaren på de två inledande frågorna reflekterar det som deltagarna primärt identifierar sig som. Här hade tydligare formuleringar såväl som motivering kring syftet med frågorna varit en fördel. Dock upptäcktes ingen skillnad eller mönster i övriga frågors svar med hänsyn till bakgrundsfrågorna.

Med en tydlig övergripande preferens, är det också intressant att förstå de motiveringar som angivits för att få en inblick i hur respondenterna tänkt. En svarande enkätdeltagare uttryckte sig enligt följande.

In general, the more game focused heuristics felt better since they gave examples and were obviously more game-focused, but they weren't as clear or summarizing as the nielsen heuristics which make them a bit confusing. (Anonym enkätdeltagare två).

En annan underströk också sin åsikt kring vikten av spelinriktade formuleringar.

Games are their own thing when it comes to IT. Some things can be generalized down and on the surface the way players interact with the system is the same as many others but how the system reacts is completely different. This means that generalized heuristics will not apply as well as heuristics specifically made for games. (Anonym enkätdeltagare tre).

Dessa två motiveringar kan möjligtvis ge insikt kring hur en spelutvecklare som genomför HU tänker och kanske framför allt på vilka formuleringar de har lättast att knyta an till. Formuleringar som är mer spelspecifika kändes av en del mer naturliga när det gällde att utvärdera spel och mer anpassade för att identifiera användbarhet inom just spel. Alltså spel och inte vilken programvara som helst.

I kontrast, så svarade en av enkätdeltagarna i den avslutande fritextfrågan så här.

There was something I noticed and that was the option often consisted on one shorter answer and one longer answer. That in itself is also a usability test of

convening information in as little time as possible. I did my thesis partly on information architecture, and the fewer words that gives away meaning will always triumph the longer answer. (Anonym enkättagare fyra).

En annan skrev följande koncisa fritextsvar i samma andemening, fast det inte förtydligas ifall kommentaren avsåg en eller båda listors svarsalternativ.

Answers where a bit to long (Anonym enkättagare fem).

Huruvida ett längre svar de facto alltid förlorar mot ett kortare svarsalternativ är en slutsats som inte kan avgöras i det här arbetet. Men det kan ge en hint om fördel och styrka hos Nielsens 10, som i samtliga kategorier hade kortare och mer koncisa formuleringar. Nielsens 10 hade samtidigt stark representation i en tredjedel av svaren från huvudfråga ett och statistiskt relevant svarsandel, om än i minoritet, på punkter från huvudfråga två.

Å andra sidan kan det argumenteras att när ett mer kortfattat språkbruk används blir avsikten mer diffus och att en längre text ger tydligare instruktioner. Inte minst visar resultaten i stort att en majoritet trots allt föredrar NHL som genomgående hade längre formuleringar. Om detta är på grund av att listpunkterna är mer spelorienterat formulerade eller ifall längden med sin något mer detaljrika beskrivning hjälper förtydliga vad som undersöks och utvärderas är inte fastlagt i detta arbete. Däremot konstateras det att enkättagare ändå givit ett uttryck för att det skulle vara en faktor som noterades, men att de kommentarerna står i kontrast till de övergripande resultaten. För att vara tydlig, så finns det inget slutgiltigt stöd i data från enkätsvaren för tanken att argumentet med kortfattat språkbruk har påverkat deltagarna i en tillräckligt nämnvärd utsträckning för att etablera ett mönster.

På det stora hela fanns det inga självklara gemensamma nämnare eller mönster för de motiveringar som enkättagare stödde sig av i sina svar. Eftersom inte ens alla deltagare svarade på fritextsvaren (endast 6 av 14 användbara svar) saknas dessutom underlaget för att identifiera ett eller flera teman som med trovärdighet är representativt för svarsgruppen. Det finns antydningar och tendenser, som tagits upp ovan men de är få i förhållande till urvalspopulationen.

I övrigt stöder enkätsvaren det förväntade resultatet när det gäller jämförelsen av Nielsens 10 och NHL. Det skulle kunna antas styrka en uttrycklig anledningen till att Pinelle, Wong och Stach (2008) ansåg att det fanns ett behov för en spelspecifik heuristiklista som verktyg till spelutvecklare som vill utvärdera sitt spels användbarhet. Samtidigt bör det poängteras att fler listor finns och att resultatet i denna jämförelse gäller mellan Nielsens 10 och NHL, inte övriga listor som är mer eller mindre inriktade mot spel.

Resultatanalys test

Testsvaret från den andra testdeltagaren skickas in till undertecknad som ställde upp dokumentet bredvid sitt eget och jämförde svaren. Markeringar gjordes var det fanns likheter eller olikheter. Både avseende utfall men även kring motiveringar. Anteckningar gjordes för att identifiera aspekter som med fördel kunde förtydligas vid den uppföljande diskussionen. Efter den uppföljande diskussionen förtydligades oklarheter och en slutgiltig sortering kring lik- och olikheter genomfördes.

Testet gav också ett tydligt resultat att NHL är väl lämpad för heuristisk användbarhetsutvärdering enligt testdeltagarnas svar. En sak som gjorde att testets resultat blev tongivande var att det konkret gick igenom ett befintligt spel och använde båda listor som verktyg för att analysera och utvärdera spelets användbarhet. Resultaten är alltså direkt kopplade till en praktisk tillämpning av teorin.

Styrkan att testet är direkt kopplat till ett genomförande kontrast av att det endast var ett spel som utvärderades. Spelets genre, att det var ett flerspelarspel och att båda testdeltagare kände till spelet sedan innan talar för att det skulle kunna producera skeva svar. Vissa antaganden eller känslor som testdeltagarna haft kring spelets styrkor och svagheter kan ha amplifierats i undersökningen, eftersom det är omöjligt att helt frikoppla sig från den typen av subjektiva åsikter, vilket Wilson (2014) poängterar. I vilken utsträckning det har påverkat de faktiska svaren är svårt att bedöma och det märks att en ansträngning har gjorts för att tänka utifrån andra, framförallt nya, spelares perspektiv.

Testdeltagarna var överens kring flera av sina motiveringar. Exempelvis fråga 3, Is the relevant game and/or system information presented to the user? Den ena förklarade:

Den visuella representationen i Dota 2 och har en klar och förstod UI som tydligt säger vart man ska handla, klicka, se, etc.

Medan den andre uttryckte sig så här:

Dota 2 ger bra och relevant information i matchen. Man kan få en överblick genom en minikarta, och specifik information om hp och liknande syns vid karaktären eller fiende/NPC.

Ett annat exempel där både svar och motivering till svar är likartade är från fråga 5 där båda riktar viss kritik till Dota 2. Detta är också ett tillfälle där den uppföljande diskussionen gav klarhet kring motiveringen. Frågan lyder, Is the game consistent and as expected? This can be in regards to input, game mechanics, story, enemy behavior etc. Båda testdeltagare ansåg att NHL hade en tydligare formulering för att fånga essensen av en spelutvärdering. Här tyckte den ena testdeltagaren:

Detta gör Dota 2 bra överlag. Ibland väljer spelet en knasig pathing, tex om jag vill gå rakt åt ett håll så går den runt ett annat objekt istället. Detta händer ibland, inte ofta. Fiender mm betar sig som de ska.

Och den andre testdeltagaren förklarade:

Dota 2 gör det mesta bra, men ibland kan pathing – dvs den valda väg som karaktären går (man klickar bara vart man vill gå, sen löser datorn hur/var vägen fram till målet är) - bli knasig och rent ut sagt fel. Tex om man är i en fight och vill retirera och klickar på en punkt bakåt, så kan ibland pathingen välja en väg som innebär ökad risk (den går till höger istället för vänster om ett träd tex), vilket inte är naturligt givet det önskade resultatet. Detta är säkerligen svårt att åstadkomma, men en brist som kan lämna spelaren frustrerad eftersom man inte detaljstyr sin karaktär.

Här finns alltså en åsikt att Dota 2 har förbättringspotential och att NHL var bäst lämpad att ta fram kritiken. Det inledande textsvaret från den andra deltagaren var lite kort formulerat med ett språkligt bruk som upplevdes lite vagt. När den uppföljande diskussionen skedde blev det tydligare och en justering av det skriftliga svaret genomfördes med godkännande av testdeltagaren.

Ett exempel på nyansskillnad i testresultatets svar var under enkätens fråga fyra. Frågan lyder: How clear are the game information aesthetics? Den handlar om hur tydligt spelets information presenteras utifrån estetik. Där ansåg den ena testdeltagaren att Dota 2 är tydlig, informativ med både fast såväl som temporär information. Förtydligande gjordes att fast information var saker som minikartan, karaktärens UI som visar hälsa och manastatus, vilka

alltid syns. Temporär information kunde vara text som dök upp mitt på skärmen eller ljud vid särskilda händelser som försvann efter några sekunder.

På samma fråga gav den andra testdeltagaren följande kommentar med motivering:

Teamwork är en viktig funktion för att kunna vinna i spelet och det är därför viktigt att information presenteras tydligt för spelaren. Detta görs bra, för vana spelare. Är man ny kanske både till Dota 2 och genren i stort, så kan det bli väldigt mycket och överväldigande. För det är mycket information som presenteras. Det här ett område som skulle kunna utvecklas än mer och bli ännu tydligare.

Här framgår det att även om båda ansåg att Dota 2 gör det bra, så finns en skillnad beroende på vilken typ av spelare som är i åtanke. Är det för en van spelare är betyget högre än för en spelare som är ny till Dota 2. Ännu större skillnad uppmärksammades om spelaren är ny till genren enligt ena testdeltagarens utvärdering. Eftersom en nyckelstrategi för att vara framgångsrik i spelet är samarbete utgör kravet på den presenterade informationens estetik en väsentlig del. Det är viktigt att alla får samma information och möjlighet till orientering och underrättelse vad som händer och var det händer. Ett ypperligt exempel på att trots goda förutsättningar kan det finnas en skillnad i tolkning på grund av subjektivitet (Wilson 2014).

Det fanns två punkter där testdeltagarna hade direkt olika åsikter. Dessa innefattade två väsentligt olika aspekter som påtalar olika delar av användbarhet. Den ena testdeltagaren ansåg att det var viktigt att få med en punkt kring fel och felmeddelanden så att både slutanvändare såväl som spelutvecklare lättare kan identifiera ursprunget för krascher eller systemfel. Spel och mjukvara i stort är inte felfritt och det är omöjligt att täcka in alla möjliga scenarios där något kan gå sönder i den testning som sker innan ett spelsläpp. Felsök är i verkligheten inte bara en del av spelutvecklingsfasen, utan hela spelets livslängd. Ju tidigare spel utvärderas, ju större chans är det att stöta på buggar och problem med mjukvaran som kan åtgärdas innan spelet publiceras för allmänheten.

Den andra testdeltagaren föredrog ett fokus på spelarens möjlighet att själv bestämma vilka delar av spelet som skulle vara tvingande. Det vill säga att den testdeltagaren ansåg att ge spelaren en större frihet att kunna välja eller välja bort ej-spelbart innehåll var mer relevant för ett högt betyg inom användbarhet när spelarens upplevelse låg i fokus.

Det finns antagligen inget rätt svar i frågan, vilken av de olika preferenserna som är viktigast, men det är värt att notera att båda har spelarens spelupplevelse, utifrån användbarhet, som fokus. Trots detta kom de fram till två olika val. Den ena valde att inkludera interaktionen med en del av spelet medan den andre hade fokus på att förenkla felsökning när något i spelet går sönder. Återigen bekräftar analysen att frågeställningen har en högre komplexitet än vad de kvantitativa resultaten återspeglar.

Även testets utfall bekräftar det förväntade resultatet och gör så med ännu tydligare (kvantitativt) besked än enkäten. Båda testdeltagare noterar att språkbruk såväl som detaljerade beskrivningar från NHL underlättade utvärdering av de olika underkategorierna som båda listor gör anspråk på.

Övergripande analys

Båda undersökningar stödjer varandras resultat, oberoende av varandra. Enkäten och testet visar att undersökning kring HU och när Nielsens 10 och NHL jämförs så blir resultaten på det stora hela samma. Denna triangulering styrker resultatens trovärdighet och validitet vilket betyder att det förväntade resultatet ifall undersökningen hade gjorts om, ger en hög sannolikhet till att producera samma svar.

En av de tydliga skillnaderna mellan resultaten från enkät och test är att enkäten gav större fördel till Nielsens 10 än vad testet gjorde. Två av sex frågor, alltså en tredjedel av svaren från första huvudfrågan skiljde i utfall. Den procentsatsen är tillräckligt hög för att vara statistiskt signifikant. Här kan en anledning vara den motivering som gavs av deltagare i fritextfrågan, nämligen att en kortare, mer koncis text skulle vara lättare att ta till sig än en längre text.

Å andra sidan kan det argumenteras att det fortfarande var en majoritet till fördel för NHL som hade längre och mer detaljerade formuleringar. Så att även om det är värt att notera den avvikelse mellan enkät och test, så kvarstår huvudresultatet.

En av fördelarna med testet, som verkligen blev tydlig i summeringen av de inhämtade resultaten, var de diskussioner som ägde rum innan såväl som efter genomförandet. Dessa samtal säkerställde att både undertecknad såväl som den andre testdeltagaren hade samsyn kring definitioner och syfte. Den uppföljande diskussionen gav möjlighet att räta ut frågetecken som fanns och flera förtydliganden tillfördes som förklarade svar och tolkningar.

En nackdel som tidigare nämnts, var att testet endast utfördes mot ett spel. Det är möjligt att med ett bredare underlag av undersökt material, skulle resultaten förändras. Detta är något som hade kunnat genomföras vid ett framtida tillfälle och ytterligare undersökning.

Slutsats

I det stora hela bekräftade både enkät och test varandra. Enkättagarna ansåg att NHL hade formuleringar som resonerade bra när en spelutvecklare utan särskild GUX-kompetens ska utvärdera ett spel utifrån användbarhet. Testdeltagarna ansåg även de att NHL hade spelspecifika formuleringar med beskrivningar och exempel som hjälpte dem att utvärdera användbarhet, inte bara i teorin utan konkret med Dota 2. De övergripande resultaten kan således anses vara triangulerade eftersom båda undersökningar styrker varandra. Båda innehöll kvantitativ och kvalitativ datainsamling så resultatbekräftelsen flyter även över olika metoder vilket ger en fördel för resultatens trovärdighet. Huruvida detta är ett konstaterande som går att extrapolera för andra listor som antingen fokuserar på spel eller är IT-agnostiska är utöver detta arbetes ramar.

En kommentar från en av de avslutande fritextsvaren i testet som genomfördes var att 10 punkter inte behövdes utan att den spelutvecklaren hellre hade velat klippa en av de sista fyra svaren för att istället få fram en niopunkts lista. Detta föranleder vidare eftertanke kring huruvida ens någon av listorna är kompletta eller ifall en övergripande justering är motiverad. Alltså trots att en given lista har fått en stark och positiv respons, kanske den inte är fulländad utan tåls att genomlysas för att eventuellt itereras för fortsatt användning. Trots att det fanns en tydlig trend och tendens i svaren, så var dessa ändå präglade av kommentarer och önskemål.

Det går inte att uteslutande resonera att NHL är en fulländad eller komplett lista och den enda som behövs för en spelutvecklare som vill tillämpa HU i spelutvecklingsfasen. Svartsvarianserna ger förståelse för att ämnet är komplext och eventuellt även i rörelse och därför fortsatt behöver ytterligare forskning. Den ständigt föränderliga marknaden skulle mycket väl kunna tänkas innebära ett behov av anpassning i takt med att allt fler spel släpps. Det senaste decenniet har sett en kraftig tillväxt i den digitala spelvolymen, och det skulle kunna spekuleras att även åren framöver kan tänkas visa en kraftig utveckling.

6 Sammanfattning och diskussion

6.1 Sammanfattning

När en spelutvecklare vill utvärdera ett spel utifrån användbarhet för att identifiera styrkor och svagheter som kan förbättras, är heuristisk användbarhetsutvärdering ett av flera verktyg som kan användas. En fördel med HU är att det snabbt visar på brister och möjligheter. Det finns flera listor som är framtagna med syfte att bistå utredningen, dessa kallas heuristiklistor. I detta arbete har två sådana jämförts, Nielsens 10 och NHL. Den ena är generell och författarna hävdar att den fungerar oavsett vilken typ av IT-programvara som avses. Detta märks genom ett mer generellt språkbruk såväl som mer agnostiska formuleringar. Den senare är framtagen specifikt för utvärdering av spel, vilket märks både på ordval såväl som vilka aspekter som tas upp. Frågan som undersöks i arbetet är vilken av de två listorna som enligt spelutvecklare utan särskild GUX kompetens bäst stödjer en användbarhetsutvärdering i ett spels utvecklingsfas.

För att nå fram till en slutsats genomfördes en enkät såväl som ett test med den relevanta målgruppen. Båda metoder ledde fram till samma övergripande slutsats, att NHL är enligt de deltagande spelutvecklarna den mest lämpliga listan av de två. Dock finns det några avvikelser och intressanta fenomen att ta i beaktning, både när resultaten generaliseras såväl som för framtida forskning.

6.2 Diskussion

Generellt

Undersökningens resultat var i stort mått enligt förväntan. Att både en kvalitativ såväl som kvantitativ metod applicerades var en styrka. Enkäten och testet, som båda innehöll båda metoder, triangulerade och bekräftade varandra, vilket höjer svarens trovärdighet. Denna tillförlitlighet är avgörande för att kunna dra relevanta slutsatser. Trots en tydlig slutsats så fanns det en bredd av kommentarer och tillägg som visar att ämnet är multi facetterat och det är inte en enkel ja och nej fråga. Det finns nyanser, det finns preferenser och det finns sammanhang som skulle kunna påverka utkomsten i en annan studie inom ämnet.

Metod

Att använda två metoder gav möjlighet att data, både kvantitativ såväl som kvalitativ, kunde trianguleras. Eftersom de två undersökningarna var olika gav resultaten möjlighet att svaren kunde stödja eller motsäga varandra. Detta var en av de stora styrkorna med arbetet överlag, att flera metoder användes för att undersöka problemformuleringen.

Samtidigt fanns det vissa brister som blev särskilt tydliga i efterhand när resultaten skulle sorteras och analyseras. Enkätens enskilda trovärdighet var inte lika hög som den hade kunnat vara. Enkäten utgick från formuleringar som spelutvecklare utan särskild GUX-kompetens skulle välja mellan rent fiktivt. De hade ingen möjlighet att konkret testa ett spel utifrån de angivna listorna så som gjordes i testet. Att låta personer tycka till om heuristiker utan att faktiskt prova och tillämpa alternativen kan anses spekulativt och dess konkreta värde ifrågasättas. Huruvida personlig bias, tidigare erfarenhet av spel eller preferens kring genre påverkade svaren och till vilken grad undersöktes aldrig.

Enkätens fritextsvar var få sett till antalet deltagare. Istället för att ha en frivillig fråga hade den, likt de andra frågorna, kunnat var tvingande. Huruvida detta bara hade gett fler svar eller

också fler relevanta svar med användbara insikter kring motiveringar är endast spekulativt. Dock är det något som hade kunnat öka chansen för att få in fler svar till den kvalitativa aspekten av insamlingsmetoden.

Testet som genomfördes hade få deltagare, där båda i stora drag hade samsyn och samma erfarenheter inom HU. Detta homogena underlag kan även det ifrågasättas om det inte hade gynnats av en mer blandad skara deltagare när heuristiker jämförs för att säkerställa att idéer och antaganden får en större bredd. Samtidigt har litteraturen pekat på att ha samsyn och en liknande förståelse för begrepp är en fördel i den här typen av metodanvändning. Det kan också ifrågasättas ifall testdeltagarna, som haft ett spel i fokus när de svarat på frågorna, gav resultat som var generaliserbara för spel i stort. Testresultatens eventuella generaliseringar kan anses begränsande till enbart ett spel eller på sin höjd eventuellt en genre. Att extrapolera breda generaliseringar innehar inte lika hög trovärdighet.

Det är också värt att notera att Dota 2 som släpptes 2013 inte anses vara i den tidiga spelutvecklingsfasen där HU generellt sett anses vara behjälplig (Nielsen 1994, uppdaterad 2024). Trots det identifierades både styrkor och svagheter och visar att spelutveckling aldrig är helt färdigt.

En annan utmaning med testet var den uttalade fokusen att utvärdera de två heuristiklistorna. I genomförandet så tappades en del fokus på den konkreta datainsamlingen utan låg lite mer abstrakt på frågan kring vilken lista bäst lämpar sig för HU. Olika heuristiker utvärderades mellan varandra och konkret i applikation mot Dota 2. Det var i sig positivt. Men testet hade kunnat vara ännu mer rigoröst genom att prova varje heuristik och lista varenda svaghet, styrka och möjlighet. Istället för att komma fram till en slutsats baserat på en mer vag testning, hade detta mer strukturerade tillvägagångssätt i större utsträckningen kunnat bidra till en större mängd data för varje heuristiklistas punkter. En större mängd data på varje enskild punkt hade som support för varje åsikt kunnat höja trovärdigheten ytterligare.

Enkäten och testet i ett större sammanhang

Ingen av deltagarna i vare sig enkät eller test har en särskild spetskompetens inom GUX. Samtliga låg inom nivå 1-3 på Nielsens skala. Resultaten visar samtidigt att en heuristisk användbarhetsutvärdering är möjlig att genomföra med en lägre UX-mognad. Detta styrker HU som verktyg för spelutvecklare och bekräftar dess nytta som hjälpmedel att identifiera styrkor och svagheter inom användbarhet i spelutvecklingsfasen.

Att HU är en snabb och effektiv metod för att ge återkoppling kring ett spels kvalitet utifrån användbarhet etablerar Omar och Jaafar (2010). I det genomförda testet blev det tydligt att det förelåg en konsensus att oavsett vilken lista som avsågs så hittades olika aspekter som antingen ansågs vara styrkor eller svagheter för Dota 2. Ibland argumenterade de två testdeltagarna lite olika men det intressanta blev diskussionen som följde. Att genomföra en HU kan därför dra nytta av ett gemensamt samtal innan resultat slutgiltigt sammanställs och forskare drar slutsatser. I spelutvecklingsfasen där feedback eftersöks med syfte att förbättra sitt spel, visar det sig att utvärdera kollektivt – alltså i samråd med andra spelutvecklare kan ge insikter som än tydligare belyser utvecklingsmöjligheter.

Strååt, Rutz och Johansson (2014) avslutade sin text med förslag på förändring av befintliga heuristiklistor. Även i denna undersökning hintades möjligheten att förkorta antalet, eller förenkla formuleringar. Som biresultat bekräftar därmed denna studie att ämnet inte kan anses vara färdigt eller definitivt. Med nya forskningsansatser och olika perspektiv, förändras behoven och det efterfrågade verktyget kan och kanske till och med bör anpassas. Som tidigare

nämnt växer branschen i snabb takt och ett resultat av en större volym spel, kan tänkas vara att nya genrer skapas. Det kan antas att fortsättningsvis kommer nya sätt att kombinera existerande genrer utforskas och att listor därefter behöver anpassas för nya behov som ännu inte är etablerade. Detta bekräftar även Liao och Shen (2012) som kom fram till slutsatsen att en heuristiklista för generell programvara inte svarar upp mot de krav och förutsättningar som den specifika genren inlärningsspel medför.

En observation från de mer tongivande tidigare granskade artiklarna är att deras underlag av forskare och testare är lågt sett till antal. I det genomgångna materialet fanns det aldrig fler än fem som deltog i studierna. En styrka som denna undersökning har är att det empiriska underlaget, sett till antal spelutvecklare som deltog i undersökningen är högre än vanligt. Detta gäller även om det kan verka lågt med sina totalt 16 deltagare, där 14 kommer från enkät och två från test. Enkäter lämpar sig väl när ett större antal deltar, eftersom de kvantifierade resultaten får en hög representation gentemot den tilltänkta målgruppen. Heuristisk användbarhetsutvärdering är trots allt ett nischat ämne och i det här arbetet med en målgrupp som är spelutvecklare utan uttalad GUX-kompetens. Det krymper både populationen och därmed den andel av population som kan förväntas delta i en studie.

Trots att antalet som faktiskt deltog i undersökningen var högre än i många andra tidigare undersökta texter, återstår ett önskemål att se ännu fler svarande för att skapa en ännu högre generaliserbarhet kring resultaten.

Avslutningsvis

Eliasson (2019) undersöker reliabilitet och för att den ska vara hög, ska resultat kunna upprepas. I samma sammanhang är det troligt att liknande resultat hade etablerats om enkät och test gjorts om igen. Men vad angår frågan om HU i stort och om det finns en lista som är bäst lämpad blir det ett för stort ställningstagande att göra eftersom det finns flera parametrar, vinklar och perspektiv att undersöka. Det finns säkerligen olika omständigheter som talar för eller emot olika listors effektivitet i användbarhetssammanhang.

6.3 Samhälleliga och etiska aspekter

I takt med den växande marknaden växer också behoven för spelutvecklare att lyckas tränga genom bruset och nå sin tänkta målgrupp. Som presenteras i bakgrunden är UX en av de avgörande faktorerna för framgång. Detta kan bli en utmaning för spelutvecklare som har en lägre UX-mognad. Denna undersökningens resultat visar att det finns verktyg som är väl lämpade för spelutvecklare utan särskild GUX-kompetens. Verktyg, specifikt den heuristiska användbarhetsutvärderingen är ett lättillgängligt, billigt verktyg som alla med fördel kan använda, oavsett UX-mognad. Därför blir värdet av HU träffande för samhället. Det är inte en metod som är förbehållen specialister allena.

För spelmarknaden och forskning överlag är det därför viktigt att forskningsresultat som denna undersökningen och andra identifierar når ut till den generella populationen. I det här fallet särskilt till spelutvecklare så att de kan ta del av och implementera kunskap till praktisk omsättning. Idag finns det inget tydligt sätt som den här typen av nya rön delas med spelbranschen och risken är att informationen stannar kvar inom den mer teoretiska forskningssfären och således inte kommer till praktisk gagn. Ifall forskning kan ge feedback till spelutveckling, och att den tas emot, kan spelutvecklingen omsätta detta i praktiken och forskningen får nytt innehåll att undersöka. Det blir en ömsesidig framåtdrift för hela branschen.

Ett annat intressant fynd var vikten av en inledande och uppföljande diskussion och det värde det tillförde analysen av testet. Detta kvalitativa tillvägagångssätt skulle mycket väl kunna tillämpas med framgång inom andra domäner. Även annan ämnesforskning kan dra fördel av att söka djupare förståelse för svar och kring resonemang som inte fångas upp av de mer summativa svaren som ges i till exempel flervalssformat.

6.4 Framtida arbete

Det hade varit intressant att undersöka problemformuleringen på ännu bredare front för att verkligen och med god akademisk trovärdighet kunna tillämpa ännu bredare generaliseringar inom ämnet. Med detta avses att införskaffa ett större empiriskt underlag, även om detta arbete redan är mer ambitiöst i sitt utformande än flera andra rådande artiklar. Då ämnet är komplext och viss förståelse för heuristisk användbarhetsutvärdering fordras när ett test genomförs antas det vara ett större projekt som antagligen skulle sträcka sig över en längre tidshorisont. Genomförandet vore högst intressant och skulle tillföra ett ännu större sammanhang till fynden såväl som högre generaliserbarhet, på grund av användandet av en större population.

Det hade även varit intressant att utforska huruvida olika genrer förhåller sig till olika heuristiklistor. Ifall varje genre behöver specialanpassade listor gör det HU till ett trubbigare verktyg för en spelutvecklare som inte besitter spetskompetens inom området, än ifall en lista kan anses vara tillräcklig oavsett genre.

Slutligen så är en av de större identifierade bristerna, överföringen av ny kunskap till den avsedda målgruppen. Det är en utmaning som all typ av forskning står inför, och spelforskning är inte ett undantag. Att undersöka möjligheten kring kunskapsdelning mellan forskning och praktiskt verksamhet hade varit ett intressant fördjupningsarbete. Hur det kan göras, vilka parter behöver vara inblandade med mera. När forskning når den praktiska branschen kan resultat omsättas i snabbare takt. Ifall information kan nå spelutvecklare med låg UX-mognad som inte redan är bekanta med forskningsområdet och inte naturligt vet var de ska leta nya rön uppstår nya möjligheter för utveckling både teoretiskt såväl som praktiskt i en snabbt växande bransch.

Referenser

- Agaté (2026). *Why does game testing matter*. <https://agate.id/why-does-game-testing-matter/> [2026-03-07]
- Buis, E.E.G., Ashby, S.S.R. and Kouwenberg, K.K.P.A. (2023) 'Increasing the UX maturity level of clients: A study of best practices in an agile environment', *Information and Software Technology*, 154. doi:10.1016/j.infsof.2022.107086.
- Borg, E. & Westerlund, J. (2014) *Statistik för beteendevetare*. (3:e uppl.) Kina: Liber.
- Chapman, L., Plewes, S. (2014). A UX Maturity Model: Effective Introduction of UX into Organizations. In: Marcus, A. (eds) *Design, User Experience, and Usability. User Experience Design Practice. DUXU 2014. Lecture Notes in Computer Science, vol 8520*. Springer, Cham. https://doi.org/10.1007/978-3-319-07638-6_2
- Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In CHI'04 extended abstracts on Human factors in computing system (pp. 1509-1512). ACM. doi:10.1145/985921.986102
- Desurvire, H., & El-Nasr, M. S. (2013). Methods for Game User Research: Studying Player Behavior to Enhance Game Design, in IEEE Computer Graphics and Applications, vol. 33, no. 4, pp. 82-87, July-Aug. 2013, doi: 10.1109/MCG.2013.61
- Desurvire, H., & Wiberg, C. (2009). Game usability heuristics (play) for evaluating and designing better games: The next iteration. In Online Communities and Social Computing (pp. 557-566). Springer Berlin Heidelberg. doi:10.1007/978-3-642-02774-1_60
- Eliasson, A. (2019). *Kvantitativ metod från början*. (4:e uppl.) Lund: Studentlitteratur.
- Gibbons, S., Moran, K., Pernice, K., Whintont, K., (2021, updated 2024). Nielsen Norman Group. Retrieved 2026-03-06, from <https://www.nngroup.com/articles/ux-maturity-model/>
- Gül, Yavuz. (2023). A theoretical perspective on survey method from quantitative research methods. *Universum*. 4. 64-68. 10.32743/UniPsy.2023.106.4.15254.
- Liao, Y.H. & Shen, C.-Y. (2012). Heuristic Evaluation of Digital Game Based Learning: A Case Study, *2012 IEEE Fourth International Conference On Digital Game And Intelligent Toy Enhanced Learning, Digital Game and Intelligent Toy Enhanced Learning (DIGITEL), 2012 IEEE Fourth International Conference on*, pp. 192-196. doi:10.1109/DIGITEL.2012.54
- Malone, T. W., (1982) *Heuristics for designing enjoyable user interfaces: Lessons from computer games*. Norwood, NJ: Ablex Publishing Corporation.
- Nielsen, J. (1994 · Updated 2024). Nielsen Norman Group. Retrieved 2026-02-05, from <https://www.nngroup.com/articles/ten-usability-heuristics>
- Nielsen, J. (1994). Nielsen Norman Group. Retrieved 2026-02-18, from <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>
- Omar, H.M.@ & Jaafar, A. (2010). Heuristics evaluation in computer games, *2010 International Conference on Information Retrieval & Knowledge Management (CAMP), Information Retrieval & Knowledge Management, (CAMP), 2010 International Conference on*, pp. 188-193. doi:10.1109/INFRKM.2010.5466921
- Richardson, M., Carter, F., Levordashka, A., Voinescu, A., Gilchrist, I.D. & Stanton Fraser, D. (2025). Methods used to study audience experience of screen-based media: A scoping review,

- Psychology of Aesthetics, Creativity, and the Arts*. doi:<https://doi.org/10.1037/aca0000686>
- SteamDB (2026). Releases. <https://steamdb.info/stats/releases/>. [2026-03-07]
- Strååt, B., Rutz, F., & Johansson, M. (2014). Does Game Quality Reflect Heuristic Evaluation?: Heuristic Evaluation of Games in Different Quality Strata. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, 6(4), 45-58. <https://doi-org.libraryproxy.his.se/10.4018/ijgcms.2014100104>
- Syddansk Universitet (2026). Games User Research. https://www.sdu.dk/en/forskning/sdumetaverselab/research_areas/games_user_research [2026-03-17].
- Valve, (2013). *Dota 2*, Washington: Valve L.L.C.
- Vetenskapsrådet, (2024). *God forskningssed 2024*. <https://www.vr.se/analys/rapporter/vara-rapporter/2024-10-02-god-forskningssed-2024.html> [2026-03-09]
- Wilson, C., (2014). *User Interface Inspection Methods, a User-Centered Design Method*. Elsevier.

Appendix A – Enkät- & Testfrågor



Heuristic Usability
Evaluation - b23gabrr

Heuristic Usability Evaluation

This survey is a part of a bachelor thesis and your answers will be used to make some evaluations and comparisons for the study. No personal information is recorded or kept in any way, the survey is anonymous. The answers provided will only be used for the purpose of this assignment. Should you have any questions or concerns, please reach out to Gabriel via email: b23gabma@student.his.se. Thanks for your time and participating!

Heuristic Usability Evaluation (HE) for games is about evaluating a game in regards to the user-game interaction. That means looking at aspects that affect the end user experience from a usability point of view. This includes visual representation, flexibility in settings, feedback on player controls and more. There are many aspects to explore!

To help game developers more easily narrow down their search for areas of improvement, there are different HE lists that each aim to point to areas or questions that will help pinpoint weaknesses. Two of the most popular lists, "Nielsen's 10 Usability Heuristics" (N10) and "Net Heuristic List" (NHL) are what will be compared in this survey. Both lists consist of 10 evaluation points. N10 list is IT-agnostic, while NHL is made specifically for the games industry. The aim of the study is to figure out which is more suitable for game developers, or if a blend of the two is most desired.

After some general information questions, the survey consists of three parts. The first is a comparison between points from each respective list, where they overlap. The second part is an evaluation of the points that are different. Lastly, you will be given an opportunity to add any thoughts you might have (optional),

Part 1: General Questions

1. Are you a...

Mark only one oval.

- Gamer
- Game developer (can also be a gamer)
- Game researcher (can also be a gamer)

2. Have you studied/used/learned about Heuristic Usability Evaluation other than from the introduction to this survey?

Mark only one oval.

- Yes
- No

Part 2: Similarities

On each question below, please choose one (1) statement that you feel best fits looking for a problem within the identified usability area.

3. Is the relevant game and/or system information presented to the user?

Mark only one oval.

- Visibility of System Status: designs should keep users informed about what is going on, through appropriate, timely feedback.
- Provide unobstructed views that are appropriate for the user's current actions. Most games provide users with a visual representation (i.e. a "view") of the virtual location that the user is currently occupying. The game should provide views that allow the user to have a clear, unobstructed view of the area, and of all visual information that is tied to the location. Views should also be designed so that they are appropriate for the activity that the user is carrying out in the game. For example, in a 3D game different camera angles may be needed for jumping sequences, for fighting sequences, and for small and large rooms.

4. How clear are the game information aesthetics?

Mark only one oval.

- Provide visual representations that are easy to interpret and that minimize the need for micromanagement. Visual representations, such as radar views, maps, icons, and avatars, are frequently used to convey information about the current status of the game. Visual representations should be designed so that they are easy to interpret, so that they minimize clutter and occlusion, and so that users can differentiate important elements from irrelevant elements. Further, representations should be designed to minimize the need for micromanagement, where users are forced to interactively search through the representation to find needed elements.
- Interfaces should not contain information that is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility. This heuristic doesn't mean you have to use a flat design — it's about making sure you're keeping the content and visual design focused on the essentials. Ensure that the visual elements of the interface support the user's primary goals.

5. Is the game consistent and as expected? This can be in regards to input, game mechanics, story, enemy behavior etc.

Mark only one oval.

- Provide consistent responses to the user's actions. Games should respond to users' actions in a predictable manner. Basic mechanics, such as hit detection, game physics, character movement, and enemy behavior, should all be appropriate for the situation that the user is facing. Games should also provide consistent input mappings so that users' actions always lead to the expected outcome.
- Consistency and standards. Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions.

6. Does the game offer flexibility for the users preferences?

Mark only one oval.

- Flexibility and efficiency of use. Shortcuts — hidden from novice users — may speed up the interaction for the expert user so that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
- Provide intuitive and customizable input mappings. Most games require rapid responses from the user, so input mapping must be designed so that users can issue commands quickly and accurately. Mappings should be easy to learn and should be intuitive to use, leveraging spatial relationships (the up button is above the down button, etc.) and other natural pairings. They should also adopt input conventions that are common in other similar games (e.g. many first-person shooters and real-time strategy games use similar input schemes). Games should allow users to remap the input settings, should support standard input devices (e.g. mouse, keyboard, gamepad), and should provide shortcuts for expert players.

7. Does the player have to remember everything or is important information readily available?

Mark only one oval.

- Recognition rather than recall. Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g. field labels or menu items) should be visible or easily retrievable when needed.
- Provide users with information on game status. Users make decisions based on their knowledge of the current status of the game. Examples of common types of information that users need to track include the current status of their character (such as their health, armor status, and location in the game world), objectives, teammates, and enemies. Users should be provided with enough information to allow them to make proper decisions while playing the game.

8. Does the game provide training and help?

Mark only one oval.

- Provide instructions, training, and help. Many games are complex and have steep learning curves, making it challenging for users to gain mastery of game fundamentals. Users should have access to complete documentation on the game, including how to interpret Visual representations and how to interact with game elements. When appropriate, users should be provided with interactive training to coach them through the basics. Further, default or recommended choices should be provided when users have to make decisions in complex games, and additional help should be accessible within the application.
- Help and documentation. It's best if the system doesn't need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks.

Part 3: Differences

From the list below, please choose four (4) statements you believe are most useful in helping explore potential usability weaknesses in a game.

9. Which four (4) of the following aspects would you find most helpful when evaluating the usability of a game?
- Match between the System and the real world. The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order.
 - User Control and Freedom. Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process.
 - Error Prevention. Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions, or check for them and present users with a confirmation option before they commit to the action.
 - Help Users Recognize, Diagnose and Recover from Errors. Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution.
 - Provide controls that are easy to manage, and that have an appropriate level of sensitivity and responsiveness. Many games allow users to control avatars such as characters or vehicles. Controls for avatars should be designed so that they are easy for the user to manage, i.e. they are not too sensitive or unresponsive. When controls are based on real world interactions, such as steering a car or using a control stick in an airplane, the game should respond to input in a way that mirrors the real world. Further, games should respond to controls in a timeframe that is suitable for gameplay requirements.
 - Allow users to skip non-playable and frequently repeated content. Many games include lengthy audio and video sequences, or other types of non-interactive content. Games should allow users to skip non-playable content so that it does not interfere with gameplay.
 - Provide predictable and reasonable behavior for computer controlled units. In many games, the computer helps the user control the movement of their character, of a small group of teammates, or of a large number of units. Computer controlled units should behave in a predictable fashion, and users should not be forced to issue extra commands to correct faulty artificial intelligence. The game should control units so that pathfinding and other behaviors are reasonable for in-game situations.
 - Allow users to customize video and audio settings, difficulty and game speed. The video and audio settings, and the difficulty and game speed levels seen in games are not appropriate for all users. The system should allow people to customize a range of settings so that the game accommodates their individual needs.

Part 4: Additional Input

Are you satisfied with your previous choices or is there something missing, that you would like to add? Any tie-breakers or other comments that you wish to provide?

10. Thoughts?