

PROSJEKTRAPPORT

INF2260 - H12



PUZZLED

Audun Karlsrud Larsen, Håvard Lundberg, Rebekka Soma, Martine Birketvedt Eklund og Denis Ligonja

25.11.12

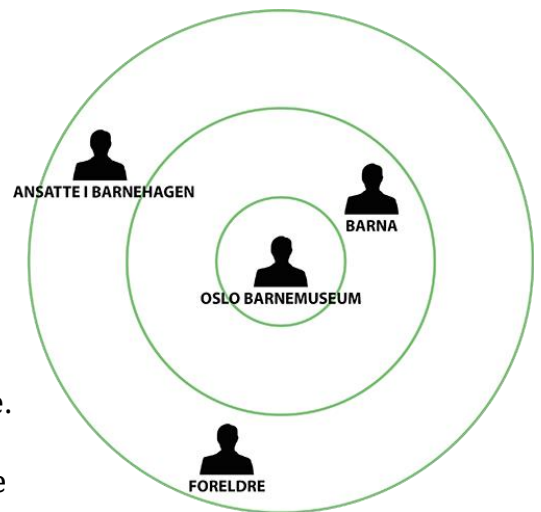
Innhold

1. Introduksjon	3
1.1 Om prosjektgruppen.....	3
1.2 Oppgavens struktur	4
1.3 Tidsplan for prosjektet	4
2. Designprosessen.....	5
2.1 Beskrivelse av prototype og prosess.....	5
2.2 Valg av målgruppe og design i henhold til denne	7
3. Metoder for evaluering.....	10
3.1 Metodevalg.....	10
3.2 Design av eksperimentet	12
3.2.1 Beskrivelse av eksperimentet.....	12
3.4 Utvalg.....	13
3.5 Oppsett.....	13
3.5.1 Teknisk opptaksutstyr.....	14
3.6 Etske aspekter ved eksperimentet.....	15
4. Resultater og analyse	15
4.1 Måling av læringsutbytte	15
4.1.1 Undersøkelse av nullhypotesen	16
4.1.2 Forskjell mellom prototypene	18
4.1.3 Regresjonsanalyse	20
4.2 Kvalitativ observasjon av atferd.....	20
5. Begrensninger ved prosjektet.....	22
5.1 Reliabilitet	22
5.1.1 Måleinstrumenter	22
5.2 Oppsett for gjennomføring av eksperimentet.....	22
5.2.1 Deltakere.....	23
5.2.2 Testlederens atferd	24
5.2.3 Omgivelsene for eksperimentet	24
5.3 Drøfting av begrensninger ved forskningsdesignet	24
5.4 Begrensninger ved prototypen	26
6. Oppsummering.....	26
7. Referanser	28
8. Vedlegg.....	30

1. Introduksjon

Vår oppdragsgiver, stiftelsen Oslo Barnemuseum, har som visjon å gi barn «'hands-on, minds on' play experiences" som hjelper dem å utvikle kreativ tenking, selvtillit og forståelse av verden (Oslo Barnemuseum, i.d.).

Gjennom prosjektkonseptet “Play power” har vi i samråd med stiftelsens grunnlegger, Katie Coughlin, diskutert – og *lekt* – oss fram til et konsept for en testbar prototype vi håper kan innfri deres visjoner. Vi anser Oslo Barnemuseum for å være vår hovedinteressent (*key stakeholder*). Vi har utviklet i henhold til deres krav, men designet med tanke på primærbrukerne; barna (se figur 1.1). Andre interessenter er de ansatte i barnehagen og foreldrene. Disse ville det vært aktuelt å involvere i større grad i neste fase av utviklingen. Eksempelvis kunne det være relevant å teste hvorvidt de ansatte lett kan håndtere prototypen. Vi utviklet prototypene med hensyn til tilbakemeldingene vi fikk fra målgruppen underveis; barn i femårsalderen. Bakgrunnen for valg av målgruppe kommer vi tilbake til i [seksjon 2.2](#). Vi har testet ut prototypen i barnehager, hvor vi primært har brukt barna som *testers*, og vil anta at det er i tilsvarende brukskontekster den vil passe best for senere bruk. Det kan også hende det blir aktuelt for Barnemuseet å benytte den i utstillingssammenheng, der barna kan interagere med den.



Figur 1.1 – Onion model av stakeholders

Oslo Barnemuseum er opptatt av læring gjennom lek og brobygging. Dette, kombinert med et ønske fra Coughlin om å ta utgangspunkt i noe «gammelt» og fornye det, inspirerte oss til å lage et interaktivt geografipuslespill, hvor tanken var at barna skulle lære mer om verden. Coughlin ønsket at prototypen skulle stimulere flere sanser og at den skulle være tilgjengelig også for barn som ikke har lært å lese. Samarbeid var et annet ønskelig aspekt ved prototypen.

Da design-briefen måtte leveres på et tidspunkt hvor vi ennå ikke hadde hatt noe individuelt møte med Coughlin, var denne temmelig generell og baserte seg i hovedsak på barnemuseets nevnte visjoner. Vi hadde ingen konkret idé, men visste at vi ville fokusere på ett av følgende temaer: stimuli av kreativitet, brobygging mellom ulike kulturer og religioner og legge til rette for læring gjennom aktivitet. Vi ønsket å utvikle konseptet vårt gjennom diskusjon med Coughlin, og det er gjennom samtaler med henne at vi har utformet overnevnte krav.

1.1 Om prosjektgruppen

Vår gruppe består av Audun Karlsrud Larsen, Denis Ligonja, Håvard Lundberg, Martine Birketvedt Eklund og Rebekka Soma.

Til sammen har vi en del erfaring med webdesign og utvikling av websider med HTML, CSS og PHP, Adobe-pakken, flash-animasjon, film og foto (klipping). Alle har tatt emnene INF1500 og INF1510, og har erfaring med Arduino-teknologien. Noen av

oss har jobbet litt med tegning og andre kreative formgivningsteknikker. Gruppen har også erfaring fra organisasjonsarbeid. Enkelte har mye intervju - og skrivetrening. Vi har også litt erfaring fra fag som psykologi, medievitenskap og entreprenørskap. Alle har gode kommunikasjonssevner og er gode på samarbeid. I tillegg er alle litt over gjennomsnittlig engasjert og nysgjerrige.

1.2 Oppgavens struktur

Denne rapporten fokuserer på hvordan vi har jobbet metodisk med dette prosjektet. Da vi under presentasjonen 22. oktober presenterte arbeidet med de ulike prototypene, hvordan disse fungerer og hvilke designprinsipp som har lagt føringer for disse, har vi fått beskjed om at dette ikke skulle omtales inngående i rapporten. I [seksjon 2](#) tar vi for oss Barnemuseets visjoner og krav og hvordan vi har utviklet i henhold til disse. Vi går også gjennom bakgrunnen for valg av målgruppe og hvilke implikasjoner dette har hatt for utviklingen av prototypene. I [seksjon 3](#) ser vi på metoder for evaluering og våre valg av metoder. [Seksjon 4](#) omhandler resultater fra våre undersøkelser og analyse av dette. I [seksjon 5](#) drøfter vi begrensninger, svakheter og styrker ved vårt prosjekt. [Seksjon 6](#) ser på mulige endringer som kunne vært gjort på prototypen basert på de erfaringene vi fikk. I [seksjon 7](#) oppsummerer vi arbeidet med dette prosjektet og de viktigste erfaringene vi har gjort oss. Vi har valgt å organisere rapporten på denne måten fordi vi mener det gir en god oppbygning og flyt.

1.3 Tidsplan for prosjektet

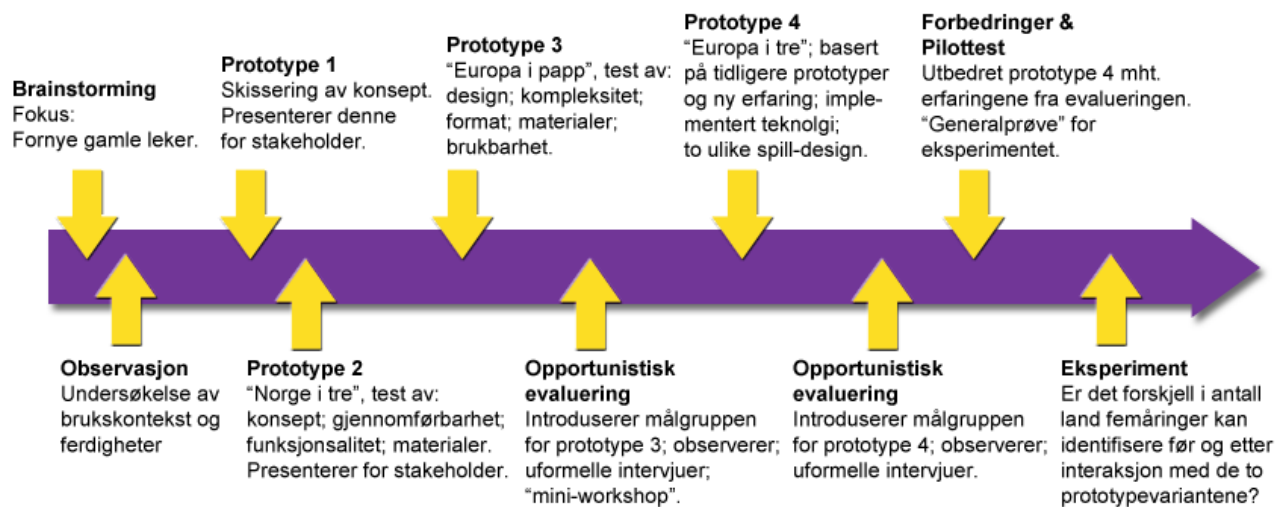
Ettersom vi ikke var helt sikre på omfanget av prosjektet, utviklet vi en foreløpig tidsplan for prosjektet tidlig i prosessen. Dette ble gjort i forbindelse med innlevering av design-briefen, på et tidspunkt vi ennå ikke hadde hatt noe møte med oppdragsgiver. Planen har blitt gjennomgående revidert etter hvert som vi fikk oversikt over hva som måtte gjøres. Det endelige plandokumentet er lagt ved som vedlegg B.

Uke	Viktige datoer	Aktivitet	Andre obliger
36		T1 - Idémyldring, brief, komme med spørsmål til møtet med Katie. Utkast til prosjektplan.	Oblig 1 - INF2260; INF3331
37		T2 - Spesifisere mål og idé. Prototyping.	Oblig 1 - INF1300; INF3331
38		T3 - Utvikle hypotese. Brukertestning. Utforme undersøkelser.	Oblig 2 - INF2260; INF3331
39		T4 - Gjennomføre undersøkelser, analysere data.	INF3331
40	Høstferie barneskole	T5 - Fortsette analyse.	Oblig 3 - INF2260; Oblig 2 - INF1300; INF3331

Figur 1.2 – Første tidsplan for prosjektet

2. Designprosessen

2.1 Beskrivelse av prototype og prosess



Figur 2.1 - Fra designprosess til eksperiment

Prototypene våre er utviklet med tanke på Oslo Barnemuseums visjon om læring gjennom lek. Ifølge Woolfolk (2004) oppstår læring "når en erfaring forårsaker en relativt permanent forandring i et individs kunnskap eller atferd." Vi utarbeidet et forskningsspørsmål som har lagt grunnlaget for arbeidet i dette prosjektet: *Kan bruk av gamification bidra til å øke geografikunnskapene til barn? Gamification er en veletablert teknikk for å øke interessen for å gjennomføre en oppgave, ved å tilføre den spillaktige trekk (Danforth, 2011). Med dette utgangspunktet, samt stiftelsens ønske om at vi skulle modernisere et analogt leketøy, konkretiserte vi visjonene våre og utviklet et interaktivt geografipuslespill.*

Mulighetene som ligger i utviklingen av applikasjoner til bærbare enheter gjorde det aktuelt å vurdere dette som plattform, men flere undersøkelser viser til at håndfaste, interaktive systemer (*tangibles*) la godt til rette for både læring og samarbeid (Marshall et al., 2003; Futurelab, 2004; Xu et al., 2007; Marshall, 2007; Horn et al., 2011). Verhaegh, Fontijn, & Jacobs (2008) peker i sin studie på at håndfaste puslespill er mer tilgjengelige for barn enn digitale varianter: Barna skjønner raskere hva som forventes og hvordan det skal oppnås, mens den digitale teknologien tidvis kommer i veien for læringen, ettersom flere av barna har problemer med å navigere grensesnittet. Artikkelen viser også til at det håndfaste systemet legger bedre til rette for samarbeid mellom barna.

At vi valgte nettopp puslespill blant leketøyene vi kunne modernisere var ikke tilfeldig, ettersom vi alle husket pusling som en utfordrende, men tilfredstillende aktivitet. Puslespill sies faktisk å være «*the preferred manipulative material, second only to Legos® in daily use*» blant unge barn (Maldonado, 2006). Ifølge Maldonado er det svært mange positive aspekter tilknyttet pusling: Barna utvikler fysiske ferdigheter som gripeteknikk, finmotorikk og øye-hånd-koordinasjon, samt kognitive ferdigheter som problemløsning og matematikkforståelse relatert til sortering, sammenligning, perspektiv og romfølelse. Ifølge Marshall et al. (2003) kan det fysiske

aspektet ved *tangibles* legge bedre til rette for «*greater insight, learning, and problem-solving than with other kinds of interfaces*».

Puslespillkonseptet passer godt inn barnemuseets krav om at prototypen bør fremme samarbeid og kommunikasjon, ettersom pusling også utvikler barnas sosiale og emosjonelle ferdigheter. Nyere læringsteorier understreker også nytteverdien av såkalt *horisontale læringsrelasjoner* der likeverdighet i relasjonene, gjennom konkurranse eller samarbeid, kan tilføre læring som formidling fra en voksen eller lærer ikke kan gi (Tetzchner, 2001, s.303). Damon og Phelps (1989) trekker blant annet frem *samarbeidslæring* i grupper av barn der de fungerer som støtte og inspirasjon for hverandre. Ifølge Tetzchner (2001) viser en rekke studier at barn som samarbeider har større mulighet for å løse oppgaver som de ikke hadde kunnskap om på forhånd.

Vi håpet at vi gjennom en geografibasert puslespillprototype kunne lære barna mer om verden, og at vi kunne stimulere læringsprosessen ved å tilføre spillet et interaktivt aspekt. Det finnes mye litteratur som bekrefter dette ståstedet, men vi vil trekke fram Stoney og Oliver ettersom de godt beskriver effekten vi håper interaktiviteten vil skape:

«Being actively engaged in a learning activity has repeatedly been shown to be beneficial for learning. Engagement also comprises cognitive engagement which increases attention to the activity, concentration and promotes “useful” learning» (Stoney & Oliver 1999 gjengitt i Price, Rogers, Scaife, Stanton, & Neale (2003)).

Vi ønsket å utforme prototypen slik at den ga økt kunnskap og forståelse om andre land og kulturer gjennom interaktiv tilbakemelding. Oppdragsgiver mente imidlertid slik informasjon kanskje ville være i overkant avansert å prosessere for målgruppen. Vi valgte derfor å utforme prototypen i henhold til dette, da vi anser oppdragsgiver som en autoritet på dette feltet. En del litteratur vi har lest, deriblant Tetzchner (2001), tyder imidlertid på at denne type informasjon kan hjelpe barna å knytte assosiasjoner til landene, og at dette kan lette læringsprosessen og gjøre det lettere å huske informasjonen. Dette er et aspekt man eventuelt måtte vurdere ved videre utvikling.

Prototype 4 (se figur 2.2) er en fullt funksjonell, high-fidelity prototype utformet i henhold til målgruppen vår (se seksjon 3.2). Som utdypet under presentasjonen 22. oktober har vi designet denne i henhold til designprinsippene vi lærte i INF1500, med særlig fokus på *feedback* (lys, lyd), *constraints* (begrensende omriss), *consistency* («gammel kunnskap, ny kontekst») og *affordance* (knotter gir indirekte hint om bruk). Vi utviklet prototypen med tanke på å stimulere flere sanser; både syn, hørsel og taktilitet.



Figur 2.2 – Prototype 4 i bruk

Brukeropplevelsesmål vi har siktet mot å oppfylle er hovedsakelig *cognitively stimulating, engaging, challenging, fun* og *enhancing sociability*. Siden konseptet vårt fokuserer på læring har vi utviklet to ulike spill-design, og det er disse vi ønsker å evaluere. Nærmere bestemt ønsker vi, ved hjelp av eksperimentell design, å undersøke om det er noen forskjell i antall land femåringer kan identifisere før og etter interaksjon med de to prototypevariantene, og hvorvidt en av de to utmerker seg. De to variantene er som følger:

Prototypevariant 1: Oppgavebasert, interaktivt geografipuslespill. Spillet er programmert slik at det selv gir barna oppgaver av typen “Finn Norge”, hvorpå barnet får feedback når landet er lagt riktig og en ny oppgave spilles av. I dette eksempelet vil prototypen gi tilbakemelding dersom en annen brikke enn Norge legges ned slik som «Dette er Danmark, prøv igjen». Videre vil det etter en satt tid (20 sekunder) spilles av et hint om hvilken farge landet har. Dette er for å forhindre at deltakerne blir oppgitt eller demotivert når de gjør feil. Når alle landene er plassert vil prototypen gi feedback på dette.

Prototypevariant 2: Interaktivt geografipuslespill, ikke oppgavebasert. Barna leker fritt med puslespillet, som gir feedback av typen “Dette er Norge!” når brikken er plassert riktig. Denne type feedback kommer hver gang en brikke plasseres i sin korrekte posisjon, og spillet har ingen definerte oppgaver.

2.2 Valg av målgruppe og design i henhold til denne

Oslo Barnemuseum ga oss ingen spesifikke føringer på hvilken aldersgruppe vi skulle designe for. Det er vanskelig å designe for en uspesifisert målgruppe, ettersom behov og ferdigheter kan variere voldsomt, særlig mellom ulike alderstrinn (Rogers, Sharp, & Preece, 2011). Vi har derfor selv valgt å fokusere på femåringer.

Det er både teknologiske og praktiske forhold som ligger til grunn for vårt valg av målgruppe. Vi valgte å benytte oss av Arduino-plattformen fordi dette er en teknologi vi har kjennskap til fra før av. Denne har imidlertid noen iboende begrensninger, blant annet antall “pinner” som i vårt tilfelle begrenset antall brikker. Det finnes løsninger for å overstige dette antallet, for eksempel kan man sammenkoble flere Arduino-brett; men førstelektor Roger Antonsen ved Institutt for informatikk frarådet oss dette, da det gjør teknologien mindre stabil.

Etttersom Arduino-kortet vårt begrenset oss til å bruke maksimalt tretten brikker måtte barna være såpass unge at tretten brikker ikke ville være veldig lett. Er oppgaven for enkel, endrer den ikke barnas mestringszone (Tetzchner, 2001). Fordi mestring bidrar til læring gjennom å øke barnas selvoppfatning og motivasjon (Haug, 2010), var det derfor viktig at vi ikke siktet oss inn mot et for høyt alderstrinn.

Brikkeantallet vi har til rådighet passer, ifølge Maldonado, generelt for fireåringer og overstiger såvidt øvre «grense» for treåringer. En femåring pusler, ifølge Maldonado, spill med 18-35 brikker, med fokus på mestringsfølelse, mens yngre barn burde pusle med et mindre antall brikker. Vi antok imidlertid at det ville bli vanskelig å evaluere prototypen med så unge barn. Disse aldersgruppene befinner seg i det Piaget kaller

den prekonseptuelle fasen (2-4 år), hvor barna fremdeles har vanskeligheter med å forstå spørsmål, formulere og evaluere svar og kommunisere svaret. I den intuitive fasen (4-7 år) utvikles disse ferdighetene (selv om språkferdigheter og verbalt minne er noe begrenset), og man kan stille dem enkle spørsmål der språket er tilpasset barnets nivå. Passende forskningsmetoder er små fokusgrupper og korte, kvalitative intervjuer, samt direkte observasjon (Tetzchner, 2001). Som en balansegang mellom barn som er for unge til å forstå spørsmålene våre, og barn som antakelig vil synes vanskelighetsgraden er for lett, tok vi en beslutning om å forsøke å designe for femåringer. Antall brikker er som nevnt noe lavt med tanke på Maldonados anbefalinger, men brikkenes komplekse form kompenserer for dette. Som Tetzchner (2001, s. 177) påpeker, og som vi naturligvis er klar over, kan selvfølgelig barn på samme alder befinne seg på ulike utviklingsstadier, og eldre og yngre barn kan også være på nivå med den gjennomsnittlige femåringen. «Femåringer» er således en *veiledende* målgruppe.

Vi har samlet informasjon om målgruppen i flere omganger, primært gjennom direkte observasjon og opportunistisk evaluering, men også gjennom eksisterende litteratur (se [referanseliste](#)) og intervjuer med barnehageansatte.

Første runde med datainnsamling, gjennomført uke 38, var en åpen observasjon hvor vi ikke interagerer med barna på noen måte, men utelukkende observerte dem leke med puslespill (noe de gjorde på eget initiativ). Fokuset for observasjonen var å få bedre forståelse for hvilke puslespill barna så ut til å like best, hvor avanserte disse spillene var og i hvilken grad de samarbeidet. Vi gjennomførte også et intervju med barnehagens pedagogiske leder der tema var femåringenes kapasitet og behov, samt diskusjon rundt prototypekonseptet vårt og råd til utforming. Informasjonen brukte vi for å utvikle de første prototypene, som vi så diskuterte med oppdragsgiver. Deretter utviklet vi prototype 3 som ble testet ved hjelp av opportunistisk evaluering med flere grupper barn, både i Forskningsparken Montessori og Uranienborg AKS.

Opportunistisk evaluering gjøres gjerne tidlig i designprosessen for å gi rask feedback på en designidé, hvilket vi trengte i forbindelse med papp-prototypen vår (prototype 3). Dette er en uformell form for evaluering hvor man i hovedsak er ute etter å bekrefte om det er verdt å utvikle prototypekonseptet videre, eller om ideen bør forkastes eller endres (Rogers et al., 2011). Til dette formålet benyttet vi hovedsakelig observasjon av barnas interaksjon med prototypen og enkle intervjuer tilpasset barnas nivå. Fordi prototypen skulle gå fra idéstadiet til high-fidelity på få uker var det viktig at datainnsamlingen ikke ble for ressurskrevende, samtidig som vi trengte bekreftelse på hvorvidt vi beveget oss i riktig retning. Med barna fra Uranienborg AKS hadde vi dessuten en liten «mini-workshop» hvor vi snakket med barna om hva de syntes ville være morsom, nyttig eller passende feedback, og hvor de hjalp oss å spille inn lydspor som brukes i den endelige prototypen.

Basert på datainnsamlingen og erfaringer vi gjorde oss, samt litteratur om emnet, utformet vi prototype 4. Vi har her brukt barna som *testers*, noe man typisk gjør for å få svar på spørsmål som hvilke deler av teknologien som er forvirrende, «buggy» eller bør endres og hvilke aspekter barn liker, samt hva man burde fokusere på å utvikle

videre (Druin, 2002). Denne formen for involvering kan ifølge Druin gi et godt grunnlag for å oppdage eventuelle problemer ved en prototype i løpet av noen timers testing. Vi ønsket samtidig å få en pekepinn på hvorvidt vi har designet i henhold til designprinsippene nevnt i [seksjon 2.1](#). I analysen av dataene har våre *key-questions* vært knyttet til overnevnte problemstillinger.

For hver observasjon- og intervjurunde har vi kategorisert dataene i henhold til et predefinert sett med kategorier, som omtalt i Taylor-powell & Renner (2003). For høyere inter-reliabilitet burde flere kodere ha kategorisert hele datasettet individuelt, men av tidsmessige årsaker ble vi nødt til å dele datasettet mellom flere personer. For at vi ikke skulle “presse” informasjon inn i kategorier som ikke passet, hadde koderne mulighet til å opprette egne kategorier for slik data. Det var imidlertid ingen som opprettet noen nye kategorier. For å få oversikt over materialet fra den aktuelle “runden” gjennomførte vi en *closed sorting*-gruppering i fellesskap, hvor koderne plasserte stikkordsbaserte *post it*-lapper i henhold til kategoriene (se vedlegg F). Deretter diskuterte vi kategoriene, betydningen av disse og hvordan funnene påvirket veien videre.

De viktigste erfaringene herfra førte til at vi utvidet den teknologiske plattformen ved å kombinere Arduino og C++ med Raspberry Pi og Python. Dette var for å gjøre teknologien mer stabil og for å forbedre tilbakemeldingen som ble gitt gjennom bedre lyd, hint og mer informativ tilbakemelding. Vi endret derfor tilbakemeldingen som ble gitt i den oppgavebaserte prototypen, slik at brukerne fikk tilbakemelding når de la feil land med informasjon om hvilket land de hadde lagt, og med en oppfordring om å prøve på nytt.

Hintene er ment å gjøre den nye kunnskapen mer tilgjengelig for målgruppen ved å fungere som det Vygotskij i pedagogisk sammenheng kaller «stillas»; en ytre støtte som bygger bro mellom det kjente og det ukjente, slik at opplæringen ikke blir frustrerende gjentatte møter med det ukjente. Slik opprettholdes problemløsningen gjennom å motivere barnet (Tetzchner, 2001, s. 302). Ifølge Sim et al. (2006) er det viktig å utforme tilbakemelding i interaktive læringssystemer på en slik måte at det ikke bare gir indikatorer på hva som er rett eller galt, men også sier noe om *hvorfor* det er galt: «*It is more effective to explain why the response is incorrect and provide the user with the correct answer.*» (Sim et al., 2006, s. 237).

Basert på observasjon av målgruppen og deres seleksjon av puslespill i barnehagen, samt forskning, valgte vi å male brikkene i sterke farger og lakkere dem blanke og appellerende. I førskolealderen foretrekker barn sterke farger uten hensyn til om de er realistiske, men utviklingen i forståelse av tegninger skjer nokså langsomt (Tetzchner, 2001, s. 129). Siden entydighet er viktig og detaljtegninger i denne alderen ikke har noen pedagogisk effekt, valgte vi bort faktiske karttegninger til fordel for ensfargede brikker.

3. Metoder for evaluering

Vi vil i denne seksjonen ta for oss metodene vi har brukt for å evaluere den siste prototypen. Valg av metode er avhengig av en rekke tema, som for eksempel formål, ressurs- og tidsbegrensninger, tilgjengelige deltakere eller brukskonteksten som skal undersøkes. Dette har vært førende for våre valg, noe vi vil gå gjennom i det følgende. Vi vil ikke diskutere formativ evaluering her, ettersom dette var deltema for presentasjonen holdt i oktober.

3.1 Metodevalg

Som tidligere nevnt har vi vært opptatt av å utvikle prototypen i henhold til vår oppdragsgivers ønske om at prototypen skal stimulere læring gjennom lek. Nettopp læring, eller læringsutbytte, vil derfor være et relevant aspekt å vurdere i evalueringssammenheng.

I artikkelen *Are Tangibles More Fun?* undersøker forfatterne forholdet mellom barns *enjoyment* og *engagement* ved bruk av ulike grensesnitttyper, og betoner at «*enjoyment and engagement are integral and prerequisite aspects of children's playful learning experiences*» (Xie, Antle, & Motamedi, 2008). Lek kan med andre ord bidra til å motivere læring, hvilket gjør det relevant å undersøke hvorvidt barna ser ut til å engasjeres av vår prototype. Læringsutbytte vil likevel være hovedfokuset for evalueringen, da vi mener dette er hovedfokuset ved prototypen.

Evaluering involverer vanligvis observasjon av deltagere og måling av deres prestasjoner gjennom brukbarhetstester (*usability testing*), eksperimenter eller *field studies* (Rogers et al., 2011, s. 434). I en undersøkelse av de mest brukte brukersentrerte metodene hvor man vurderte disse metodenes nytte, ble *field studies* rangert svært høyt på nytteskalaen (Vredenburg, Mao, Smith, & Carey, 2002). På bakgrunn av faglitteraturen bestemte vi oss likevel tidlig for å se bort ifra *field studies* som evalueringmetode, da denne metoden har sine begrensninger når det gjelder å teste spesifikke hypoteser og vurdere brukbarheten til produktet. I tillegg er *field studies* en metode som genererer store mengder med data som fort blir u håndterlig for studenter uten videre erfaring fra analysearbeid. Det er dermed også en svært tidkrevende metode (Rogers et al., 2011, s. 491).

Brukarhetstesting benyttes ofte for å avdekke og forbedre feil ved brukergrensesnittet (Lazar, Feng, & Hochheiser, 2010, s. 254). ISO-standarden 9241-11 (ISO, 1998) definerer *usability* som i hvilken grad et produkt kan bli brukt av spesifikke brukere for å oppnå spesifikke mål knyttet til *effectiveness*, *efficiency* og *satisfaction* i en bestemt brukskontekst. I og med at vi ønsket å vurdere mulig læringsutbytte ved bruk av vår prototype kom vi frem til at det var mindre hensiktsmessig for oss å benytte brukbarhetstesting, som fokuserer på andre parametre. Eksempelvis beskriver *efficiency* blant annet hvor effektivt en oppgave kan gjennomføres, ofte relatert til tid (Lazar et al., 2010, s. 26). Dette var noe vi ikke ønsket å måle ettersom det for læringsprosessens sin del er viktigere å komme frem til *hvordan* et problem løses heller enn hvor lang *tid* det tar. Bearbeidningshastighet alene behøver ikke å si mye om barns evne til å tenke og resonnerer (Tetzchner, 2001, s. 187).

Der *usability testing* primært benyttes for å evaluere grensesnitt og andre løsninger (Lazar et al., 2010, s. 255), brukes eksperimentell design i hovedsak til å forstå problemer og undersøke forskningsspørsmål. Vi anser derfor eksperiment for å være en mer passende metode for å evaluere våre spørsmål knyttet til læringsutbytte. Hovedgrunnen for dette er at eksperimentell design gjør det mulig å undersøke kausale relasjoner mellom ulike faktorer, noe deskriptive og relasjonelle metoder ikke muliggjør alene (Lazar et al., 2010, s. 27). Vi mente dette var avgjørende da vi ville undersøke mulig læringsutbytte ved interaksjon med *tangibles* (her puslespill).

Det er generelt utfordrende å utføre brukertesting med barn, blant annet fordi deres tilbakemelding kan formes av konteksten de er i. Hanna, Ridsen, & Alexander skriver i artikkelen *Guidelines for Usability Testing with Children* (1997) at barn i alderen 2-5 år ofte ønsker å tilfredsstille voksne; de kan la være å si noe de tenker fordi de ikke ønsker å skuffe en voksen, og andre ganger kan de si ting de ikke egentlig mener fordi de tror at det er det de voksne vil høre. Barn kan òg ha vanskeligheter med å uttrykke seg grunnet begrenset ordforråd, sjenanse eller rastløshet. De samme problemene kan naturligvis oppstå også ved bruk av eksperimentell design, men tilbakemeldingene vi er avhengig av i vårt spesifikke forskningsdesign vil antagelig ikke påvirkes i samme grad av barnas ønske om å tilfredsstille de voksne. Dette fordi et barn enten innehar kunnskapen om et land og dets plassering, eller det mangler denne kunnskapen. Hvis barna sier det de tror vi vil høre, kan de likevel ikke påvirke resultatene for denne type eksperiment, ettersom svaret i dette tilfellet vil være rett eller galt. Derimot vil det fremdeles være en fare for at de unnlater å svare på spørsmål og på den måten påvirker eksperimentet.

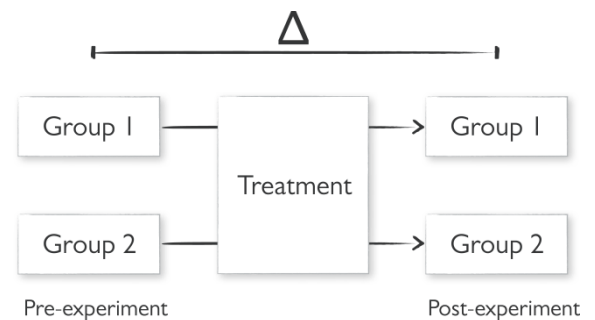
Utover å evaluere læringsutbytte, var vi interessert i å se hvorvidt prototypen ser ut til å engasjere barna. Basert på litteratur på området kom vi imidlertid frem til at dette er en krevende øvelse, spesielt når det gjelder barn (MacFarlane, Read, Höysniemi, & Markopoulos, (2003) gjengitt i Sim et al., 2006). Vi kom over ulike metoder og verktøy som er utviklet med tanke på å måle dette, som observasjonsmetoden Fun Toolkit (Read et al., 2002, gjengitt i Sim et al., 2006) og en Likert-skala med smileys (Sim et al., 2006). Sistnevnte har vi forstått det er mye uenighet rundt, blant annet ettersom yngre barn har en tendens til å gi mer positiv feedback enn eldre barn (Obrist et al., 2009).

Det er åpenbart utfordrende å tolke barns kroppsspråk og vurdere nyansene i dette på en objektiv, sammenlignbar og systematisk måte. Vi valgte derfor å gjøre kvalitative observasjoner underveis i eksperimentet, og se om barna likte eller mislikte prototypen, etter mal fra Price, Rogers, Scaife, Stanton, & Neale (2003). Som Lazar et al. snakker om forskerens erfaring og domenekunnskap, samt ferdigheter til å gjenkjenne det som er verdt å merke seg (2010, s. 376). Vi besitter imidlertid ikke den samme kompetansen, og å vurdere nyanser i engasjement hos barn føler vi oss ikke kvalifiserte til. Vi håpet imidlertid å få en pekepinn om hvorvidt underholdningsaspektet ved spillet er vellykket. Basert på dette måtte vi eventuelt foretatt en påfølgende undersøkelse for å avdekke hva konkret barna ville ha annerledes. Resultatene fra denne kvalitative vurderingen er gjengitt i [seksjon 4](#).

3.2 Design av eksperimentet

På bakgrunn av de vurderingene vi har gjengitt over valgte vi eksperimentell design for å gjennomføre vår evaluering av prototypen. Da utgangspunktet for vårt eksperiment har vært å prøve å evaluere hvorvidt interaksjon med de ulike prototypevariantene gir læringsutbytte og om det eventuelt er noen forskjell mellom de to variantene, valgte vi between-group for å unngå læringseffekten som kan oppstå ved *within-group* design. Siden vår brukergruppe er unge barn ga dette valget også en klar fordel i forhold til within-group design, da vi unngikk at de måtte utføre flere oppgaver som kunne føre til utmattelse og frustrasjon noe som igjen kunne påvirke resultatene våre (Lazar et al., 2010, s. 47)

Barna vil, som beskrevet nedenfor, utføre eksperimentet parvis. Vi valgte denne tilnærmingen av flere årsaker. Et viktig moment ved vår brukergruppe har vært å sørge for at de føler seg trygge under eksperimentet. Ved å være to har de noen de kjenner i møte med ukjente voksne og føler kanskje ikke like stor prestasjonsangst som de ville gjort alene. Et annet viktig poeng er at prototypen skal legge til rette for samarbeid og at den reelle brukskonteksten involverer at mer enn ett barn interagerer med prototypen. Ifølge Xie øker det å jobbe i små grupper barnas tilfredshet, engasjement og motivasjon (Xie et al., 2008). Basert på antagelsen om at samarbeid påvirker disse faktorene, ville vi antakeligvis ende opp med resultater som med lavere økologisk validitet dersom barna måtte pusle alene (Tetzchner 2001, s.46).



Figur 3.1 – Kvasi-eksperimentelt design hentet fra Mutlu (2011).

3.2.1 Beskrivelse av eksperimentet

Vi ønsker gjennom dette eksperimentet å kartlegge om det er noen forskjell i antall land femåringer kan identifisere før og etter interaksjon med de to prototypene våre. Vi ønsker å se om barna har større læringsutbytte med én av de to variantene, i form av antall land de er i stand til å identifisere etter lek med prototypevarianten. Med læringsutbytte mener vi altså ervervet kunnskap. Dette vil vi undersøke gjennom et between-group design hvor én gruppe barn får leke med den ene varianten og én gruppe med den andre. Barna vil utføre eksperimentet parvis, og vil i forkant av interaksjon med prototypen bli bedt om å identifisere de ulike landene i prototypen, slik at vi kan gjennomføre den samme testen etter interaksjonsrunden og se om det er noen forskjell i antall land barna kan identifisere. Barna får interagere med den aktuelle prototypevarianten frem til alle brikkene er lagt.

Hypoteseformulering

H0: Det er ingen forskjell i antall land femåringer kan identifisere før og etter interaksjon med de to prototypevariantene.

H1: Det er forskjell i antall land femåringer kan identifisere før og etter interaksjon med prototypevariant 1.

H2: Det er forskjell i antall land femåringer kan identifisere før og etter interaksjon med prototypevariant 2.

Uavhengige variable: Prototypevariant

Avhengige variabler: Antall korrekt identifiserte land

Kontrollerte variabler: Oppgavene barna får

Vilkår:

Prototypevariant 1: Oppgavebasert, interaktivt geografipuslespill

Prototypevariant 2: Fri lek med interaktivt geografipuslespill

3.4 Utvalg

Vårt utvalg består av barn i vår målgruppe: 18 jenter og 22 gutter i femårsalderen. Disse ble valgt fra fire forskjellige barnehager; Friggfeltet barnehage, Barnas barnehage, Bamsebo barnehage og Gydas vei barnehage. Barnehagene ble valgt hovedsakelig med tanke på enkel transport av prototypen, det vil si at den måtte enten være i nærheten av Ole Johan Dahls hus eller i umiddelbar nærhet ved hjelp av kollektivt transport. Utvalget vårt kan dermed sies å være *convenience sampling*, siden vi har brukt de barna vi av praktiske årsaker fikk tilgang til, heller enn et strategisk eller helt randomisert utvalg (Rogers et al., 2011, s. 224).

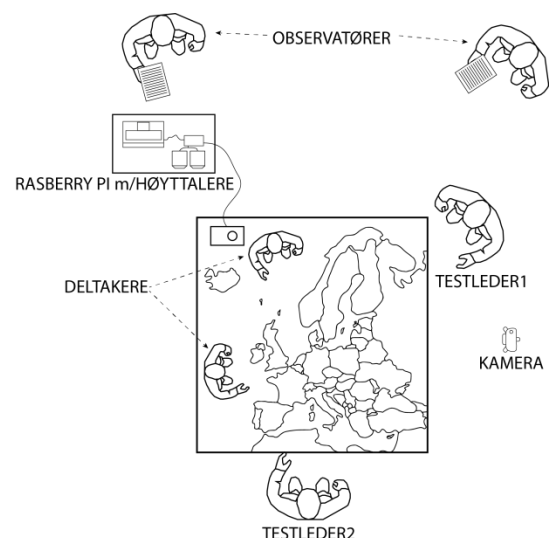
Ved gjennomføringen av eksperimentet ble barna valgt ut av bestyreren, i dialog med de foresatte som måtte godkjenne deltakelse gjennom et samtykkeskjema. På bakgrunn av dette kan vi ikke hevde at eksperimentet er et ekte eksperiment, men heller et kvasi-eksperiment da utvelgelsen ikke var tilfeldig (Lazar et al., 2010, s. 42).

En annen begrensning ved vårt eksperiment er antall deltagere. En viktig forutsetning for representativitet er at utvalget er stort nok. En liten gruppe øker sannsynligheten for at resultatene ikke er representative for hele populasjonen (Tetzchner, 2001, s. 53).

3.5 Oppsett

Man streber etter å gjennomføre HCI-eksperimenter der man i størst mulig grad kan kontrollere alle variable, for eksempel i laboratorium. Vi gjennomførte derimot eksperimentene ute i barnehagene. Grunnen til dette var at vi ønsket å øke resultatenes økologiske validitet ved å gjennomføre eksperimentene i en tilnærmet reell brukskontekst. Samtidig vil dette kunne bidra til at barna føler seg tryggere, når de er i sitt vante miljø (Lazar et al., 2010, s.265; Hanna et al., 1997).

Testoppsettet var bestemt på forhånd (se figur 3.2) og dette ble forsøkt gjennomført på lik måte i hver av de fire barnehagene i den grad det lot seg gjøre med tanke på romstørrelse. Ved gjennomføringen av pilot-eksperimentene, erfarte vi at to testledere fungerte bedre enn kun én. Dette var hovedsakelig på grunn av at testlederens spørsmål til deltakerne kunne omformuleres raskere dersom barna hadde vanskeligheter med å svare på spørsmålet (Tetzchner, 2001, s.46). Dette bidro til å legge mindre press på testlederne, og åpnet for en bedre dynamikk mellom dem og barna. Zaman påpeker i artikkelen *Evaluating games with children* (2005) at "two



Figur 3.2 - Testoppsett

researchers are needed when doing user tests with children”. Han sier videre at *“the quality of the information gained by the user’s answers during the tests, depends on whether the child feels at ease. If children feel uncomfortable, they become silent and important information will be lost”*.

I tillegg til testlederne var det to observatører for å øke reliabiliteten, ved at samme situasjon ble registrert av flere (Lazar et al., 2010, s. 212). Ideelt sett burde ikke observatørene ha vært synlige da dette kan medføre at deltakerne blir mer stresset i en allerede ukjent situasjon med fremmede (Wilson et al., 2010). Likevel ville dette vært vanskelig å gjennomføre ettersom eksperimentet ikke ble gjennomført i et laboratorium med tilgang til enveisspeil eller skjult kamera.

Deltakerne gjorde seg kjent med prototypen (heretter kalt puslespillet) ved at det lå ferdig puslet på bakken når eksperimentet startet. Da hensikten med eksperimentet ikke er å se om deltakerne kan løse puslespillet, men heller gjenkjenne et visst antall land, vil det på denne måten være enklere for deltakerne å løse puslespillet ettersom det er en fordel å kunne huske tilbake til en episode da puslespillet var helt (Tetzchner, 2001, s. 212). For å gjøre eksperimentet så likt som mulig hver gang var det satt opp et manuskript med spørsmål (se vedlegg C), samt en kort introduksjon til prototypen og deltakernes oppgaver (Hanna et al., 1997) som testlederne skulle gjennomføre før deltakerne fikk mulighet til å pusle. Dette er kodet som “forkunnskap” i våre resultater ([seksjon 4](#)) og observatørene hadde som oppgave å notere ned svarene. Det samme spørsmålssettet ble benyttet når puslespillet var ferdig løst og dette er kodet som «etterkunnskap» i våre resultater.

Spørsmålene er formulert på enkel måte slik at barn på denne alderen ikke skulle ha noen problemer med å forstå disse (Tetzchner, 2001, s.47). Det var samtidig lagt opp til at testledernes tonefall skulle gjøres mest mulig naturlig og undrende for at barna genuint skulle tro at dette var noe vi lurte på. Rekkefølgen på spørsmålene ble forsøkt holdt så konsekvent som mulig, men en videre diskusjon rundt utfordringene vi møtte på kan sees under [seksjon 5.2](#).

Som tidligere nevnt var det to observatører med i rommet. Observatørene konsentrerte seg hovedsakelig om to ting under puslingen: notering av brukbarhetsproblemer ved puslespillet og puslebrikkene, samt notering av indikatorer på glede og engasjement som for eksempel tegn på positivt kroppspråk, kommentarer, smil og latter. Eventuell frustrasjon ble også notert. Indikatorer på dette kunne for eksempel være å se seg rundt i rommet, flytte blikket vekk fra puslespillet eller å sukke. Slike atferdsmessige tegn kan være en mer pålitelig kilde enn barnas svar på direkte spørsmål om hvorvidt de var fornøyd eller likte puslespillet (Hanna et al., 1997).

3.5.1 Teknisk opptaksutstyr

Vi benyttet oss av to kameraer for å dokumentere eksperimentet. Det ene ble benyttet til å ta stillbilder og det andre var montert på en tripod rettet mot puslespillet og testlederne. Sistnevnte kamera tok video- og lydopptak for å kunne kvalitetssikre notatene i etterkant og til bruk ved en eventuell kvalitativ analyse.

3.6 Ethiske aspekter ved eksperimentet

Deltakerne i vårt eksperiment er en gruppe som kan være spesielt sårbar ved at de vanskelig kan forsvare sine interesser eller være like tydelig på grensesetting (Fangen, K., 2010). Vi har tatt hensyn til prinsippene til Society for Research in Child Development (SRCD, 2007) og de forskningsetiske retningslinjene til Forskningsetisk komité for humaniora og samfunnsfag (NESH, 2006). Disse understreker behovet for å tilrettelegge forskningsmetodikken til de individuelle behovene hos barna. Et annet viktig element er informert samtykke. Da våre deltakere er under 16 år har vi innhentet dette gjennom samtykkeskjema som de foresatte har skrevet under på (se vedlegg A).

Det er viktig å ha i bakhodet at det ikke er så lett å forutse hvordan barn reagerer, og vi kan heller ikke vite noe om hva de føler og tenker, verken under eller etter eksperimentet. Utfordringer vedrørende mobbing eller negative tanker knyttet til egne prestasjoner er det dessverre er vanskelig for oss å forutse og motvirke effekten av. Vi forsøker naturligvis å oppfordre til en vennlig, morsom tone barna imellom, men hvordan barna samhandler kan vi vanskelig gripe inn i.

4. Resultater og analyse

For å undersøke hypotesene vi har jobbet utfra har vi som sagt gjennomført et kvasi-eksperiment. Vi har utført 10 observasjoner for hver av vilkårene i kvasi-eksperimentet, med til sammen 40 deltakere (18 jenter og 22 gutter). Disse var fordelt på seks par med gutter, fire par med jenter og ti par med en jente og en gutt. Disse ble jevnt fordelt på hver av gruppene i eksperimentet.

4.1 Måling av læringsutbytte

Ifølge Sim et al. (2006) kan et eventuelt læringsutbytte av et system måles med før- og ettertester. Som nevnt tidligere gjennomførte vi slike tester i form av muntlige spørsmål til barna om hvilke land de gjenkjente før og etter at de brukte prototypen (se [seksjon 3](#)). Disse målte hvor mange av de 13 landene som deltagerne kunne identifisere før og etter å ha brukt prototypen. Læringsutbytte ble kalkulert ut i fra differansen mellom resultatene fra for- og ettertestene (tabell 4.1.1). Begrensninger ved denne metoden er diskutert i seksjon fem.

Tabell 4.1.1: Gjennomsnittsverdier for læringsutbytte i de to ulike gruppene

Gruppe	Forkunnskap	Etterkunnskap	Læringsutbytte
Fripusling	0,6	2,3	1,7
Oppgavebasert	1,1	4,8	3,7

Tabell 4.1.1 viser at det er forskjell i gjennomsnittsverdien for læringsutbytte i de to gruppene. Forskjeller i snittverdiene gir ikke alene grunnlag for å trekke noen konklusjoner, men en indikator på at det kan være interessante forskjeller.

Som tidligere diskutert hadde vi ikke mulighet til selv å bestemme sammensetningen av individene i utvalget for vårt eksperiment, da dette ble avgjort av barnehagene og hvilke foreldre som ga informert samtykke. Vi har derfor ikke satt sammen gruppene

ut i fra nivå på forkunnskap om europeiske land. Basert på gjennomsnittsverdiene for de to gruppene ser vi at det er en forskjell, om enn liten, der gruppen som brukte den oppgavebaserte modulen ser ut til å ha høyere forkunnskap.

For å være sikker på at det ikke var signifikant forskjell mellom *utgangspunktet* til de to gruppene utførte vi derfor en t-test. Vi brukte t-test fordi utformingen av vårt eksperimentelle design har én uavhengig variabel (prototypevariant) med to vilkår (oppgavebasert og fri lek) og Lazar et al. (2010, s. 75) anbefaler t-tester for slike eksperiment. Da målingene kommer fra forskjellige utvalg har vi benyttet oss av uavhengige, tohalede t-tester (Langdridge, 2006, s. 170-171).

Vi antok at det ikke var noen forskjell mellom forkunnskapene til deltakerne i de to gruppene. En t-test (tabell 4.1.2) for to utvalg med antatt ulike varianser viser at det ikke er noen signifikant forskjell mellom forkunnskapene om europeiske land mellom de to ulike gruppene ($t(13) = -0,98$, $p < 0,05$). T-testen ga oss en p-verdi på 0,34 som tilsier at det er stor sannsynlighet for at den forskjellen som ble observert i snittene ville forekomme som følge av tilfeldige feil.

Tabell 4.1.2: t-Test: To utvalg med antatt ulike varianser

Forkunnskaper	Fripusling	Oppgavebasert
Gjennomsnitt	0,6	1,1
Varians	0,488888889	2,1
Observasjoner	10	10
Antatt avvik mellom gjennomsnittene	0	
fg	13	
t-Stat	-0,982682673	
P(T<=t) ensidig	0,171849769	
T-kritisk, ensidig	1,770933396	
P(T<=t) tosidig	0,343699539	
T-kritisk, tosidig	2,160368656	

4.1.1 Undersøkelse av nullhypotesen

Først ønsket vi å se om vi kunne forkaste nullhypotesen som sa følgende: *Det er ingen forskjell i antall land femåringer kan identifisere før og etter interaksjon med de to prototypevariantene.*

Ved å se nærmere på ulik deskriptiv statistikk (tabell 4.1.1.1) for vårt datagrunnlag ser vi at verdiene for gjennomsnitt og median for etterkunnskap er tilnærmet lik for begge gruppene (0,3 i forskjell for begge), mens det er noe større forskjell i snitt og median for forkunnskap for den oppgavebaserte gruppen. Modus (mode) ligger også tett opp mot disse verdiene for etterkunnskap. Dette kan passe med en normalfordeling av verdiene for det vi har ønsket å måle, nemlig etterkunnskap og læringsutbytte. Dette er viktig når man skal gjennomføre parametriske tester for å undersøke signifikans (Langdridge, 2006, s. 169).

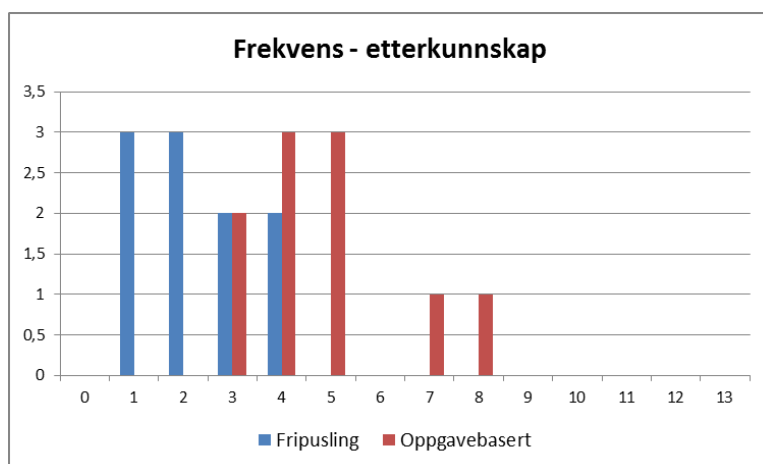
For å undersøke dette har vi sett hvor hyppig de ulike verdiene for etterkunnskap forekommer for hver av gruppene og lagt dette inn i histogrammet under (Figur

4.1.1.1). Her ser vi at hovedandelen av verdiene faller rundt gjennomsnittet til de to ulike gruppene (2,3 og 4,8). Når vi har sett på fordelingen av data ser vi at om lag 2/3 av disse faller mellom gjennomsnitt – standard avvik og gjennomsnitt + standard avvik for etterkunnskap for hver gruppe. Dette kan tyde på at dataene våre er tilnærmet normalfordelt, men det er vanskelig å være sikker på grunn av omfanget på utvalget vårt. Som man også kunne se ut av den deskriptive statistikken er det større spredning blant verdiene i gruppen som brukte den oppgavebaserte modulen, enn for gruppen som brukte fripusling.

Tabell 4.1.1.1: Måling av for- og etterkunnskaper og læringsutbytte

	Fripusling			Oppgavebasert		
	For-kunnskaper	Etter-kunnskap	Lærings-utbytte	For-kunnskaper	Etter-kunnskap	Lærings-utbytte
	1	3	2	0	4	4
	0	2	2	0	5	5
	0	1	1	0	4	4
	0	1	1	3	8	5
	0	2	2	2	5	3
	1	4	3	0	3	3
	2	4	2	1	3	2
	1	1	0	0	5	5
	1	2	1	1	4	3
	0	3	3	4	7	3
Count	10	10	10	10	10	10
Sum	6	23	17	11	48	37
Mean	0,6	2,3	1,7	1,1	4,8	3,7
Median	0,5	2	2	0,5	4,5	3,5
Mode	0	2	2	0	4	3
Range	2	3	3	4	5	3
Variance	0,489	1,344	0,90	2,100	2,622	1,12
Standard deviation	0,699	1,160	0,95	1,449	1,619	1,06

Våre alternative hypoteser omhandler læringsutbyttet ved hver enkelt modul. Som vist tidligere så vi en forskjell i gjennomsnittsverdier for før- og etterkunnskap for hver gruppe. For å undersøke om det var en signifikant forskjell mellom disse verdiene gjennomførte vi toaledede t-tester med gjennomsnitt for to parvise utvalg.



Figur 4.1.1.1: Frekvens - etterkunnskap

En t-test (tabell 4.1.1.2) med gjennomsnitt for to parvise utvalg viser at det er en signifikant forskjell på for- og etterkunnskapene for gruppen som brukte modulen med fripusling ($t(9) = -5,66$, $p < 0,05$). Denne t-testen ga en p-verdi på 0,0003, hvilket betyr at det er svært lav sannsynlighet for å oppnå en slik endring mellom gjennomsnittene ved en tilfeldighet.

Tabell 4.1.1.2: T-Test: Gjennomsnitt for to parvise utvalg

Fripusling		
	Forkunnskaper	Etterkunnskap
Gjennomsnitt	0,6	2,3
Varians	0,488888889	1,344444444
Observasjoner	10	10
Pearson-korrelasjon	0,575612567	
Antatt avvik mellom gjennomsnittene	0	
fg	9	
t-Stat	-5,666666667	
P(T<=t) ensidig	0,000153511	
T-kritisk, ensidig	1,833112933	
P(T<=t) tosidig	0,000307022	
T-kritisk, tosidig	2,262157163	

En t-test (tabell 4.1.1.3) med gjennomsnitt for to parvise utvalg viser at det er en signifikant forskjell på for- og etterkunnskapene for gruppen som brukte den oppgavebaserte modulen ($t(9) = -11$, $p < 0,05$). Denne t-testen ga en p-verdi på 0,000001, som betyr at sannsynligheten for at en slik forskjell i gjennomsnittsverdier skulle oppstå ved en tilfeldighet er betydelig liten.

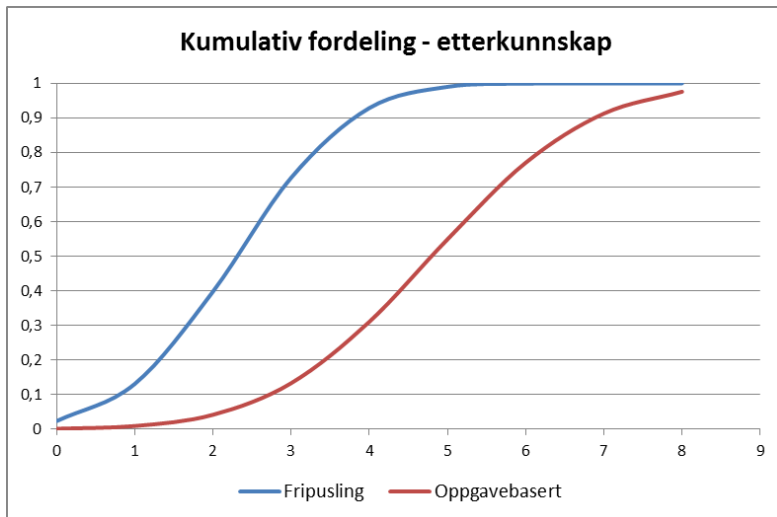
Tabell 4.1.1.3: T-Test: Gjennomsnitt for to parvise utvalg

Oppgavebasert		
	Forkunnskaper	Etterkunnskap
Gjennomsnitt	1,1	4,8
Varians	2,1	2,622222222
Observasjoner	10	10
Pearson-korrelasjon	0,767057836	
Antatt avvik mellom gjennomsnittene	0	
fg	9	
t-Stat	-11,04491281	
P(T<=t) ensidig	7,77769E-07	
T-kritisk, ensidig	1,833112933	
P(T<=t) tosidig	1,55554E-06	
T-kritisk, tosidig	2,262157163	

Disse to testene viser at det er forskjell på læringsutbyttet ved bruk av begge prototypevariantene og tyder på at vi kan forkaste nullhypotesen vår.

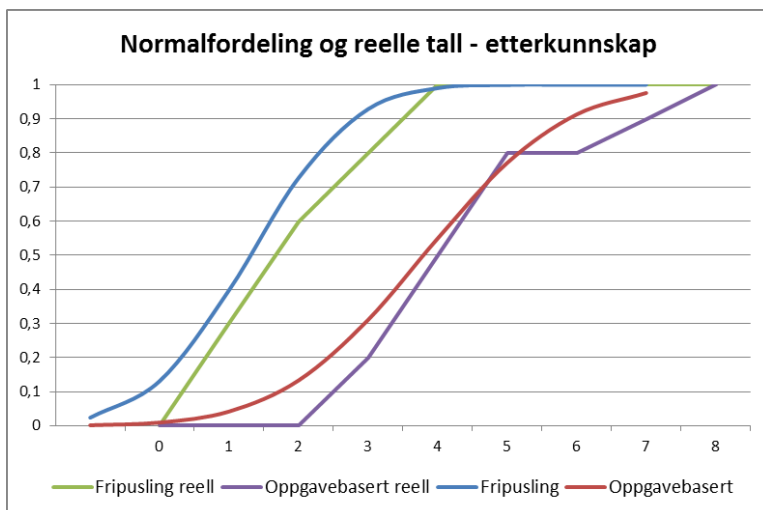
4.1.2 Forskjell mellom prototypene

Selv om vi ikke hadde formulert hypoteser for om det var forskjell i grad på læringsutbytte mellom de to prototype-variantene, syntes vi det var interessant og også å undersøke dette. For å lettere se om det er en markant forskjell mellom disse to har vi plottet verdiene inn i en kumulativ kurve:



Figur 4.1.2.1: Kumulativ fordeling - etterkunnskap

Ser vi på gjennomsnittet for etterkunnskap til den oppgavebaserte gruppen (4,8) og sammenligner med den kumulative kurven for fripusling ser vi at om lag 98 prosent havner under denne verdien. Dette må kunne sies å være en markant forskjell. Dette kan tyde på at det er forskjell i læringsutbytte mellom de to spill-modulene, men man kan ikke bare basere seg på gjennomsnittsverdier alene når dette skal undersøkes. Noe av grunnen til dette er at vårt utvalg ikke er tilfeldig og at størrelsen på utvalget ikke er spesielt stort (som diskutert i seksjon 3). Dette ser vi tydeligere når vi sammenstiller den kumulative fordelingskurven og reelle prosenttall (frekvens av verdi delt på antall forekomster). Selv om kurvene følger hverandre nokså likt, er det tydelig at vi har for få målinger til at de skulle følge den kumulative fordelingskurven nøyaktig.



Figur 4.1.2.2: Normalfordeling og reelle tall - etterkunnskap

For å undersøke om vårt begrensede utvalg er representativt for en større populasjon er vi nødt til å benytte oss av signifikanstesting mellom læringsutbytte ved bruk av de to ulike modulene (fripusling og oppgavebasert).

En tohalet t-test (tabell 4.1.2.1) for to utvalg med antatt ulike varianser viser at det er en signifikant forskjell på læringsutbytte oppnådd mellom gruppen som brukte fripuslings-modulen og gruppen som brukte den oppgavebaserte modulen ($t(18) = -4,45$, $p < 0,05$). Denne t-testen ga en p-verdi på 0,0003 som tilsier at en slik forskjell i gjennomsnittene bare kan oppstå tilfeldig med en sannsynlighet på 0,0003. Dermed er forskjellen signifikant.

Tabell 4.1.2.1: t-Test: To utvalg med antatt ulike varianser

Læring	Fripusling	Oppgavebasert
Gjennomsnitt	1,7	3,7
Varians	0,9	1,12222222
Observasjoner	10	10
Antatt avvik mellom gjennomsnittene	0	
fg	18	
t-Stat	-4,4474959	
P(T<=t) ensidig	0,000155518	
T-kritisk, ensidig	1,734063592	
P(T<=t) tosidig	0,000311035	
T-kritisk, tosidig	2,100922037	

4.1.3 Regresjonsanalyse

Selv om en t-test viser at det er signifikant forskjell mellom snittene er det vanskelig å si sikkert om det er vår uavhengige variabel (prototype) som har ført til denne endringen. Da barna som deltok alle var i femårsalderen og hadde omtrentlig lik forkunnskap, prøvde vi å undersøke om det var noe forhold mellom etterkunnskap og kjønn. Hver av deltakerparene ble kodet som 2 (gutt), 1 (jente) og 0 (gutt og jente). En regresjonsanalyse som så på forholdet mellom etterkunnskap og kjønn ga en R-kvadrat verdi på 0,001 og en justert R-kvadrat verdi på -0,05, på en skala fra -1 til 1. P-verdien for denne testen var på 0,87 som tyder på at det er svært høy sannsynlighet for at observerte sammenhenger ville forekomme tilfeldig. Dette tyder på at det ikke er en sammenheng mellom kjønn og etterkunnskap, og kan sånn sett styrke våre funn om at vår uavhengige variabel har hatt en effekt på læringsutbytte. Med vårt begrensede utvalg er det likevel svært vanskelig å påvise en slik sammenheng.

4.2 Kvalitativ observasjon av atferd

Samtlige eksperimenter ble logget av to observatører, som blant annet observerte kroppsspråket til barna og noterte hva de sa underveis i spillet for å prøve å vurdere hva de mente om prototypen og leken. For de barna der vi fikk godkjenning av de foresatte har vi konsultert videomaterialet der dette var nødvendig.

Generelt observerte vi at barna var veldig engasjerte og deltok aktivt underveis i spillet. Dette ga seg uttrykk gjennom ulik oppførsel; noen var veldig aktive og rastløse, mens andre forholdt seg rolig og var svært fokuserte. Noen fremstod som undrende når de fant frem til riktig land og det virket som om det gikk opp et lys for dem: “Åja, så det er dette som er Hellas” sa en jente.

For flere av deltakerparene observerte vi en høy grad av samarbeid, spesielt i den oppgavebaserte modulen. Eksempelvis, når et deltakerpar ble bedt om å finne Spania, som den ene gutten satt med, utbrøt jenta: “Det er Spania!”. Under en annen

observasjon, der de blir bedt om finne Polen, går det en stund slik at puslespillet gir hint om at landet er grønt. Da plukker den ene gutten opp Polen, men plasserer det feil. Den andre gutten plukker opp Norge, som også er grønt, men gutten med Polen forklarer at “Nei, det er Norge”. Så legger han Polen rett. Når oppgaven senere er å legge Norge gjør de dette umiddelbart. En tredje situasjon der de blir bedt om å finne Tyskland plukker den ene opp Hellas og spør den andre “Er dette det [landet]?”. Den andre sier, “Det er Hellas”, som han har prøvd å legge tidligere.

Flere av barna kom med stadige gledesutbrudd underveis i spillingen; innimellom kunne vi se dem løfte armene triumferende i været eller klappe i hendene. Eksempelvis var det to gutter som ropte “Vi er de kuleste! Vi klarte nesten alle!” etter at de var ferdige. En annen gutt var tydelig fornøyd når han fant ut hvilke land han skulle legge, og løp bort til kameraet for å vise det frem. En jente ropte henrykt “Dette landet skal jo der!” mens hun holdt Sverige i været og pekte på omrisset. En gutt sa: “Det var morsomt å pusle puslespill!”.

Flere ganger virket det som om dimensjonene på puslespillet lå til rette for en større innsikt i størrelsen på landene og avstanden mellom dem. En jente sa for eksempel: “Island? Er *det* Island? Så langt borte?”, mens en gutt sa: “Danmark, er det så lite?” og “Norge, det er enda større! Enda større enn andre land!”. En jente kunne forklare at man måtte ta to fly til Frankrike fordi: “Frankrike er langt borte fra Norge, ser dere”.

Vi observerte også barn som strevde med å finne frem til rett land og som ble synlig frustrerte. Eksempelvis sa en jente: “Dumme puslespill, hvorfor må det være så vanskelig? Dette er veldig vanskelig, altså!” da hun hadde brukt lang tid på å plassere Tyskland. Frustrasjonene så imidlertid ut til å være glemt med en gang hun fikk neste oppgave, og hun utbrøt “Åh, det er kanskje denne!”, om en potensiell brikkekandidat.

4.3 Funn

Basert på målingene av for- og etterkunnskap for hver gruppe har vi ved hjelp av signifikanttester kunnet forkaste nullhypotesen som vi jobbet utfra. Vi mener å ha validert de alternative hypotesene som sier at det er forskjell på hvor mange land deltakerne kan før og etter at de har benyttet prototypen. En regresjonsanalyse som så på forholdet mellom etterkunnskap og kjønn tyder på at kjønn ikke er en variabel som påvirker antall rette land. En signifikanttest mellom resultatet av læringsutbytte for hver av gruppene tyder på at det er en signifikant forskjell mellom snittene, som tyder på at de som brukte den oppgavebaserte modulen kunne identifisere flest land. Dette kan være fordi det var en høyere grad av tilbakemeldingsfunksjoner i denne, som hint og feilmeldinger.

Alle våre signifikanttester er gjennomført med et signifikansnivå på 5 % for å forminske muligheten for en type I-feil der vi antar at vår uavhengige variabel har hatt en effekt, men at denne kommer av tilfeldige feil (Langdridge, 2006, s. 169). Når man reduserer sannsynligheten for å gjøre en type I-feil øker man risikoen for en type II-feil, som består i å ikke forkaste nullhypotesen. Generelt anser man type I-feil til å være verre enn type-II feil, fordi det kan føre til at man kommer frem til konklusjoner som egentlig ikke lar seg forsvare.

Selv om vi har påvist signifikante forskjeller i gjennomsnittsverdiene mellom for- og etterkunnskap og for læringsutbytte mellom de to prototypene er det viktig å understreke at dette ikke nødvendigvis sier noe om størrelsen eller viktigheten av våre funn. Som nevnt tidligere har ikke våre data nødvendigvis høy ekstern validitet som følge av begrensninger i vårt utvalg. Videre hefter det flere usikkerhetsmomenter ved hvorvidt vår prototype faktisk har ført til et varig læringsutbytte hos deltakerne, men dette har ikke vært omfanget for disse testene. Dette momentet diskuteres i [seksjon 5](#).

Basert på ansiktsuttrykk, kroppspråk og verbale uttrykk dannet vi oss en oppfatning om at det oppgavebaserte puslespillet klarte å kapre og vekke barnas interesse på en bedre måte enn fripuslingsmodulen. Det kan være flere årsaker til dette. Eksempelvis kan det skyldes en høyere grad av interaksjon med puslespillet eller at den oppgavebaserte versjonen i større grad trigger konkurranseinstinktet hos barna. I denne versjonen fikk også barna hint underveis og mer tilbakemelding når de gjorde feil, noe som kan ha bidratt til større interesse.

5. Begrensninger ved prosjektet

I denne seksjonen går vi gjennom de ulike begrensningene ved vårt arbeid i dette prosjektet. Vi ser på de innledende observasjonene, opportunistisk evaluering, pilotundersøkelsen og kvasi-eksperimentet.

5.1 Reliabilitet

Et viktig tema for HCI-forskning er hvorvidt resultatene fra forskningen kan bli gjenskapt av andre forskere under lignende omstendigheter. Denne reliabiliteten kan svekkes ved at det gjøres systematiske feil, såkalte bias, under forskningsprosjektet. Ifølge Store Norske Leksikon brukes begrepet bias om: *“utsagn, valg og vurderinger som på systematisk måte avviker fra det som er faktisk korrekt.”* Det er fem hovedkategorier for systematiske feil og vi vil gjennomgå disse der de har vært relevant for vårt prosjekt.

5.1.1 Måleinstrumenter

Måleinstrumenter som benyttes for å samle inn data under et eksperiment kan brukes feil, ha feil ved seg eller på annen måte føre til at data ikke blir registrert korrekt (Lazar et al., 2010, s.59). I vårt prosjekt har observatørene under eksperimentet vært vårt primære måleinstrument gjennom å dokumentere resultatene og relevant data som ble observert under eksperimentene. En person tolker og oppfatter subjektivt, og dermed kan bias oppstå. For å minske risikoen for dette, har vi benyttet oss av to observatører som begge har samlet inn data. I tillegg har vi benyttet oss av videoopptak for å underbygge observasjonene. Dette ble kun benyttet når vi hadde anskaffet tilstrekkelig samtykke fra de foresatte.

5.2 Oppsett for gjennomføring av eksperimentet

Oppsettet for eksperimentet og gjennomføringen er et kritisk punkt. I dette stadiet var det en rekke utfordringer som kunne lede til bias. Basert på erfaring fra INF1500 og INF1510 var vi klar over nødvendigheten av å gjennomføre en pilot-test. Dette er et

essensielt punkt i alle eksperimenter, og kan bidra til å avdekke feil og mangler ved eksperimentet (Lazar et al., 2010, s. 60). Basert på resultatene av pilot-testen oppdaget vi flere punkter vi måtte ta til etterretning for å eliminere bias. Disse gjennomgås under.

Vi så at det var behov for at alle barna som skulle bruke prototypen startet med samme fysiske utgangspunkt, ved at de så puslespillet fra samme vinkel. Årsaken til dette er at de fleste er vant til å se kartet med Norge lengst mot nord. I eksperimentene startet derfor alle barna på «sør-siden» av kartet slik at de så puslespillet fra en naturlig vinkel.

Instruksjonene som deltagerne mottar spiller en viktig rolle i utførelsen av eksperimentet, og dette er spesielt viktig når man arbeider med barn. Formuleringene i instruksjonene må være nøye utarbeidet og testet. Hvordan vi skulle formulere oss på en best mulig måte for å presisere oppgavene vi ønsket at deltakerne skulle utføre var et viktig moment i pilot-eksperimentet. Vi løste dette ved å utarbeide et skriftlig dokument som inneholdt all informasjonen vi ønsket å gi før, etter og under eksperimentet. Da barna ikke kunne lese hadde vi likevel en utfordring med å sikre at informasjonen til samtlige deltakere ble gitt på lik måte. Testlederne fikk som oppgave å gjengi denne informasjonen. Under selve eksperimentet måtte vi noen ganger avvike fra malen litt avhengig av situasjonen. Ideelt sett burde eksempelvis rekkefølgen på spørsmålene ha vært så konsekvent som mulig, men dette var i noen tilfeller krevende da barn har en tendens til å vandre av sted i sin egen tankeprosess på grunn av deres korte oppmerksomhetsspenn (Hanna et al., 1997). Tetzchner viser til at en lite fleksibel testsituasjon kan gjøre barn usikre, og at usikkerheten kan føre til at de misforstår eller «lukker» seg, samt at usikkerheten blir forsterket når barna oppdager at de misforstår (Tetzchner, 2001, s. 47). Derfor var det plass for noe improvisasjon i rekkefølgen og lett omformulering av spørsmål i møte med det enkelte barnet, så lenge testleder gikk gjennom samtlige spørsmål. Det kan diskuteres hvorvidt testlederne og observatørene burde benyttet noe lengre tid med deltakerne før eksperimentet ble igangsatt slik at barna fikk en økt trygghetsfølelse rundt testlederne (Hanna et al., 1997).

5.2.1 Deltakere

Som vektlagt i [seksjon 3.4](#) har vi benyttet oss av *convenience sampling*. På bakgrunn av dette er det sannsynlig at utvalget vårt har demografiske begrensninger, og resultatene dermed ikke kan generaliseres (Lazar et al., 2010, s. 60). Selv om utvalget vårt består av både offentlige og private barnehager, befinner alle barnehagene seg på Oslos vestkant. Denne gruppen er trolig ikke direkte representativ for femåringer generelt. Det pedagogiske opplegget mellom forskjellige barnehager kan dessuten variere, og dermed påvirke barnas ferdigheter. Vi har ingen forutsetninger for å vite hvor ulike barn/barnehager befinner seg på en ferdighetskala, og måtte derfor tatt det samme forbeholdet angående mulig avvik fra målgruppen også om vi jobbet med andre barnehager. Det samme kan sies om datainnsamlingen vi gjorde i forbindelse med prototypeutviklingen, som vi gjennomførte i samarbeid med Forskningsparken Montessori.

Utvelgelsen av hvilke barn som skulle danne par ble gjort av bestyreren i hver enkelt barnehage, og kan derfor ikke kalles randomisert. Bestyreren kan ha lagt ulike vurderinger til grunn for utvelgelsen, eksempelvis barnas personlighet, ferdigheter innen pusling eller interesseområder. Dette påvirker trolig resultatene. I retrospektiv ser vi at vi kunne ha spurt bestyrerne på hvilket grunnlag de valgte ut barna. Barns evne til å huske varierer fra domene til domene, og bearbeiding av kjent informasjon er raskere og bruker mindre kognitive ressurser enn bearbeiding av ukjent informasjon (Tetzchner, 2001, s. 186). Derfor kan barn som har et forhold til de ulike landene ha lettere for å feste seg ved hvor disse er plassert.

I tillegg kommer utfordringene knyttet til barns ønske om å tilfredsstille voksne som allerede er drøftet i [seksjon 3.1](#).

5.2.2 Testlederens atferd

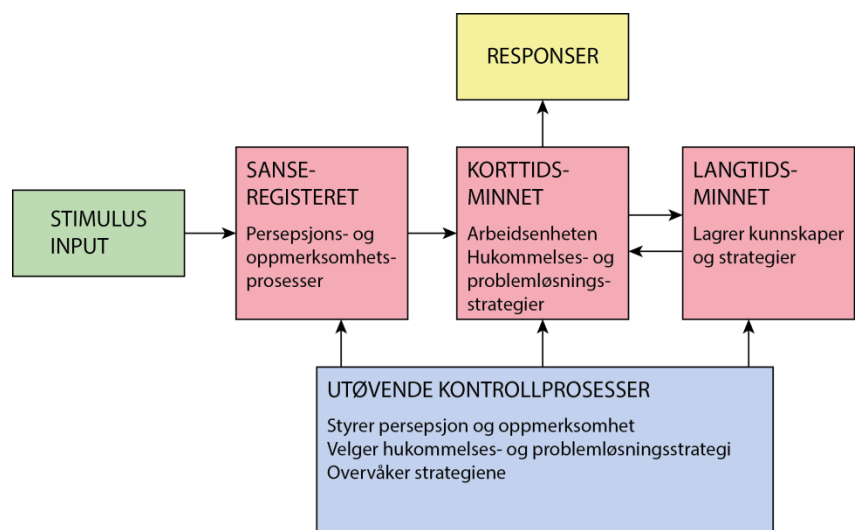
«Barn lærer bedre når den voksne er en engasjert og aktiv deltaker i barnas utforskning» (Bråten og Thurmann-Moe 1998; Henderson, 1984; gjengitt i Tetzchner, 2001, s. 301). Vi forsøkte under eksperimentene å opprettholde et konsekvent kroppsspråk, ansiktsuttrykk og tonefall, men det er realistisk sett ikke mulig å være fullstendig konsekvent. Med andre ord kan vi være en faktor som har påvirket barnas læringsprosess. Som beskrevet i [seksjon 5.2](#) har det også vært behov for å gjøre små, individuelle tilpasninger i formulering av spørsmålene.

5.2.3 Omgivelsene for eksperimentet

For å bygge videre på de eksisterende, trygge omgivelsene i barnehagen forsøkte vi å skape et eksperimentmiljø barna var komfortable med. Dette kan ha bidratt til å senke den interne validiteten. Dette gjør det vanskeligere å konkludere med at en effekt kan tilskrives den årsaken vi tror, eller om det kan være andre utenforliggende faktorer som er like sannsynlige. Vi sørget for å ha det stille rundt eksperimentet og at lyssettingen var konsistent og klar. Dette hjalp til å minimere risikoen for ukonsentrerte handlinger (Lazar et al., 2010 s. 63). Samtidig var det viktig å presisere ovenfor barna at vi ikke tester *dem*, men ønsket hjelp til å løse puslespillet (Lazar et al., 2010. s 61).

5.3 Drøfting av begrensninger ved forskningsdesignet

Den korte varigheten til eksperimentet betyr at det er vanskelig å kartlegge det faktiske læringsutbyttet hos deltakerne. Dette er en begrensning i våre data. Som nevnt under [seksjon 3.2.1](#) definerer vi læringsutbytte som ervervet kunnskap. Det kan imidlertid diskuteres om det at barna umiddelbart etter interaksjon med prototypen kan identifisere ulike land, betyr at de har ervervet denne kunnskapen. At de kan gjengi



Figur 5.1 – Informasjonsbehandlingsmodell hentet fra Evenshaug & Hallen (2000)

den der og da, betyr ikke at kunnskapen vil lagres i langtidshukommelsen. Det er først og fremst arbeidshukommelsen som trigges under arbeid med puslespillet, som når barnet leter etter en bit som passer i puslespillet (se figur 5.1).

Arbeidshukommelsen har dog begrenset lagrings- og bearbeidingskapasitet og innholdet kan derfor bli raskt redusert om det ikke blir gjentatt. Følgelig er den avhengig av ny informasjon eller informasjon hentet fra langtidshukommelsen (Tetzchner 2001, s. 209). Vi burde derfor ideelt sett latt barna leke med puslespillet gjentatte ganger over et lengre tidsrom, men på grunn av tidsbegrensninger hadde vi ikke mulighet til dette. Som Sim, MacFarlane og Read påpeker i artikkelen "*All work and no play: Measuring fun, usability, and learning in software for children*", vil det imidlertid være vanskelig å utføre et slikt langvarig eksperiment. Dette fordi det er mange ukontrollerbare variabler som kan påvirke barnas læring, eksempelvis lesing av bøker eller diskusjon i hjemmet. Vi merket oss blant annet en samtale mellom en mor og et av barna som hadde puslet, hvor moren foreslo at de skulle se på kart når de kom hjem.

En svakhet ved vårt eksperiment er at vi ikke har hatt noen kontrollgruppe. Vi kan derfor ikke si helt sikkert om en eventuell økning i antall land barna kan identifisere i testen av etterkunnskap skyldes interaksjon med noen av prototypene, eller om barna rett og slett presterer bedre fordi de er tryggere på oss og dermed tør å snakke mer åpent. Som vi tidligere har pekt på kan barn holde tilbake informasjon når de er usikre på seg selv eller situasjonen. Det at vi ikke har hatt en kontrollgruppe skyldes først og fremst praktiske og logistikkmessige utfordringer. Vi hadde begrenset med tid, hvilket gjorde det vanskelig å få tak i et høyt nok antall femåringer. Ettersom barna dessuten skulle pusle i par, innebar det at vi trengte dobbelt så mange barn som ellers per observasjon. Når vi totalt klarte å gjennomføre testing med 40 femåringer, fordelt på fire barnehager, ville vi trengt ytterligere 20 femåringer om vi også skulle hatt en kontrollgruppe, ettersom gruppene bør være av samme størrelse for at de skal kunne sammenlignes statistisk. Vi kunne i teorien latt en tredjedel av disse 40 barna fungere som en kontrollgruppe, men ettersom vi ønsket minst 10 observasjoner per vilkår, som tilsvarer 20 barn, fokuserte vi i første omgang på å nå dette målet.

En annen utfordring er hvordan vi eventuelt skulle gjennomført en observasjon med en kontrollgruppe i dette tilfellet. Vi kunne gjennomført test av forkunnskap, hvorpå vi lot barna pusle med et helt vanlig puslespill, mens vi forsøkte å interagere med dem på samme måte som med de barna som fikk interagere med prototypen. Vi kunne så gjennomført testen av etterkunnskap og sett om det var noen forskjell i antall land barna kunne identifisere. Vi håpet i utgangspunktet at det skulle holde å undersøke kunnskapsnivået i forkant av interaksjon med prototypene, siden vi tenkte på dette som en form for kartlegging av kunnskapsnivå hos barn som ikke har interagert med prototypen. I retrospektiv at det ville vært hensiktsmessig med kontrollgruppe av overnevnte grunner.

Lazar et al. beskriver forskeren som en «scientific expert»; han har ekspertise, erfaring og domenekunnskap som gjør ham i stand til å forstå de ulike aspektene ved eksperimentet og ferdigheter til å gjenkjenne det som er verdt å merke seg (2010, s.

376). Vår erfaring fra “forskningsprosjekter” begrenser seg til INF1510 og vi kan derfor ikke sies å være i besittelse av noen tilsvarende kompetanse. Dette innebærer at vi kanskje har oversett relevante aspekter ved forskningsdesignet, under datainnsamlingen og ved analysen.

5.4 Begrensninger ved prototypen

Basert på observasjonene vi gjorde da barna interagerte med prototypene under eksperimentet avdekket vi noen svakheter ved prototypen. Disse punktene vil dermed danne grunnlaget for eventuell videre utvikling av prototypen.

Lampe på selve brikken (for både en visuell og auditiv feedback)

Det kan være hensiktsmessig å ha en eller to lamper på selve brikken istedenfor en kontrollboks som på det nåværende tidspunkt er plassert øverst til venstre på prototypen. Dette for at brukeren av puslespillet skal kunne få en umiddelbar visuell og auditiv feedback uten å måtte bevege fokuset vekk fra brikken som blir lagt ned (Rogers et al., 2011, s. 26). Underveis i prosjektet fant vi ut at en slik plassering ville legge bedre til rette for læring gjennom en eksplisitt sammenheng mellom handling og visuell tilbakemelding (Rogers et al., 2011, s. 207).

Endring av feedback

I den oppgaveorienterte prototypen kan det være hensiktsmessig å endre den auditive feedbacken som gis ved riktig brikke til en som er mer bekreftende for brukeren. Ettersom både den korrekte feedbacken og feedback ved feil brikke begge starter med «Dette er x-land», vil brukeren kunne få raskere respons ved at lyden endres til for eksempel «Riktig! Dette er x-land».

6. Oppsummering

I dette prosjektet har vi forsøkt å undersøke forskningsspørsmålet vårt om *bruk av gamification kan bidra til å øke geografikunnskapene til barn*. For å gjøre dette ønsket vi å utvikle et håndfast, interaktivt system (*tangibles*) for barn i femårsalderen. Vi valgte å utvikle et fysisk puslespill fremfor en digital versjon fordi vi ønsket å legge til rette for læring gjennom lek og samarbeid mellom barn. Gjennomgang av eksisterende litteratur på området tydet på at dette var mål som bedre kunne oppnås gjennom bruk av *tangibles*.

Vi utviklet fire prototyper gjennom en iterativ utviklingsprosess. I prosessen benyttet vi oss av observasjoner av brukerne i deres reelle brukskontekst, formative evalueringer, uformelle intervjuer og “mini-workshop” med brukergruppen for å få innspill til utviklingen. Som nevnt har vi utviklet i henhold til designprinsippene, med særlig fokus på *feedback*, *constraints*, *consistency* og *affordance*. På bakgrunn av erfaringene vi fikk fra en oppurtunistisk evaluering og pilot-test designet vi et kvasi-eksperiment for å undersøke om vår endelige prototype kunne ha en positiv læringseffekt, og om barna likte å leke med den.

Vi designet prototypen med tanke på brukeropplevelsesmål som *cognitively stimulating*, *engaging*, *challenging*, *fun* og *enhancing sociability*. De kvalitative resultatene tyder på at barna som interagerte med prototypen på mange måter var

begeistret for puslespillet, både på grunn av utformingen og tilbakemeldingen de fikk. Underveis i leken så vi at de samarbeidet om å løse problemer, fikk ny innsikt og aha-opplevelser og uttrykte glede når de klarte å pusle rett. Vi observerte også barn som tidvis mistet interessen når de ikke klarte å pusle riktig, eller når det gikk lang tid mellom tilbakemelding fra puslespillet. Flere uttrykte ønske om å få pusle flere runder etter at det var ferdig, som kan være et tegn på at de likte puslespillet.

De kvantitative resultatene fra eksperimentet tyder på at interaksjonen med puslespillet førte til et læringsutbytte for begge gruppene som deltok med hver sin variant av prototypen. Selv om vi fant signifikante forskjeller i snittene mellom for- og etterkunnskap for hver av gruppene, og mellom læringsutbyttet mellom gruppene betyr ikke det nødvendigvis at funnene våre har signifikant betydning. Det er blant annet begrensninger ved utvalget vårt, både i antall og bakgrunn, som gjør at det ikke er representativt og dermed begrenser den eksterne validiteten ved våre funn. Funnene fra både den kvalitative og kvantitative undersøkelsen tyder på at det var den oppgavebaserte modulen som var mest engasjerende og som ga et størst læringsutbytte.

En annen utfordring har vært selve kjernen ved prosjektet vårt: *læring*. På grunn av tidsbegrensninger har vi ikke hatt mulighet til å gjøre tidsserier med målinger av læringsutbytte. Det er en klar svakhet at vi ikke har hatt mulighet til å gjennomføre mer omfattende undersøkelser av dette momentet, og slik sett er det vanskelig å konkludere om barna faktisk har tilegnet seg ny kunnskap.

Uavhengig av dette har vi oppnådd ny innsikt i hvordan man utvikler systemer som skal brukes av barn til læringsformål, hvordan man undersøker og forstår bruk i sin brukskontekst og metodikk for å evaluere og forstå adferd og kausale sammenhenger.

7. Referanser

- Damon, W., & Phelps, E. (1989). *Strategic uses of peer learning in children's education*. In T. Berndt. & G. Ladd (Eds.), *Peer relationships in child development* (pp. 135-157). New York: Wiley.
- Danforth, L. (2011). *Gamification and libraries*. In *Library Journal*. 136.3 (pp. 84)
- Druin, A. (2002). The Role of Children in the Design of New Technology. *Behaviour and Information Technology (BIT)*, 21(1), 1–25. Retrieved from http://www.umiacs.umd.edu/~allisond/child_info_tech/Druin-BIT-Paper2002.pdf
- Hanna, L., Ridsen, K., & Alexander, K. J. (1997). Guidelines for Usability Testing with Children, (october), 9–14. Retrieved from <https://www.comp.glam.ac.uk/blackboardAT/CS/CS3S03/CourseMaterial/Geneen/06-07/kidsusability.pdf>
- Horn, M. S., Crouser, R. J., & Bers, M. U. (2011). Tangible interaction and learning: the case for a hybrid approach. *Personal and Ubiquitous Computing*, 16(4), 1–11–11. doi:10.1007/s00779-011-0404-2
- ISO. (1998). INTERNATIONAL Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11 : Guidance on usability. Retrieved from <http://www.it.uu.se/edu/course/homepage/acsd/vt09/ISO9241part11.pdf>
- Lazar, J., Feng, J. H., & Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction* (1st ed.). West Sussex: Wiley.
- Maldonado, N. (2006). Set the Puzzles : table for. *Every Child*, 16(3), 24–31. Retrieved from <http://search.informit.com.au/documentSummary;dn=961499564013996;res=IELHSS>
- Marshall, P. (2007). Do tangible interfaces enhance learning? *Proceedings of the 1st international conference on Tangible and embedded interaction TEI 07*, 12(1), 163. Retrieved from <http://oro.open.ac.uk/19538/>
- Marshall, P., Price, S., & Rogers, Y. (2003a). Conceptualising tangibles to support learning. *Proceeding of the 2003 conference on Interaction design and children IDC 03*, (c), 101. Retrieved from <http://discovery.ucl.ac.uk/1324127/>
- Marshall, P., Price, S., & Rogers, Y. (2003b). Conceptualising tangibles to support learning. *Proceeding of the 2003 conference on Interaction design and children IDC 03*, (c), 101. Retrieved from <http://discovery.ucl.ac.uk/1324127/>
- NESH. (2006). Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi - Forskningsetiske komiteer. Retrieved October 14, 2012, from <http://www.etikkom.no/no/Forskningsetikk/Etiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/>

- Obrist, M., Igelsböck, J., Beck, E., Moser, C., Riegler, S., & Tscheligi, M. (2009). “ Now You Need to Laugh !” - Investigating Fun in Games with Children, 81–88.
- Price, S., Rogers, Y., Scaife, M., Stanton, D., & Neale, H. (2003). Using “tangibles” to promote novel forms of playful learning. *Interacting with Computers*, 15(2), 169–185. Retrieved from <http://opus.bath.ac.uk/9551/>
- Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction Design. Beyond Human-Computer Interaction* (3rd ed.). Sussex: Wiley.
- Sim, G., Macfarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education*, 46(3), 235–248. doi:10.1016/j.compedu.2005.11.021
- SRCD. (2007). Ethical Standards for Research with Children. Retrieved October 14, 2012, from http://www.srcd.org/index.php?option=com_content&task=view&id=68&Itemid=102
- Taylor-powell, E., & Renner, M. (2003). *Analyzing Qualitative Data*.
- Tetzchner, S. von. (2001). *Utviklingspsykologi. Barne- og ungdomsalderen* (1st ed.). Oslo: Gyldendal Norsk Forlag.
- Verhaegh, J., Fontijn, W., & Jacobs, A. On the Benefits of Tangible Interfaces for Educational Games. , 2008 Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning 141–145 (2008). Ieee. doi:10.1109/DIGITEL.2008.37
- Vredenburg, K., Mao, J.-Y., Smith, P. W., & Carey, T. (2002). A survey of user-centered design practice. *Proceedings of the SIGCHI conference on Human factors in computing systems Changing our world, changing ourselves - CHI '02*, (1), 471. doi:10.1145/503457.503460
- Wilson, C., Nielsen, J., Stone, D., Jarrett, C., Woodroffe, M., Minocha, S., & Kuniavsky, M. (2010). *User Experience Re-Mastered*. Burlington: Morgan Kaufmann Publishers.
- Xie, L., Antle, A. N., & Motamedi, N. (2008). Are tangibles more fun?: comparing children’s enjoyment and engagement using physical, graphical and tangible user interfaces. *Proceedings of the 2nd international conference on Tangible and embedded interaction* (pp. 191–198). ACM Press. doi:10.1145/1347390.1347433
- Xu, D., Read, J. C., Mazzone, E., & Brown, M. (2007). Designing and testing a tangible interface prototype. *Proceedings of the 6th international conference on Interaction design and children - IDC '07*, 2(1), 25. doi:10.1145/1297277.1297282

8. Vedlegg

Vedlegg A - Samtykkeskjema



Samtykkeskjema for deltakelse i pilotprosjekt for utvikling av et interaktivt puslespill

Studenter i interaksjonsdesign fra Institutt for informatikk ønsker innspill fra barn i sine prosjekter for design av interaktive, lærerike aktiviteter for barn som utvikles på oppdrag av Oslo Barnemuseum. Hensikten er å samle informasjon som kan bidra til å utvikle en prototype for museet gjennom uttesting av en prototype. Vi har utviklet denne prototypen med mål om at den skal være inspirerende og gi en morsom opplevelse, slik at barna kan utvikle kreativ tenkning, selvtillit og forståelse av verden.

HOVEDANSVARLIGE:

Vi er en gruppe studenter som tar bachelorprogrammet Informatikk: design, bruk og interaksjon ved Institutt for Informatikk på Universitet i Oslo.

Prosjektgruppa består av: Håvard Lundberg, Rebekka Soma, Martine Birketvedt Eklund, Audun Karlsrud Larsen og Denis Ligonja.

Kontakt e-post: havard.lundberg@gmail.com

Mål: Målet med observasjonen er å teste en prototype av et interaktivt puslespill med hensyn til læringseffekt og brukervennlighet.

Gjennomføring: Vi kommer til å gjennomføre en observasjon av brukerens interaksjon med prototypen, der brukeren får noen konkrete oppgaver av oss. I etterkant kommer vi til å stille noen spørsmål for å forsøke å måle læringseffekten. Hele sesjonen vil ikke overstige mer enn en time

Fordeler: Vi håper at resultatet i denne studien kan være nyttig for evaluering av aktiviteter vi har laget. Vi håper også informasjonen vi innhenter vil hjelpe oss til å forbedre aktivitetene og hjelpe Oslo Barnemuseum i å utvikle sitt tilbud.

Risiko/ubehag: Deltagerne vil ikke bli utsatt for noen form for risikofylte aktiviteter.

Taushetsplikt: All informasjon som blir samlet inn under studieundersøkelsen vil være konfidensiell. Testresultater som fremkommer gjennom testing vil være anonyme og vil ikke kunne knyttes opp mot den enkelte deltaker.

TILLATELSE:

Det vil bli tatt bilder og videoopptak under testingen. Informasjonen vil kun bli brukt i kursrapport og vil ikke bli publisert.

Jeg gir med dette samtykke til mitt barns deltagelse.

Barnets navn	Foresattes underskrift	Dato

Tusen takk for deres støtte! Ved spørsmål angående prosjektet, vennligst kontakt hovedansvarlige.

Vedlegg B – Revidert plan

Uke	Viktig	Aktivitet	Annet:
36		Idémyldring, design brief, komme med spørsmål til møtet med Katie. Utkast til prosjektplan.	INF2260 - oblig 1; INF3331
37		Spesifisere mål og idé, forskningsspørsmål, eventuelle hypoteser. Finne litteratur. Kontakte barnehager for datainnsamling.	INF1300 - oblig 1; INF3331
38		Finne litteratur. Utforme undersøkelser. Møte hos Forskningsparken Montessorri. Første observasjonsrunde i barnehagen. Prototyping (skisser).	INF2260 - oblig 2; INF3331
39		Analyse av data. Utvikling av prototype 2, "Norge i tre", test av konsept, gjennomførbarhet, funksjonalitet, materialer. Presenterer for stakeholder.	INF3331
40	Hele uka: Høstferie barnskole	Jobbet på AHO: materialtester, innkjøp, utvikling av prototype 3 ("Europa i papp", test av: design, kompleksitet, format, materialer, brukbarhet). Observasjon ved Forskningsparken Montessorri. Planlegging av evaluering av prototype 3. Gjennomføring med Montessorri og Uranienborg AKS. "Mini-workshop".	INF1300 - oblig 2; INF3331
41		Analyse av data. Utforming av hypoteser/forskningsdesign. Materialvurdering og innkjøp. Planlegging av prototype 4, igangsetting av byggeprosess (tegning, saging, Arduino-programmering).	INF2260 - oblig 3; INF3331
42		Fortsetter bygging av prototype 4 (pussing, maling, lakkering, tusje landegrensener, montere bakplate, Arduino-programmering). Forberede presentasjonen.	INF3290 - innlevering deltapport 1; INF3331
43	Presentasjon	Presentasjon. Planlegge evaluering av prototype 4.	INF2260 - oblig 4; INF3331
44		Evaluering av prototype 4.	INF1300 - oblig 3; MEVIT1700 - hjemmeeksamen; INF3331
45		Videreutvikling av prototype 4 (bytte av teknologisk plattform, nye ledninger, implementering av hint-funksjon, forbedret feedback-løsning). Pilottest. Kontakte barnehager for gjennomføring av eksperiment.	INF2260 - oblig 5; ENT1000 - hjemmeeksamen; INF3331
46		Gjennomføring av eksperiment i fire ulike barnehager. Koding og analyse av data. Finne ytterligere litteratur. Begynner på rapport.	INF3290 - innlevering deltapport 1; ENT1000 - eksamen; INF3331
47	Rapport.	Ferdigstilling av rapport. Forberede presentasjon til konkurransen.	MEVIT1110 - kvalifiseringssprøve

Vedlegg C – Spørsmålssett

Mal for spørsmål som skal stilles under eksperimentet

Deltakernummer:

Alder:

Kjønn:

Type modul:

Sted for observasjon:

Tid for observasjon:

- Test av forkunnskap begynner:
- Pusler:
- Test av etterkunnskap begynner:

Kunnskapstest:

<i>Forkunnskap</i>	<ul style="list-style-type: none">• Vet dere hva dette er (peker på puslespillet)?• Vet du hvilket kontinent det er?• Vet du navnet på noen land i Europa?• Kan du peke på de landene du kjenner til? <p>(Testen ender når barna ikke kan flere land, ettersom vi sier «La oss spille og finne ut!»)</p>
<i>Etterkunnskap</i>	<ul style="list-style-type: none">• Kan du fortelle oss hvilke land du kan/vet hvor ligger nå?

Vedlegg D – Bilder fra eksperiment



Vedlegg E – Bilder fra eksperiment



Vedlegg F – Closed sorting med post-its

