

# Kan distraktionsuppgifter ersätta tid?

## En användbarhetsstudie med inriktning på lärbarhet

Kristian Almgren, Ellen Ekström, Anna Ericsson,  
Jonatan Lidström, Camilla Niklasson, Joacim Wirebrand  
Handledare: Kristoffer Månsson och Johan Åberg  
Linköpings Universitet 2012-06-08

---

Rapporten syftar till att utvärdera systemet Commit utifrån användbarhet för förstagångs användare och användare med viss vana av systemet. Rapporten undersöker också huruvida lärbarhet kan mätas mer effektivt genom att använda en distraktionsuppgift istället för att vänta ett dygn mellan två testfällen. De faktorer som mätts för att utvärdera lärbarheten är tid, korrekthet och handlingar. De huvudsakliga problemen med systemet kopplat till användbarhet presenteras och förslag på förbättringar ges. Resultaten indikerar att det skulle kunna gå att ersätta ett dygns väntan med en distraktionsuppgift då motsatsen inte kunde bevisas. Dock har studien för låg statistisk styrka för att ge tillräckligt starka bevis för detta. Vi kan alltså inte med säkerhet säga att det går att ersätta tid med en distraktionsuppgift. För att studera detta vidare krävs studier med högre styrka.

---

## Inledning

Commit är ett webbaserat verktyg som är utvecklat för att användas av psykologer och deras patienter som komplement vid kognitiv beteendeterapi (kbt). Systemet har utformats av Kristoffer Månsson och Gerhard Andersson på Institutionen för beteendevetenskap och lärande vid Linköpings universitet. Commit är ett nyutvecklat system som hittills endast testats i en pilotstudie där psykologstudenter använt systemet vid behandling av patienter i ca 10 veckor innan denna studie utfördes. Således har ingen användbarhetsstudie tidigare utförts på Commit.

En viktig aspekt som undersöks vid en användbarhetsstudie är systemets lärbarhet, hur lätt det är att lära sig använda ett system. Lärbarhet kan mätas genom att en användare utför en viss uppgift vid två tillfällen där skillnaden i prestation visar hur mycket inläring som skett och därmed hur lätt systemet är att lära sig. Tullis och Albert (2008) presenterar två tillvägagångssätt att testa lärbarhet vid användbarhetsstudier; genom att separera testen med minst ett dygn eller med en distraktionsuppgift. Vad som inte framgår är om dessa tillvägagångssätt är likvärdiga.

## Syfte

Syftet med denna rapport är att utvärdera systemet Commit med avseende på användbarhet för nya användare i patientrollen samt mer vana användare i psykologrollen. Förbättringsförslag kommer att redovisas som avser åtgärda de allvarligaste problemen och de funktioner som flest deltagare haft problem att använda. Som en del av denna

utvärdering kommer två ansatser för att mäta lärbarhet mätas och jämföras.

Mer konkret kommer följande frågeställningar att undersökas:

1. Hur användbart är systemet Commit för förstagångs användare i patientrollen samt för psykologer med en viss vana av systemet?
2. Med avseende på mätning av lärbarhet, hur pass bra fungerar det att ersätta en dags vila med en distraktionsuppgift mellan två testfällen?

## Bakgrund

Den internationella standardiseringsorganisationen definierar användbarhet som *“the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”* (ISO 9241-11). Vid en användbarhetsstudie är det viktigt att undersöka om systemet som utvärderas uppfyller användarens behov, om systemet går att använda på ett effektivt sätt, om systemet är tilltalande att använda och om det är lätt att lära sig använda systemet (Tullis & Albert, 2008).

Lärbarhet är en vital komponent av användbarhet. Det saknas även en klar definition på vad för mått som ska användas för att mäta lärbarhet. Grossman preciserar att *“We were unable to find a single collection of learnability metrics, similar to the published collections of usability metrics”* (2009:651). Grossman (2009) har dock funnit olika aspekter att ta hänsyn till för att kunna kvantifiera lärbarhet presenterat i tabell 1. Dessa kan vidare omformuleras till att mäta lärbarhet

**Tabell 1.** Aspekter att ta hänsyn till för att kvantifiera lärlärdhet (Grossman, 2009)

Task Metrics: Metrics based on task performance	
T1.	Percentage of users who complete a task optimally. [28]
T2.	Percentage of users who complete a task without any help. [28]
T3.	Ability to complete task optimally after certain time frame. [4]
T4.	Decrease in task errors made over certain time interval. [32]
T5.	Time until user completes a certain task successfully. [33]
T6.	Time until user completes a set of tasks within a time frame. [33]
T7.	Quality of work performed during a task, as scored by judges. [9]

Någon tidigare forskning huruvida en distraktionsuppgift och ett dygns väntan är likvärdiga har inte påträffats och inte heller någon studie där dessa två mätmetoder jämförs. Ebbinghaus presenterade under sent 1800-tal teorin att glömska är en funktion av tid (Cowan et al. 2006) men enligt Fu och Liu (2007) kan en distraktionsuppgift även ha effekt på minnet. Retroaktiv interferens är ett distraktionsfenomen som innebär att om en uppgift utförs som skiljer sig från den uppgift som tidigare utfördes så försämrar förmågan att utföra den första uppgiften på nytt (Sternberg, 2009).

Utformningen för den distraktionsuppgift som valts för denna studie är inspirerad av en så kallad Word Fragment Completion Task (WFCT). WFCT har tidigare använts i minnesforskning relaterat till primingeffekter.

## Commit

Commit är uppdelat i en patientdel och en psykologdel där tillgängliga funktioner skiljer sig något åt mellan delarna. Patientsidans startsida består av åtta olika boxar, det vill säga en ikon med titel, som var och en kan expanderas till egna specifika funktioner.

Det finns även två ikoner längst ner på sidan, vilka leder till behandlaren respektive patientens kontaktuppgifter. Här kan även textstorleken ändras. Högst upp på sidan finns en rullista där patienten kan ändra behandlingsträff, samt ikoner för att gå till startsidan och logga ut.

På psykologens startsida finns fyra olika boxar, även här ikoner med titel och specifika funktioner. Utöver dessa boxar finns även en rullista med alla psykologens patienter, vilka leder till den senaste behandlingsträffen med den givna patienten. När psykologen är inne på en patients behandlingsträff är alla boxar desamma som på patientsidan med undantaget att boxen inkorg. Överst på sidan finns det även en knapp där psykologen kan lägga till nya användare/patienter.

## Metod

Ansatsen till att undersöka frågeställning 1 (hur användbart är Commit för nya användare i patientrollen, samt för mer vana användare i psykologrollen) bygger på tre huvudsakliga aspekter. Den första aspekten är resultaten av SUS enkäten och värdena som frambringades av både vana och ovana användare. Den andra aspekten är intervjuerna som psykologerna deltog i och den sista aspekten är prestationen hos de ovana användarna vid test 1.

Ansatsen till att undersöka frågeställning 2 (hur pass bra fungerar det att ersätta en dags vila med en distraktionsuppgift mellan två testtillfällen) bygger på ett experiment med en mellangrupsdesign, där forskningsdeltagarna slumpmässigt delats in i två grupper. Grupp A genomförde ett användbarhetstest (test 1), direkt följt av en distraktionsuppgift, varpå de gjorde om samma test på nytt (test 2). Deltagarna i grupp B gjorde samma test (test 1), väntade ett dygn, och kom sedan tillbaka för att göra om samma test (test 2). Skillnaden mellan de två grupperna är alltså huruvida de fick ett dygns vila mellan testen eller om de fick göra en distraktionsuppgift istället.

Mer konkret har frågeställning 2 operationaliserats enligt följande. Finns det en skillnad mellan de båda grupperna med avseende på:

- Skillnaden i korrekt utförda uppgifter mellan test 1 och test 2, uttryckt i antal rätt för samtliga uppgifter (variant av T2 i tabell 1).
- Skillnaden i tid per uppgift mellan test 1 och test 2, uttryckt i medelvärde för samtliga uppgifter (variant av T6 i tabell 1).
- Skillnaden i klick per uppgift mellan test 1 och test 2, uttryckt i medelvärde för samtliga uppgifter (variant av T7 i tabell 1).

Vår nollhypotes är att det inte finns någon skillnad i ovan nämnda punkter.

## Mätning av användbarhet

Genom att mäta användbarhet kan händelser i interaktionen mellan människa och system påvisas (Tullis & Albert, 2008) och det som framkommer visar vad som är bra respektive dåligt med ett system. Vad som ska mätas och vilka mått som användas bör därför klargöras.

- Korrekt utförda uppgifter (task success) - detta är ett mått är att undersöka om en uppgift genomförts med lyckat resultat eller ej.

- Tid per uppgift (time-on-task) – detta är ett mått på effektivitet och anger hur lång tid det tar att genomföra en uppgift.
- Klick per uppgift (efficiency) - detta är ett effektivitetsmått som anger hur många handlingar (klick) en uppgift kräver.

För att utvärdera tid per uppgift och klick per uppgift och kunna jämföra uppgifter med varandra behövs en så kallad guldstandard, vilket är det optimala resultatet för uppgiften. Guldstandarden för antal klick är det minsta antal klick som krävs för att utföra en uppgift. Detta subtraheras med antalet klick forskningsdeltagare som lyckas med uppgiften gör i genomsnitt. Guldstandarden för tiden bestämdes av att en expert som gjort testet fem gånger. Dennes genomsnittstid för en specifik uppgift subtraherades sedan med tiden det tog för forskningsdeltagare som lyckats med samma uppgift.

## Pilottest

För att testa designen för experimentet, innan de riktiga testerna, utfördes ett pilottest på fem personer som fick testa patientdelen av Commit. Denna resulterade i ändringar i instruktioner och uppgifter, samt förenkling av navigeringen på datorn. Antalet försöksledare som skulle närvara i rummet samtidigt som testet utfördes fastställdes till max tre stycken.

## Genomförande av test

För att testa nya användare på patientsidan fick forskningsdeltagarna utföra 15 uppgifter, vilka täckte de flesta funktioner i systemet. Deltagarna bestod av 28 personer och valdes genom bekvämlighetsurval. Efter test 1 fick deltagarna i grupp A göra en distraktionsuppgift följt av test 2 och grupp B fick återkomma ett dygn senare för att göra test 2. Ett program spelade in all aktivitet på skärmen och ett annat program räknade antalet musklick. Detta sammanställdes i Excel för att sedan analyseras i tre 2 x 2 ANOVA-tester, utifrån mätfaktorerna antal musklick, tid per uppgift och antalet korrekt utförda uppgifter.

## Självskattning

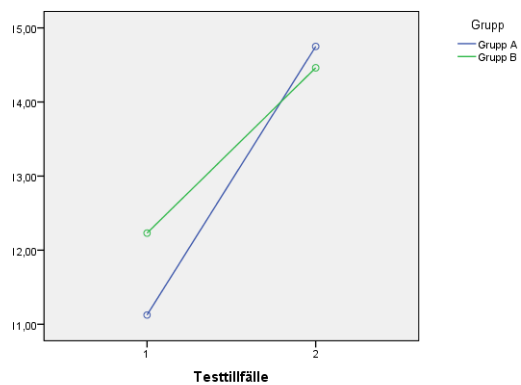
För att få ut ett självskattningsvärde på hur tillfredsställande systemet var att använda användes en System Usability Scale-enkät (SUS-enkät) som deltagarna fick fylla i direkt efter test 2. SUS ger ett värde mellan 0-100 på vilken tillfredsställelse en användare får av ett system.

## Intervjuer

Fem psykologstudenter intervjuades i en semistrukturerad intervju med fokus på hur systemet fungerar att använda för en användare med viss vana. De inspelade intervjuerna transkriberades och analyserades utifrån grundad teori, det vill säga texten som producerades kodades och kategoriserades för att få fram essensen i psykologernas svar. Utifrån vad som framkom som problematiskt med systemet sammanställdes förslag på förbättringar.

## Resultat och diskussion

### Lärbarhet och distraktionsuppgift

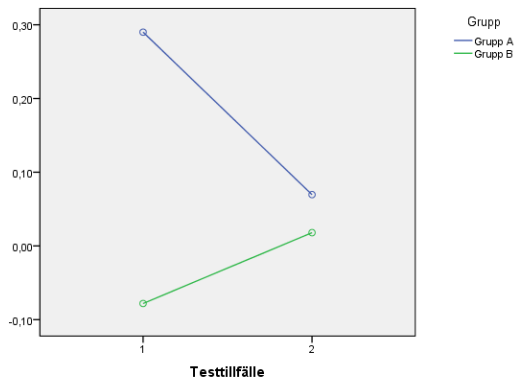


Figur 1. Korrekt utförda uppgifter

Datat från ANOVA-testet (korrekt utförda uppgifter) visar ingen statistisk signifikans med ett p-värde på 0.264 ( $F(1,19) = 1.324$ ,  $p > 0.05$ ). Vid test 1 hade grupp A ett medelvärde på 11,1 antal korrekt utförda uppgifter och grupp B ett medelvärde på 12,2. Vid test 2 hade grupp A ett medelvärde på 14,8 och grupp B ett medelvärde på 14,7.

Grafen i figur 1 visar att grupp A genomförde testet med en större förbättring mellan test 1 och test 2. Båda grupper får nära samma antal korrekt utförda uppgifter vid test 2.

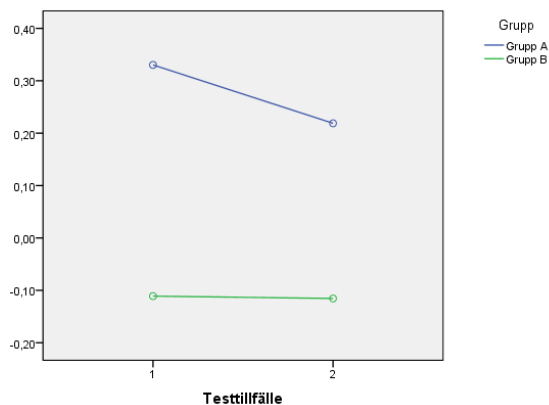
Utifrån resultatet på antal korrekt utförda uppgifter kan vi konstatera att efter ha använt systemet en gång klarar de flesta deltagare av att utföra de flesta uppgifter. Även tiden de tar för dem att slutföra uppgifter visar sig minska, så också antalet klick. Commit är med andra ord väldigt lätt att lära sig.



**Figur 2.** Tid per uppgift

Datat från ANOVA-testet visar ingen statistisk signifikans med ett p-värde på 0,095 ( $F(1, 20) = 3,066$ ,  $p > 0.05$ ). Vid test 1 hade grupp A ett z-medelvärde på 0,29 vilket motsvarar 28,7 sekunder och grupp B hade ett z-medelvärde på -0,078 vilket motsvarar 27,7 sekunder. Vid test 2 hade grupp A ett z-medelvärde på 0,07 vilket motsvarar 15,6 sekunder och grupp B hade ett z-medelvärde på 0,018 vilket motsvarar 15,0 sekunder. Se figur 2.

Grafen i figur 2 visar att grupp A har gjort större förbättring än grupp B mellan test 1 och test 2. Båda grupperna får ungefär en lika lång medeltid vid test 2. Det är viktigt att påpeka att grafen är baserad på z-medelvärden som har räknats utifrån hela populationens prestation. Därför ser det ut som att grupp B försämrar sin tid vid test 2 vilket inte är fallet. Det som grafen visar är gruppernas förbättringar jämfört med varandra, alltså har även grupp B förbättrat sin tid vid test 2 men inte gjort lika stor förbättring som grupp A.



**Figur 3.** Klick per uppgift

Datat från ANOVA-testet visar ingen statistisk signifikans med ett p-värde på 0.504 ( $F(1, 16) = 0,468$ ,  $p > 0.05$ ). Vid test 1 hade grupp A ett z-medelvärde 0,330 vilket motsvarar 6,2 antal klick och grupp B hade ett z-medelvärde på -0,111 vilket motsvarar 5,6 antal klick. Vid test 2 hade grupp A ett z-medelvärde på 0,219 vilket motsvarar 5,1 antal

klick och grupp B hade ett z-medelvärde -0,115 vilket motsvarar 4,7 antal klick.

Precis som grafen i figur 2 är grafen i figur 3 baserad på z-medelvärden. Grafen visar att grupp A gjort större förbättringar än grupp B. Här ser vi även att resultatet vid test 2 skiljer sig ganska mycket mellan grupperna.

Eftersom skillnaderna i alla tre mätningar är såpass liten vid test 2, oavsett hur prestationen såg ut vid test 1 och oavsett om det var en distraktionsuppgift eller om det var ett dygn som separerade testen, tycker vi att detta tyder på att distraktionsuppgiften vi använt kan ersätta ett dygns väntan vid mätning av de faktorer vi mätt.

Eftersom forskningsdeltagarna i grupp A och B vid test 1 har exakt samma förutsättningar borde deras resultat i teorin vara ungefär detsamma i test 1. I graferna kan vi dock se att resultaten mellan grupperna skiljer sig mer åt i test 1 än test 2. Detta kan förklaras genom ett statistiskt fenomen, så kallat regression to the mean. Detta förekommer i alla statistiska tester där två gruppers prestationer mäts där resultaten inte utfaller exakt samma vid båda testerna. Om en grupp utför två tester kommer inte de 10 % som presterat sämst i test 1 nödvändigtvis prestera sämst i test 2 (Trochim, 2006).

I vårt test kan vi tydligt se detta fenomen i grafen för tid per uppgift (figur 2) på grund av att grupp A och B slumpmässigt har delats in i grupper som presterar olika vilket påverkar vilken grupp som gör större förbättringar mellan test 1 och test 2. Om det varit grupp B som slumpmässigt varit en sämre presterande grupp vid test 1 tror vi att det kunde ha varit grupp B som gjorde större förbättringar mellan test 1 och test 2. Vi hade kunnat påverka regression to the mean genom att låta alla deltagare göra test 1 och efter det beroende på hur de presterat dela in dem i grupp A och B så att dessa gruppers medelvärde i test 1 var så lika som möjligt. Vi tror att resultaten då hade blivit mer pålitliga då resultaten för test 1 korrelerat bättre grupperna emellan. Detta är dock något som praktiskt sett är omöjligt att utföra i vår studie, då distraktionsuppgiften ska ges direkt efter test 1. Det skulle alltså kräva att vi sammanställde resultatet från alla deltagare efter första testet på 0 sekunder.

Nollhypotesen säger att det inte är någon skillnad mellan grupp A och grupp B. Vanligtvis syftar statistiska tester till att påvisa en skillnad, det vill säga att förkasta nollhypotesen. Det betyder dock inte att nollhypotesen kan bekräftas om ickesignifikant resultat uppnås, det en studie kan göra är att påvisa att det troligtvis inte är möjligt att förkasta den (Keppel, 1991).

Vi kan inte förkasta nollhypotesen då vårt test uppnår ett ickesignifikant resultat. Våra resultat saknar den statistiska styrka (power) som krävs för att bevisa nollhypotesen som ickeförkastlig. Dock ger våra resultat en fingervisning för hur resultaten kan se ut vid vilket kan vara användbart vid eventuella framtida studier inom detta område.

## Resultat av självskattning

Resultat från de SUS-enkäter som grupp A och grupp B fick fylla i gav ett medelvärde på 63,4 av 100. Medelvärdet ligger därmed inte över gränsen för det värde som Tullis och Albert (2008) hänvisade till vilket skulle tolkas som mycket bra (80) på resultat från SUS-enkäter men det ligger inte heller under det värde (60) då resultatet bör tolkas som dåligt. Även psykologstudenterna fick fylla i en SUS-enkät och där blev medelvärdet 77 av 100.

Dessa resultat visar alltså att tillfredsställelsen är relativt hög hos vana användaren som använder systemets psykologdel, men något lägre hos nya användare som använder systemets patientdel.

## Resultat från intervjuer

Fyra övergripande kategorier framkom från analys av intervjuerna:

- Vilka funktioner som använts mest och minst.  
De funktioner som huvudsakligen användes i systemet av psykologerna var att skicka meddelanden, dela uppgifter/filer, sätta upp saker på agendan och att gå in i systemet tillsammans med sina patienter. Något som inte användes var att ladda upp filer genom Ipads då det inte fanns stöd för denna funktion.
- Vad systemets primära syfte ansågs vara.  
Två av de främsta syftena med Commit ansågs vara att effektivisera behandlingen och att kunna ha kontakt med patienterna mellan mötessessioner.
- Vad som ansågs bra med systemet.  
De positiva kommentarer kring Commit som framkom var att det är lätt att lära sig, det tog ungefär lika lång tid att lära sig alla olika funktioner och att det är lättbegripligt och enkelt att förstå. Systemet ansågs uppfylla sitt syfte. Det var bra att kunna dela information via biblioteket, att den kunde delas direkt utan fördröjning, och att det var praktiskt att kunna ha allt material samlat där. Ikonerna ansågs vara tydliga och pedagogiska, och kommunikationen sågs som säker. Alla systemets funktioner sågs

inte som direkt nödvändiga, men ändå ett bra komplement i behandlingen.

- Vad som ansågs var dåligt med systemet:

### Startsidan

Psykologens startside är för lik patient-översikten, vilket försvårar att avgöra vilken sida som visas.

### Mitt bibliotek

Två boxar ("Ladda upp" och "Mitt bibliotek") har liknande funktioner på psykologsidan vilket skapar förvirring. Det är även svårt att veta om en patient har tillgång till en fil som är uppladdad i systemet. Enkäter som fanns i biblioteket var i ett filformat som inte gick att fylla i online, de behövdes alltså skrivas ut för att fyllas i.

### Uppgifter

Det skulle varit bättre om gjorda uppgifter sorteras efter något annat än datum. Vid långtgående uppgifter skulle det även vara bra att se hur långt patienten hunnit med uppgiften och inte bara om den är slutförd eller inte. Boxen uppgifter bör tituleras hemuppgifter.

### Meddelande

Funktionen att skicka meddelanden kan göras från två ställen, vilket även det kan åstadkomma förvirring. En notifiering via sms att ett nytt meddelande anlant var önskvärt.

### Design

Flera tyckte att det var svårt att veta vad en box innehöll utifrån dess titel, då namnen på dessa inte passade och var för lika varandra.

### Övrigt

Automatisk utloggning sker om systemet inte använts under en tid vilket inte fungerar så bra under behandlingsträffar. Om en patient avbokar en träff via systemet får inte psykologen någon notifiering via telefon, vilket var önskvärt.

En kommentar som sammanfattar psykologstudenternas uppfattning var att systemet ser enkelt ut men ändå inte är så intuitivt och att administrativa funktioner var svåra att förstå sig på.

## Förbättringsförslag

### Patientdelen

Följande förslag är baserade på de uppgifter som flest forskningsdeltagare misslyckades med under test 1 samt våra egna observationer.

- I målfunktionen bör knappen för att lägga till ett mål vara tydligare exempelvis med en knapp där det står "Lägg till" istället för, som det ser ut nu, en knapp med endast ett plustecken.

- I meddelandefunktionen bör användaren kunna klicka på ett oläst meddelande och endast få upp detta meddelandet, alternativt att det är tydligt när

ett meddelande öppnas vilket som är nytt respektive gammalt i meddelandekonversationen.

- Bredvid rullistan där de olika behandlingsträffarna kan väljas bör en förklarande titel finnas, till exempel "Behandlingsträff".

- "Mina uppgifter" och "Min behandlares uppgifter" är svåra att få syn på. Det kan åtgärdas genom att skriva en förklarande text under ikonerna alternativt ta bort den svarta bakgrunden som vi tror gör att ikonerna är svåra att upptäcka. En annan anledning kan vara att ikonerna är otydliga.

- En bekräftelse bör visas när en fil laddas upp under "Personligt bibliotek".

## Psykologdelen

Följande förslag är baserade på intervjuerna med psykologstudenterna:

- Psykologens egen startsida bör göras mindre lik psykologens patientöverblick.

- "Mitt bibliotek" och "Ladda upp" har liknande funktioner och skulle troligtvis kunna slås ihop till en. Det skulle till exempel kunna finnas bara ett "Bibliotek" som psykologen både använder till att ladda upp filer som han/hon vill ha tillgängliga för delning och för att kunna dela filer till patienter.

- Filer bör kunna laddas upp direkt till en patient, utan att behöva ta vägen via "Ladda upp" först.

- Uppladdningsfunktionen i "Mitt bibliotek" behöver förbättras genom att placera en "ladda-  
upp"- knapp bredvid varje fil.

- Visa i en lista vilka filer som redan har delats till respektive patient samt ge en bekräftelse på att en fil delats till en patient.

- Enkäter bör kunna redigeras digitalt, det vill säga ej vara i pdf-format.

- Markeringen som visar att en uppgift utförts bör kunna avmarkeras vid långtgående uppgifter.

- Patientens box "Uppgifter" bör byta namn till "Hemuppgifter".

- Meddelandefunktionen bör bara finnas på ett ställe.

- Det bör synas om patienten läst ett meddelande skickat från psykologen.

- Psykologen bör få ett sms när patienten avbokat en kommande träff.

- Meddelandefunktionen borde utformas så att när ett meddelande väljs är det endast detta meddelande som visas och eventuellt även en knapp som visar resten av konversationen.

- Tiden innan automatisk utloggning sker bör förlängas.

## Slutsats

Denna studie har undersökt några aspekter av användbarhet och tagit fram resultat kring dessa. Huruvida dessa resultat ses som bra eller dåliga beror på kundens krav för systemets användbarhet.

Vi anser dock att Commit ligger på en acceptabel nivå, det är inte ett optimalt användbart system men det är heller inte helt oanvändbart och det finns utrymme för förbättringar.

Det har visat sig att psykologstudenterna är nöjda med funktionerna som finns och fördelaktigt använder dem som komplement i sitt arbete. Det poängterades att de hellre vill att funktionerna förfinas snarare än att Commit utökas med fler funktioner.

Patientdelen av systemet har en höglärlarbarhet. Det är för en ny användare lätt att lära sig vilket har visat sig då resultaten närmar sig guldstandard efter endast två interaktioner.

Våra resultat indikerar att det skulle kunna gå att ersätta ett dygns väntan med en distraktionsuppgift då motsatsen inte kunde bevisas. Dock har vår studie för låg statistisk styrka för att ge tillräckligt starka bevis för detta. Vi kan alltså inte med säkerhetsäga att det går att ersätta tid med en distraktionsuppgift. För att studera detta vidare krävs studier med högre styrka.

## Referenser

Cowan N., Della Sala S., Dewar M., (2006) *Forgetting due to retroactive interference: a fusion of Müller and Pilzecker's (1900) early insights into everyday forgetting and recent research on anterograde amnesia*, Cognitive Neuroscience, Psychology, Edinburgh: University of Edinburgh.

Fu X., Liu Y. (2007). How Does Distraction Task Influence the Interaction of Working Memory and Long-Term Memory?. 7th International Conference on Engineering Psychology and Cognitive Ergonomics, EPCE 2007, Beijing, 22 July 2007 through 27 July 2007, Code 70837, vol. 4562, s.366-374.

Grossman, T., Fitzmaurice, G., & Attar, R. (2009). A Survey of Software Learnability: Metrics, Methodologies and Guidelines. *Conference on Human Factors in Computing Systems*. Boston: CHI 2009 Conference Proceedings, s. 649-658.

Keppel, G. (1991). *Design and analysis, a researcher's handbook*. New Jearsey: Prentice-Hall.

Trochim, W. M., (20-10-2006). *Regression to the Mean*. Research methods Knowledge Base, <http://www.socialresearchmethods.net/kb/regrmean.php> [30-05-2012].

Tullis T., Albert B., (2008). *Measuring the user experience: collecting, analysing and presenting usability metrics*, Burlington: Elsevier.