

UNIVERSITY NAME

DOCTORAL THESIS

Thesis Title

Author:

John SMITH

Supervisor:

Dr. James SMITH

*A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy*

in the

Research Group Name
Department or School Name

May 2016

Declaration of Authorship

I, John SMITH, declare that this thesis titled, 'Thesis Title' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

“Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism.”

Dave Barry

UNIVERSITY NAME (IN BLOCK CAPITALS)

Abstract

Faculty Name

Department or School Name

Doctor of Philosophy

Thesis Title

by John SMITH

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...

Acknowledgements

The acknowledgements and the people to thank go here, don't forget to include your project advisor...

Contents

Declaration of Authorship	i
Abstract	iii
Acknowledgements	iv
Contents	v
List of Figures	vii
List of Tables	viii
Abbreviations	ix
Physical Constants	x
Symbols	xi
1 Prototype II	1
1.1 Prototype Development Iteration II	1
1.2 Prototype Evaluation Description	1
1.3 Prototype Evaluation Methods	1
1.4 Findings	1
2 Summative Evaluation	2
2.1 Recruitment of Participants	2
2.2 Experiments	3
2.2.1 Experiment Design	3
2.2.2 Data Collection Methods	4
2.2.2.1 Family Wellness App Logs	4
2.2.2.2 Questionnaires	4
2.2.2.3 Interviews	6
2.3 Findings	6
2.3.1 Analysis of Usage Trend	6
2.3.2 User Experience of Intermediaries	9

Bibliography

12

List of Figures

List of Tables

2.1	Usage comparison between Logbook and Gamified systems for 14 pairs of users	7
2.2	Usage comparison between Logbook and Gamified systems for 11 pairs of users	8

Abbreviations

LAH List Abbreviations **Here**

Physical Constants

Speed of Light $c = 2.997\,924\,58 \times 10^8 \text{ ms}^{-\text{s}}$ (exact)

Symbols

a	distance	m
P	power	W (Js^{-1})
ω	angular frequency	rads^{-1}

For/Dedicated to/To my...

Chapter 1

Prototype II

1.1 Prototype Development Iteration II

1.2 Prototype Evaluation Description

- Location:
- Recruitment of Participants:
- Participants Demography

1.3 Prototype Evaluation Methods

1.4 Findings

Chapter 2

Summative Evaluation

2.1 Recruitment of Participants

With help of a research assistant who was a resident of Langa, we managed to recruit a total of fourteen adult participants (beneficiary users). We recruited these participants from two Cape Town's townships: Langa, and Athlone. In Langa there were six adult participants while in Athlone there were nine adult participants. The average age of these adult participants was 44.21 years with a standard deviation (S.D) of 9.99 years. The youngest adult was 26 years of age while the oldest was 60 years of age. Thirteen participants were females.

Each adult participant elected one of their children/grand children to become their intermediary user hence form a pair. Each pair was required to work together in managing the wellness of one member of a pair (a beneficiary user). All beneficiary users were working with their children but one whom was working with her grand child. The average age of children participants (intermediary users) was 15.42 (S.D=2.06) years. The youngest intermediary user was 12 years of age while the oldest was 20 years of age. The number of females and males intermediary users were equal.

I gave detailed information of what the study was all about to both intermediary and beneficiary participants. I informed them about different modes of which I will collect data. All beneficiary participants signed informed consent forms to agree to be part of the study. Since all intermediaries were under 21 years of age, they signed assent forms which were also signed by their parents/guardians who were part of the study.

I allocated one day to teach intermediary participants on how to use the "Family Wellness App". In addition, each intermediary was given a user manual. After the training, I gave out one Android phone Samsung(GT-S5300) to each pair of participants. These

phones were installed with two native apps. The first one was a pedometer and the second one was the main "Family Wellness App". The "Family Wellness App" loaded all its content from a web application hosted remotely. The app was used for a total period of six weeks. Each pair of participants provided the service provider's number of their SIM Card that will be inserted on their given Android phone. I allocated 1.3 GB of data to each SIM CARD. In addition each beneficiary participant was given a total of ZAR 240 as a compensation for transport and time throughout the duration of the study. The details of the experiments are outlined on the next section.

2.2 Experiments

This phase of the study evaluated the effectiveness of gamification/rewards in motivating both intermediaries and beneficiaries to engage with the family wellness application (Family Wellness App). I compared two versions of the applications. The first version of the application was simply a logbook or journal that allows a pair of users to record and view wellness data of a beneficiary member of the pair. With the logbook app users could view physical activity feedbacks and recording and viewing summaries of nutrition components of food consumed by a beneficiary within a pair. The second version of the application was an extension of logbook to include rewards/gamified subsystem. I carried this evaluation for a period of six weeks. I carried out the experiments from the mid October 2015 to the end of November 2015. The details of how experiments were designed and how data were collected are presented on the next sub-sections.

2.2.1 Experiment Design

I used "within-group" design for my experiments. In within-group design, the same group of participants is exposed to different experimental conditions. This helps to minimize the number of groups needed to test hypotheses as only one group is used for control and intervention. Another advantage of within-group design is that it minimizes the effect of confounding factors. The only problem with this approach is the learning effect and it lengthens the duration of the study. In order to minimize the impact of the learning effect on the outcome, I randomly assigned pairs of participants to two groups referred to as experimental sequences. The first experimental sequence started with the "Logbook App" and finished with the "Gamified App". The second experimental sequence started with the "Gamified App" and finished with the "Logbook App". I used the following abbreviations "LG" and "GL" to refer to the first and second experimental sequences respectively.

A total of seven pairs of participants were assigned to the LG group while the remaining seven pairs were assigned to the GL group. Both groups spent the first four weeks in their first experimental conditions of which the “Logbook App” for the LG group and the “Gamified App” for the GL group. After 27 days each group was switched to a different experiment condition. The LG group started using the Gamified App while the GL group started using the Logbook App. The second phase of the experiment lasted for a total of 14 days. In the next sub-section, I provide details of how data were collected during the duration of 41 days (6 weeks) of running the experiments

2.2.2 Data Collection Methods

Data collection was a triangulation of application’s logs, questionnaires and interviews.

2.2.2.1 Family Wellness App Logs

Application’s Logs consisted of information regarding the time there were users activities on the app, the pair that was accessing the app, and the functionality that was being accessed by that pair. Logs were categorized to their respective experimental condition.

2.2.2.2 Questionnaires

I administered questionnaires at the baseline, mid-line (during switching of experimental conditions), and end-line. The list of questionnaires is provided below.

Baseline Questionnaires

At baseline both intermediary and beneficiary participants filled their respective questionnaires.

- **Intermediaries:** Intermediaries participants’ baseline questionnaire had three sections. The first section captured demographic information such as age, gender, and number services/apps used on cellphones. The second section included an IMI (Intrinsic Motivation Inventory) questionnaire to assess participants’ intrinsic motivation in using cellphones. The third section included an IMI questionnaire to assess participants’ intrinsic motivation in helping their parents with cellphone based tasks.

- **Beneficiaries:** Beneficiary participants' baseline questionnaire had four sections. The first section captured demographic information such as age, gender, and number services/apps used on cellphones. The second section included an IMI questionnaire to assess participants' intrinsic motivation in using cellphones. The third section included an IMI questionnaire to assess participants' intrinsic motivation in self-monitoring of diet/nutrition. The third section included an IMI questionnaire to assess participants' intrinsic motivation in self-monitoring of physical activity.

Midline Questionnaires Also at midline both intermediary and beneficiary participants filled their respective questionnaires.

- **Intermediaries:** Intermediaries participants' midline questionnaire had only one section which included an IMI questionnaire to assess participants' intrinsic motivation in using the family wellness app.
- **Beneficiaries:** Beneficiary participants' baseline questionnaire had four sections. The first section included an IMI questionnaire to assess participants' intrinsic motivation in using the family wellness app. The third section included an IMI questionnaire to assess participants' intrinsic motivation in self-monitoring of diet/nutrition. The third section included an IMI questionnaire to assess participants' intrinsic motivation in self-monitoring of physical activity.

Endline Questionnaires At endline both intermediary and beneficiary participants filled their respective questionnaires.

- **Intermediaries:** Intermediaries participants' endline questionnaire had only one section which included an IMI questionnaire to assess participants' intrinsic motivation in using the family wellness app.
- **Beneficiaries:** Beneficiary participants' endline questionnaire had three sections. The first section included an IMI questionnaire to assess participants' intrinsic motivation in using the family wellness app. The third section included an IMI questionnaire to assess participants' intrinsic motivation in self-monitoring of diet/nutrition. The third section included an IMI questionnaire to assess participants' intrinsic motivation in self-monitoring of physical activity.

I developed the IMI questionnaires with guidance of materials found on "Self-Determination Theory"¹ website which is maintained by the researchers working on the theory including Richard Ryan and Edward Deci (Deci and Ryan, 1985) whom were early pioneers in

¹<http://www.selfdeterminationtheory.org/intrinsic-motivation-inventory/>

developing the theory. I pretested these questionnaires during the informative evaluation of prototype II in chapter 1. The most important sub-scales for our theoretical construct were perceived competence and autonomy which are part of the three basic psychological needs. The relatedness sub-scale is not yet validated but it was included in all questionnaires. Other sub-scales that were included in all questionnaires are perceived enjoyment, and perceived usefulness. In each question from the IMI sub scales, respondents were supposed to rate their experience in a scale of 1 to 7 points which means that 1 implies the statement is "not true at all" and 7 means the statement is "very true".

Perceived enjoyment is the only direct measure of intrinsic motivation while perceived competence and perceived autonomy are predictors of intrinsic motivation. Self-Determination theory suggests that a behaviour can be started as externally motivated and if external motivators support the three basic psychological needs which are relatedness, competitiveness, and autonomy then a behaviour that was once externally motivated can be internalized and users will start doing it because it is a good thing to do but not because of external rewards (Deci and Ryan, 1985). Perceived usefulness measures this level of internalization.

2.2.2.3 Interviews

I also conducted short unstructured interviews at midline and endline. I only picked fewer intermediaries and beneficiaries for the interviews. Interview responses were important in supplementing data collected through questionnaires and application's logs.

2.3 Findings

There are two primary outcomes in analysing the findings. The first primary outcome is usage of both systems. The second primary outcome is user experience/intrinsic motivation in using the both systems. There are also secondary outcomes and these are: intrinsic motivations of beneficiaries in using the wellness application; self-monitoring of diet/nutrition; and self-monitoring of physical activity.

2.3.1 Analysis of Usage Trend

The average number of days on which pairs used both versions of the application was 10.5 (SD=7.39) days. The most active usage was from a pair that utilized the app for a

total of 26 days. The less active usage was from a pair that had used the app for only two days out of 41 days.

I compared usage between a gamified system and logbook system by counting both the number of days and total number sessions between the two experimental conditions. Since the number of days and total sessions were relative to how many days users were exposed to a particular experimental condition, I divided the number of sessions with total days on which users were exposed to a particular experimental condition to get the ratio of days with respect to the total number of days and number of sessions per day. Therefore, comparison was based on a relative number and not an absolute. For instance, lets say pair one is from a LG group which they had access to a logbook system for 27 days and a gamified system for 14 days. Suppose the number of sessions in Logbook is 50 and the number of sessions in Gamified system is 30 then the relative number of sessions per day spent on logbook is 50 divide by 27 days and the relative number of session spent on a gamified system is 30 divide by 27.

Before comparing the two versions of the system on usage, I had to test if the differences of both the ratio of days and number of sessions per day between the two systems followed a normal distribution in order to identify a proper statistical test to use. The difference between logbook and gamification didn't have a normal distribution in both the ratio of days and number of sessions per day. I applied a natural logarithm transformation on the data and the difference of logs of both cases followed a normal distribution. To test for a normal distribution I used "Shapiro-Wilk Normality Test"².

TABLE 2.1: Usage comparison between Logbook and Gamified systems for 14 pairs of users

Log mean	Logbook App	Gamified App
Ratio of days	M=0.16;SD=0.15	M=0.24 ;SD=0.18
	t(13)= 1.4927 ; p=0.1594 ; 95% CI= -0.184 to 0.034	
Number of sessions/day	M=0.22 ;SD=0.22	M=0.36;SD=0.33
	t(13)=1.5519 ; p=0.1447 ; 95% CI= -0.341 to 0.056	

I carried out a student t-test on logarithmic transformed data. On comparison of usage between a gamified system and logbook system, the former had a higher mean log of ratio of days and mean log of number of sessions per day compared to the latter without statistical significance. The results are shown on Table 2.1 above.

²<http://sdittami.altervista.org/shapirotest/ShapiroTest.html>

There were three pairs with cases of technical problems with the app and this affected both their motivation and ability to participate or access the system. **Pair A** from the “GL” group had a problem with the app where it failed to load every time their intermediary user tried to start it. What I observed from the house where this pair lived in is that there was a poor internet signal hence the app was always failing to load most of the time. The other two pairs (**Pair B** and **Pair C**) from the “LG” group had their pedometer not transmitting steps data that were crucial in advancement of points and badges in gamification. One pedometer never transmitted any readings to the server. While the other one stopped transmitting on the fourth week of running experiments and this was before this particular user was switched to gamification condition. These two users with pedometer problems were close friends and they started doing informal comparison to each other even before they were switched to gamification condition. Since gamification depended on transmission of steps to the server, pedometer problems affected their motivation to participate in gamification phase although they were active in logbook phase. The learning effect coupled with problems with their pedometers mediated their decrease in motivation to continue engage with the app. A different user from a pair that didn’t have any problems with the app happened to live close to the two aforementioned users who problems with their pedometer, shared her concerns about the app during interviews. She didn’t appreciate her advancement in badges because she admitted that her peers (the two users) did more efforts than her but they were not getting anything so she didn’t understand why she was ahead of them. She was referring to their usage during logbook condition as they were both using the “Logbook App” at the same time. This proves that usability problems played a role to the some extent in demotivating participation in gamification condition of intermediary users from **Pair B** and **Pair C**. They only used the app for one day while in gamification condition.

I repeated the above analysis on usage without the three aforementioned pairs(**Pair A**, **Pair B** and **Pair C**). The results of the analysis are presented on Table 2.2. A student t test demonstrated that the means log of both the ratio of days and number of sessions per day were significantly higher in the Gamified system compared to the Logbook system.

TABLE 2.2: Usage comparison between Logbook and Gamified systems for 11 pairs of users

Log mean	Logbook App	Gamified App
Ratio of days	M=0.14;SD=0.14	M=0.27 ;SD=0.18
	t(10)= 2.5893 ; p=0.0270 ; 95% CI= -0.246 to -0.0177	
Number of sessions/day	M=0.18 ;SD=0.20	M=0.43;SD=0.34
	t(10)= 2.7479 ; p= 0.0206 ; 95% CI= -0.442 to -0.046	

2.3.2 User Experience of Intermediaries

I examined user experience through IMI(Intrinsic Motivation Inventory) questionnaires and interviews. The IMI questionnaire that assessed intrinsic motivation in using the “The Family Wellness App” had five sub-scales which were perceived competence, perceived autonomy, perceived enjoyment, perceived usefulness, and perceived relatedness. Our hypotheses of interest for intermediaries were

1. Hypothesis 1

- H_0 : There is no difference in perceived competence in using between a Logbook app and Gamified app
- H_A : There is a difference in perceived competence in using between a Logbook app and Gamified app

2. Hypothesis 1

- H_0 : There is no difference in perceived autonomy in using between a Logbook app and Gamified app
- H_A : There is a difference in perceived autonomy between a Logbook app and Gamified app

3. Hypothesis 3

- H_0 : There is no difference in perceived enjoyment in using between a Logbook app and Gamified app
- H_A : There is a difference in perceived enjoyment in using between a Logbook app and Gamified app

4. Hypothesis 4

- H_0 : There is no difference in perceived usefulness in using between a Logbook app and Gamified app
- H_A : There is a difference in perceived enjoyment in using between a Logbook app and Gamified app

The “Gamified App” was designed in such a way that a pair will earn rewards based on usage points and the average number of steps walked by a beneficiary participant who is a member of the pair. The purpose of rewards was to foster users’ intrinsic experiences such

as competitiveness and a sense of autonomy which are predictors of intrinsic motivation. A gamified system had a detrimental effect on intrinsic motivation of some users. These users reported higher scores of enjoyment while using the “Logbook App” but their scores were lower when they used the “Gamified App”. These included two intermediary users from the two pairs (**Pair A** and **Pair C**) out the aforementioned three pairs that had technical/usability problems. The third user from **Pair B** whom also had a problem of with his beneficiary’s pedometer, only experienced technical problems from the fourth week prior to switching to gamification hence he already had some steps data in the system that helped to put him on top of other users on the score board therefore he reported a higher perceived enjoyment in a “Gamified App” compared to a “Logbook App”. So I considered the latter intermediary user as some one who had a positive experience with the “Gamified App” although he had used it for only one day. The two intermediary users from **Pair A** and **Pair C** had a severe negative experience as the result of usability problems.

Apart from the aforementioned users who had negative experiences as the result of usability problems, there was a very intriguing phenomenon from other two intermediary users from **Pair D** and **Pair E**. These two users had used the app more often in gamification condition compared to when they were in logbook condition but had reported lower scores in perceived enjoyment when in gamification condition compared to when they were in Logbook condition. Two of these intermediary users were in LG and GL groups respectively. I examined the performance of these two users in gamification rewards and discovered that two users never managed to make any progress in attaining a single reward despite their efforts in using the “Gamified App”. This harmed both their perceived enjoyment and perceived competence to use the “Family Wellness App” while in gamification condition. One of these two intermediary participants sent an SMS to the researcher asking what he was supposed to do in order to advance in badges and he was informed that all the information was specified in their given user manual. In addition, SMS reminders were sent out to all intermediaries to inform them of how far they have gone in achieving rewards and what is remaining in terms of usage and steps in order for them to reach the next badge. Their beneficiary participants were not walking enough steps despite the fact that these intermediaries had put efforts in engaging with the App. One beneficiary who was working with one of these intermediary users also reported in interviews that there was always a contention with her intermediary when this beneficiary wanted to see what was going in the app as the intermediary was not voluntarily willing to help some of the time.

On comparison between Logbook and Gamification on IMI sub-scales, I excluded a total of four intermediary users. Two out of those four intermediaries had a severe negative experience as the result of usability problems, whereby one them, the app failed to load

when he attempted to use it at the beginning of the study, and another one whom his pedometer failed to transmit any steps from the beginning of the study. The remaining two had a negative experience as the result of failure of our gamification design to match challenges with abilities. i.e efforts of beneficiaries differed hence challenges should match abilities of a pair. When challenges are too difficult as they don't match users' skills, end users can become demotivated ([Zhang, 2008](#)). As result only intermediary from ten pairs were considered for analysis out of intermediaries in 14 pairs. I did a comparison on perceived competence, perceived autonomy, perceived enjoyment and perceived usefulness of between the "Logbook App" and the "Gamified App". The distribution of the differences of all the scores from the four aforementioned sub-scales, followed a normal distribution hence met a condition for using a student t-test.

Bibliography

- Edward L Deci and Richard M Ryan. *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media, 1985.
- Ping Zhang. Technical opinion motivational affordances: reasons for ict design and use. *Communications of the ACM*, 51(11):145–147, 2008.