

# Gamifying Online Marketplaces to Overcome Supply and Demand Imbalances

a Design and Evaluation Study at OLX India

Bauke Buikstra

Delft University of Technology



# Gamifying Online Marketplaces to Overcome Supply and Demand Imbalances

*a Design and Evaluation Study at OLX India*

by

**B.J. (Bauke) Buikstra**

in partial fulfilment of the requirements for the degree of

**Master of Science**

in Systems Engineering, Policy Analysis & Management

at the Delft University of Technology,

to be defended publicly on Tuesday June 23, 2015 at 15:00h.

## **Author**

*Name* Bauke Buikstra  
*Student Number* 1521136  
*E-mail* b.buikstra@outlook.com; b.j.buikstra@student.tudelft.nl  
*Faculty* Technology, Policy & Management – Delft University of Technology

## **Graduation Committee**

*Full Professor (chair)* Prof. Dr. Ir. Alexander Verbraeck, dep. Systems Engineering, TPM, TU Delft  
*First Supervisor* Dr. Rens Kortmann, dep. Systems Engineering, TPM, TU Delft  
*Second Supervisor* Dr. Laurens Rook, dep. Economics of Technology and Innovation, TPM, TU Delft  
*External Supervisor* Eric Klaassen, BLOOM – online business development

## **Document**

*Date* 16-06-2015  
*Number of words* 72,128  
*Keywords* Gamification, gameful design, controlled randomised experiment, online marketplace, user-generated content, network effects

*An electronic version of this thesis is available at <http://repository.tudelft.nl/>.*

## **Gamification**

Noun - gam·i·fi·ca·tion - \,gā-mə-fə-'kā-shən\

*“A cynical practice by corporate douches where workers are supposedly motivated to work even harder on menial, pointless tasks by rewarding them with lame titles, meaningless rankings, coupons or a variety of other real-life trash loot.”*

Manylaugh

July 10, 2012

Urban Dictionary (website)

<http://www.urbandictionary.com/define.php?term=Gamification>

---

## **<front page photograph>**

‘Indian Street Market Chaos’

Robin Baumgarten

Amritsar, India – January 6, 2013

<https://www.flickr.com/photos/robinbaumgarten/8402189946/in/set-72157632374090212>

Creative Commons License

## PREFACE & ACKNOWLEDGEMENTS

---

Thank you for being interested in my work and reading at least this sentence of the preface. Hopefully you will continue all the way towards the end and become one of the honourable few who can say they have read my entire MSc thesis (who knows, there might be a prize). It's been quite a journey, with ample amounts of time, energy and thoughts spent – some of it wisely and some of it a little less wisely. I am glad to have a page on which I can write some thoughts from my own perspective, so I will try to make good use of it.

First of all: this is my MSc thesis report, describing a research project conducted at OLX India, with help of BLOOM (where I wrote this thesis as a graduate intern), in order to complete the programme Systems Engineering, Policy Analysis and Management at the faculty of Technology, Policy and Management of Delft University of Technology. This thesis is accompanied by an article, which was submitted to the CHI PLAY 2015 conference. Now that this is clear and you know what's coming, I would like to say some thanks and share some of my thoughts; selected from the messy, fuzzy, but also crystal clear, nostalgic, proud and slightly apocalyptic cocktail of seemingly random brainwaves that my mind has become at the end of this thesis-writing-experience.

Thank you Rens, for meeting with me on a regular basis and being my main sparring partner, not only content-wise (by suggesting papers that became the most influential for this research), but also process-wise. Laurens, thank you for being critical and laughing at my ignorance now and then. Thanks also for helping me out with my methodological quest, especially concerning the statistical tests, even though I did not really stick with your advice in the end ☺. Alexander, thank you for cutting away more than half of the uncertainty and paving way for the main research structure during every meeting we had. The three of you have been a role model graduation committee, as far as I'm concerned. Also, even though they probably will not read it, thanks to Ivo Wenzler and Igor Mayer, for introducing me into the world of serious games and gamification. Overall, the TPM faculty has offered me an excellent BSc and MSc programme. I truly feel that I am ready to 'go out there'.

Speaking of which, I owe many many thanks to the BLOOM team. Eric, you made this research possible and helped me realise what I want to do for the coming few years of my professional career. You were a great help and advisor in dire times. Merlin, Jim, Onno, Timo, Steven and Michelle: thanks for letting me do my thing at the office and listening to the stories about my graduation perils.

OLX is a good example of a company with a whole lot of interesting potential to serve as a case for academic research and has proven to be very cooperative and brave to team up with me. Miguel, Casper and Bram, you have been very kind and open to help me out. The India team: Vineet, Sarabjit, Rakesh, Deepali and especially Akanksha, I'm grateful for your patience, work and the opportunity I got to conduct such a big experiment. I hope some of my work helps you in developing the concept of second-hand goods marketplaces all over the world.

Furthermore, thanks to my family, friends and Lotte for being interested (or at least pretending to be) in my research and the progress, as well as watching out for my physical and mental health by providing me with the necessary real life relaxation moments (and not behind a PC, eating out the internet while procrastinating endlessly and pretending to have 'a short break' from typing). Finally, I must say my respect for all fellow students, researchers, PhD candidates, professors and also writers in general has grown tremendously. It's a special process.

Keep calm and game on,

Bauke Buikstra

## EXECUTIVE SUMMARY

---

### Challenges for Online Marketplaces

In recent years, several traditional economic models are being challenged by start-ups with disruptive new business models, where offer and demand are connected in a flexible and direct way, through an online platform. Consumers interact with each other and form both sides of the market. Airbnb is a commonly known example, as well as Uber, Kickstarter, Snappcar and Craigslist. Their main value for society is that the search and/or transaction costs for participants are considerably reduced (Hagiu, 2014; Seamans & Zhu, 2014). These are online marketplaces: an online platform that purely facilitates communication of two user sides through online content created by these users. Online marketplaces rely on online content provided by the supply side of their user base. This content is generally called user-generated content (UGC) (Albuquerque, Pavlidis, Chatow, Chen, & Jamal, 2012; Bakos & Katsamakos, 2008; Khatibloo, 2011). Characteristic for these online marketplaces is the chicken-egg problem that occurs, often in an early stage. Caillaud and Jullien describe the chicken-egg problem as follows: “to attract buyers, an intermediary should have a large base of registered sellers, but these will be willing to register only if they expect many buyers to show up” (2003, p. 310). The chicken-egg problem can hinder platform growth and is one of the main issues for online marketplaces (Armstrong, 2006; Caillaud & Jullien, 2003; Eisenmann, Parker, & Alstyne, 2006; Hagiu, 2014). The underlying phenomena for the chicken-egg problem are network effects, meaning that the supply and demand sides are always interdependent and can both limit and stimulate each other to grow (Katz & Shapiro, 1985). Online marketplaces heavily rely on (positive) cross-side network effects for growth, where the content provided by sellers is usually the limiting factor and thereby affects the increase of buyers (Eisenmann et al., 2006; Jordan & Hariharan, 2015). Sellers have to supply user-generated content, for which they have to be able and willing to do so. On the long term, after overcoming the initial growth phase, there is a continuous imbalance between the two sides, which needs to be resolved in order to grow. Stimulating growth of the user side lacking in numbers is vital (Albuquerque et al., 2012; Armstrong, 2006; Bakos & Katsamakos, 2008; Cambini et al., 2011; Eisenmann et al., 2006; Goldsmith, Pagani, & Lu, 2013; Hagiu, 2014; J.-C. Rochet & Tirole, 2003).

### Gamification and its Possibilities

Furthermore, ‘gamification’ has recently gained attention. A simple and illustrative definition of gamification is: “*the use of game design elements in non-game contexts*” (Deterding, Dixon, Khaled, & Nacke, 2011, p. 7), also known as the traditional ‘elemental definition’. Where serious games combine all types of gaming elements into a whole, in order to provide an immersive and ‘other world’ experience, gamification uses the strategies of persuasive technology to affect behaviour during a task, thus serving a facilitating role in a service a product, rather than the game elements being the core of it (Deterding et al., 2011). The main potential and added value of gamification is the increased user engagement because of the elements of play, fun and competition, which can be added by gaming, creating a ‘gameful experience’. Increased engagement can result in an increased or transformed motivation to, depending on the system where it is implemented, participate, use, learn, have social interaction or perform tasks (Deterding et al., 2011; Groh, 2012; Hamari, Koivisto, & Sarsa, 2014; Hense et al., 2014; Huotari & Hamari, 2012; Nicholson, 2012; Thiebes, Lins, & Basten, 2014; Werbach & Hunter, 2012). When looking back at the main challenge for online marketplaces, the potential of gamification to aid in this challenge can easily be discerned. UGC is driven by user activity, where the decision to place content is not only dependent on the demand for content (the network effect) but also on the level of engagement of a user with the online marketplace. The users themselves must supply UGC and be intrinsically willing to do this. The main potential and added value of gamification fit well into the main challenge of online marketplaces: engaging users with a gameful experience, in order to stimulate their activity and post their content on the platform.

However, the gamification field contains strongly divided opinions, definitions and movements (Deterding, 2014a; Groh, 2012). Some see it as a re-invented marketing tool to trick and exploit

customers in order to increase profits, while others speak of a true enhancement of user engagement by creating valuable gameful experiences (Farrington, 2011; Hamari, 2013). A lack of scientific literature, a lack of research results and its young age indicate that gamification is currently still in its infancy. The design principles are scattered and diverse, there are multiple definitions, the number of valid evaluation studies is scarce, there is a rampant growth of self-proclaimed gamification experts and some have deemed gamification as a marketing buzzword. This is also commonly recognised within the main gamification literature, where the need for more empirical research on gamification design and evaluation is expressed (Deterding et al., 2011; Deterding, 2014c; Groh, 2012; Koivisto & Hamari, 2014; Thiebes et al., 2014).

### Research Question

The potential combination of gamification and online marketplaces, in the context of network effects and forthcoming chicken-egg problem, has not been explicitly identified before. It is worth investigating, as it is an interesting combination of a relevant societal subject and a promising but disputed method. The main research question that follows the research problem and corresponding research goal is: *what is the suitability and capability of gamification to increase the amount of user-generated content on online marketplaces?*

### Theoretical Perspective on Gamification

Sebastian Deterding published work in which he states: “I suggest expanding the remit of gamification from the structuring of objects to the framing of contexts, and from game design elements to motivational affordances” (2014a, p. 307). He proposes to combine the existing elemental (Deterding et al., 2011) and experiential perspectives (Huotari & Hamari, 2011) into a more socio-technical view, which recognises motivational affordances and also the user context, the design process and motivational factors: “Understood as such – a unified whole of restructuration and reframing – gamification is a holistic socio-technical systems design practice [...] one that understands humans interacting with technology as assemblages, activity systems, or ecologies of heterogeneous and intertwined actors” (Deterding, 2014a, pp. 312–313). This view was adopted throughout this research, including the related gamification design method (Deterding, 2014c), since it is the only available scientifically published method. Also, it is very valuable from a theoretical perspective, because it synthesises existing design methods and an array of gamification theory into one.

### Gamification Treatment Design

The design method was applied to a case: the mobile website of OLX India, which is an online marketplace where users can buy and sell second-hand goods. The marketplace copes with the network effects problems, in the sense that it is trying to grow its user base and the share of users that posts listings (creates content) is very low. Increasing the number of listings will also have a positive effect on the other type of users (buyers). Therefore, the challenges of OLX very closely resemble the more general online marketplace problems that are mentioned in literature (Cambini et al., 2011; Hagiu, 2014; Seamans & Zhu, 2014). The main goal for OLX was to increase the number of new listings posted per user. The main issue for OLX users is that they do not know what to sell. They do not think they have objects of value to others in their home, thus do not post a listing. This was derived from OLX internal reports, interviews with OLX employees, a workshop with OLX users and a workshop with an OLX expert. The design method was focused on resolving this challenge and through workshops with ideation and iterative prototyping, a final design was created: the Selling Assistant. The idea is that appealing explicitly to the identified challenge with a call to action (for instance: ‘I don’t know what to sell’) will engage users. Breaking the challenge into smaller steps by providing guidance and a limited number of choices for items to sell originates from the game design lenses used, which are part of the design method (Deterding, 2014c). The next best action is to select an item from a small list, rather than to ‘sell something’. For each category featured in the list, the number of ‘average views per listing’ is shown, which indicates the demand for a new listing in that category. By showing this number to users, they might be convinced of new products, which are

eligible for sale on OLX. On the long term, this mechanism can be used by OLX to drive the number of listings specific categories and effectively matching supply and demand dynamically.

### Assessing the Treatment Effects in an Experiment

To evaluate the Selling Assistant an online, double blind, randomised controlled experiment (Ron Kohavi, Longbotham, Sommerfield, & Henne, 2008) was used, because is used in most scientific gamification evaluation studies (Denny, 2013; Farzan et al., 2008a, 2008b; Hamari, 2013). The experiment was live for exactly 7 days. 51,103 OLX users were randomly selected to be included in the experiment and randomly divided over an Original variant (control group), a Suggest variant (with 1 version of the Selling Assistant) and a Know variant (with another version of the Selling Assistant).

The Selling Assistant caused an increase in the amount of UGC. Suggest and Know yielded respectively 18% and 19% more new listings than the control group in the Original variant. These percentages correspond with the amount of listings that were posted through the Selling Assistant pages, suggesting that the treatment actually helped users in selecting an item to sell. Within both the Suggest and Know variant, users who actively engaged with the Selling Assistant posted more than 6 times as many listings on average during the experiment than users who did not engage with the Selling Assistant, controlling for other predictor variables such as location, browser, number of page views and source. However, because only 2.1% and 2.4% of the users interacted with the Selling Assistant, the overall differences between the Selling Assistant groups (Know and Suggest variants) and the control group (the Original variant) were not significant. The extra productivity of users who engaged with the Selling Assistant contributes only in a small degree to the overall number of listings, especially considering the fact that the vast majority of users did not post a listing at all. Thus, the results of the experiment in this study suggest that the gamification treatment was effective only for a small proportion of users, who either coped with exactly the challenge that the treatment tried to solve and/or are willing to actively interact with such a treatment. Almost exactly the same was by concluded by Juho Hamari, who implemented badges as game elements (using a more traditional gamification approach) on an online marketplace (2013). The question as to whether these interacting users within Suggest and Know were stimulated by the Selling Assistant to post more listings or start posting a listing in general (become a lister) could not be further filled in. Looking at new and returning visits of users during the experiment learns that users generally post a listing in a returning visit, not in their first visit.

### Limitations and Avenues for Further Research

Some research limitations were identified, regarding the experiment and the qualitative design method application and evaluation. They mostly consist of minor issues, which do not compromise the possibilities of answering the research questions. The limitations were used as input for the future research recommendations. First of all, measuring only the behavioural effect of a treatment that was made with a user-centred design method seems, in hindsight, like an incomplete evaluation of its effect. Not only a behaviourist view but also a cognitivist view on gamification is important: find out what the design did with people, did they like it? How did their intentions and attitudes change? This way, the effect of the treatment can actually be connected to changes in motivational elements of competence, relatedness and autonomy of users (Deterding, 2014a; Groh, 2012; Ryan & Deci, 2000; Werbach, 2014b). A mixed methods approach, including qualitative observational data, psychometrics and historical data on user behaviour, was not possible in this study. However, this is strongly recommended for future research. A trade-off has to made between working with a small user base that is aware of being in an experiment (not double-blind, so more bias) and being able to do psychological measurements, versus working with a large user base in a double-blind experiment, measuring only behavioural data. Furthermore, it is recommended to use a second control group with a dummy intervention and to have an experiment timeline of at least two weeks, to correct for 'novelty' or 'Hawthorne' effects (users displaying an inherent positive attitude towards new features) (Hamari et al., 2014; Ron Kohavi et al., 2008; Lieberoth, 2014). Also, questions remain regarding the type of users that was mostly affected. This results in the recommendation to further investigate user

segments and their different responses to gamification, for example by segmenting on demographics (Koivisto & Hamari, 2014; Lieberoth, 2014). A last and important point to investigate further is the difference of applying gamification from an elemental (Deterding et al., 2011), experiential (Huotari & Hamari, 2012) or 'new' process (or socio-technical design) perspective (Deterding, 2014a; Werbach, 2014a), preferably within the same case. The fact that there are multiple views and methods within gamification could be a positive development, in the sense that gamification is evolving from a single method into a research field. However, it also creates confusion, because the term gamification refers to multiple things, making it harder to identify relevant scientific literature. Related to the previous point, this research could be validated more by applying the same design method again in a different case, with the same general subject and context. Thorough documentation of the design method and process in this research allows easy reproduction within another case and increases comparability of future empirical gamification studies, which is currently a problem (Deterding, 2014a; Farzan & Brusilovsky, 2011; Groh, 2012; Hamari et al., 2014; Nicholson, 2012; Thiebes et al., 2014).

## Conclusions

Now, the capability and suitability of gamification to help solve imbalances on online marketplaces (by stimulating the amount of UGC) can be assessed, in order to answer the main research question.

Regarding the suitability of gamification to online marketplaces, Deterding provided a structured design method, which fits quite seamlessly into the way of working of OLX and many other online marketplaces. Its elements of continuous improvement and iterative prototyping, as well as frequent user contact with user-centered design, are well known concepts for such organisations (Klaassen, 2014; Ries, 2011). An added value of the design method is its way to connect user motivations to activities, which forces the designer to connect with the marketplace user and to structurally assess the marketplace from a game designer perspective (with novel insights and ideas that lie outside the normal thought realms of most marketplace employees). The design method synthesises an existing variety of methods and design practices into one and is able to cope with most of the known gamification criticism. This method creates a potential single starting point for structured gamification research, which can be compared and evaluated on the same grounds. Studies using the design method are more comparable than studies using the game element approach, because game elements are infinite, not mutually exclusive and highly variable across available gamification theory sources.

Looking at the capability of gamification to increase the amount of UGC on online marketplaces, the results of the experiment are matched with literature on the network effect related problems that online marketplaces cope with. Once a problem is identified, gamification allows for very specific motivational affordances to be implemented, stimulating the users who cope with exactly this problem. When the problem or challenge that is focused on is very delineated, the result will likely also be very delineated. This makes gamification less useful for chicken-egg problems in an early stage, as the this stage requires the overall user base to be increased and activated (Bakos & Katsamakas, 2008; Jordan & Hariharan, 2015; J. C. Rochet & Tirole, 2006). When looking to solve imbalances on an online marketplace in a later phase of growth, tackling specific issues for specific groups users that do not create content, due to a shared challenge they face, is a good method (Albuquerque et al., 2012; Hagiu, 2014; J. C. Rochet & Tirole, 2006). Using gamification to tackle specific problems seems more valuable in such a later stage. More specifically, a solution such as the Selling Assistant can both nudge users to create more content (or nudge more users to create content) and nudge them into the right type of content, which can specifically solve the imbalances on the platform. However, not having something to offer can be a very practical hurdle to create content on an online marketplace, regardless of a user's motivation and the trigger that stimulates him/her to want to create content (with or without gamification). This could make it harder for gamification to be successful on an online marketplace than on general online platforms such as social networks.



Overall, it can be concluded that gamification is capable - if certain preconditions regarding the marketplace growth stage and the willingness of users to interact are in place - and suitable to increase the amount of user-generated content on online marketplaces. Nevertheless, the exploratory nature of this study should not be forgotten. The adopted gamification perspective and corresponding design method are novel, within the already novel and manifold gamification field. Conclusions made should therefore not necessarily be regarded definitive, but as starting points for further research.

## TABLE OF CONTENTS

---

<b>Preface &amp; Acknowledgements .....</b>	<b>iv</b>
<b>Executive Summary .....</b>	<b>v</b>
<b>Table of Contents .....</b>	<b>x</b>
<b>List of Tables .....</b>	<b>xii</b>
<b>List of Figures .....</b>	<b>xiii</b>
<b>List of Abbreviations.....</b>	<b>xv</b>
<b>1 Introduction.....</b>	<b>1</b>
1.1 Problem Outline .....	1
1.2 Research Foundations .....	6
1.3 Report Structure .....	9
<b>2 Theoretical Perspective on Gamification .....</b>	<b>10</b>
2.1 Gamification Defined .....	10
2.2 Lessons from Past Gamification Studies.....	16
2.3 Underlying Theory of Gamification Perspective.....	19
2.4 Gamification Design Method.....	20
<b>3 Gamification Design .....</b>	<b>24</b>
3.1 Case Introduction: OLX.....	24
3.2 Applying the Gamification Design Method .....	26
3.3 Final Gamification Treatment: The Selling Assistant.....	34
<b>4 Experiment with Gamification Treatment .....</b>	<b>38</b>
4.1 Expected Effects of the Selling Assistant.....	38
4.2 Experiment Setup .....	40
4.3 Data Preparation, Validation & Statistical Methods.....	43
4.4 Results.....	45
<b>5 Discussion, Limitations &amp; Future Research.....</b>	<b>60</b>
5.1 Experiment with the Selling Assistant .....	60
5.2 Application of the Design Method .....	64
5.3 Theoretical Implications .....	66
5.4 Managerial Implications .....	72
<b>6 Conclusions &amp; Reflections .....</b>	<b>75</b>
6.1 Conclusions .....	75
6.2 Personal Reflections .....	78
<b>Bibliography .....</b>	<b>80</b>

<b>APPENDICES .....</b>	<b>87</b>
<b>A Relevant Empirical Gamification Research .....</b>	<b>88</b>
<b>B Extensive Description of Gamification Design Method as Applied in this Research .....</b>	<b>90</b>
<b>C Chronological Walkthrough of Ideation &amp; Iterative Prototyping Step in Design Method .....</b>	<b>93</b>
<b>D Selling Assistant Details .....</b>	<b>113</b>
<b>E Experiment Details .....</b>	<b>115</b>
<b>F Excluded Metrics and Hypotheses in Experiment .....</b>	<b>118</b>
<b>G Extensive Data Preparation .....</b>	<b>121</b>
<b>H Extensive Data Validation .....</b>	<b>130</b>
<b>I Justification of Statistical Distributions and Tests Used .....</b>	<b>137</b>
<b>J Extensive Results of Statistical Tests .....</b>	<b>143</b>

## LIST OF TABLES

---

Table 1: Experiment results of Original, Suggest and Know in terms of average new listings per user .....	46
Table 2: Results of Mann-Whitney U tests to verify hypothesis 1 .....	46
Table 3: Results of Mann-Whitney U tests to verify hypothesis 2 .....	50
Table 4: Results of Z-tests for proportions to verify hypothesis 3 .....	52
Table 5: Results of Mann-Whitney U tests to verify hypothesis 4 .....	54
Table 6: Differences between new and returning visits of users who clicked the home page button, split for Suggest and Know .....	57
Table 7: Overview of tested and (un)supported hypotheses .....	58
Table 8: Participant data, workshop 1 .....	99
Table 9: Data sample of experiment data, organised per user .....	128
Table 10: Data sample of experiment data, organised per page load .....	129
Table 11: Number of users with a certain number of posts, within different user segments .....	135
Table 12: Descriptive statistics of new listings per user, split per experiment variant .....	137
Table 13: Chi-Square goodness of fit test to Poisson distribution for new-listings-per-user-data per variant and for all .....	140



## LIST OF FIGURES

Figure 1: Report Structure.....	9
Figure 2: Game elements as defined by Deterding et al. (2011, p.12) .....	12
Figure 3: Gamification examples from a service marketing perspective (Huotari & Hamari, 2012, p. 20) .....	13
Figure 4: Differentiating gaming from playing, and full games from gamification. Adapted from (Deterding et al., 2011, p. 13) .....	15
Figure 5: General form of a skill atom (Deterding, 2014c) .....	22
Figure 6: The gamification design method (Deterding, 2014b) with amendments, as applied in this research .....	23
Figure 7: OLX India mobile website homepage (OLX, 2015a) .....	25
Figure 8: Behaviour chain for OLX users who want to post a listing .....	28
Figure 9: Skill atom for posting a listing on OLX mobile website .....	31
Figure 10: Ideation and iterative prototyping, step 4 of the gamification design method .....	33
Figure 11: Workshop with OLX expert (bottom right photo) and workshop with OLX users (other photos) .....	34
Figure 12: Final design of gamification treatment 'The Selling Assistant', with home page (suggest version), product page (version A) and posting page, from left to right. ....	36
Figure 13: Data preparation steps performed .....	43
Figure 14: Visual inspection of the effect of predictor variables on the mean listings per user .....	48
Figure 15: Home Page Know Variant .....	49
Figure 16: The number of new listings per user per home page variant, segmented in users who did and did not visit the product page .....	50
Figure 17: The proportion of users that is a lister, per experiment variant .....	52
Figure 18: Visual inspection of the effect of predictor variables on the mean number of listings per lister .....	53
Figure 19: The number of listings per home page variant, segmented in different listing types .....	55
Figure 20: Listings per user, only for users who visited the product page, split for user groups, segmented in different listing types.....	56
Figure 21: 95% confidence intervals for the average number of new listings per user, split per home page variant, segmented for new and returning visits .....	58
Figure 22: Results of initial ideation .....	95
Figure 23: Selling Assistant prototype version 1 .....	96
Figure 24: OLX challenge mock-up version 1 .....	97
Figure 25: Workshop 1 photographs.....	100
Figure 26: Selling Assistant mock-up version 2 .....	103
Figure 27: The OLX Challenge mock-up version 2.....	104
Figure 28: Choose Your Goal mockup version 1 .....	105
Figure 29: Workshop 2 photograph .....	106

Figure 30: Selling Assistant mock-up version 3 .....	108
Figure 31: Choose Your Goal mockup version 2, variation B.....	110
Figure 32: Choose Your Goal mockup version 2, variation A.....	110
Figure 33: Selling Assistant mock-up version 4.....	111
Figure 34: Game design lenses used for the Selling Assistant (Deterding, 2014c) .....	113
Figure 35: Full experiment workflow of the Selling Assistant.....	114
Figure 36: Sampling method of Optimizely, used in experiment, from <a href="https://help.optimizely.com/hc/en-us/articles/200040335">https://help.optimizely.com/hc/en-us/articles/200040335</a> .....	115
Figure 37: Home page variant user distribution .....	130
Figure 38: Product page variant user distribution, by home page variant .....	131
Figure 39: Number of listings posted per user during experiment .....	133
Figure 40: Number of listings per user, within different user segments.....	134
Figure 41: Number of pages loaded by users during experiment. Left: per user. Right: per user per variant.....	135
Figure 42: Number of users with a certain number of new listings posted, split per experiment variant .....	137
Figure 43: QQ-plots of variant quantiles (y-axis) against theoretical normal quantiles (x-axis) .....	138
Figure 44: Number of users with a certain number of new listings posted, split per experiment variant and theoretical Poisson distribution per variant .....	140
Figure 45: Poissonness plots for the new-listings-per-user data of Original, Suggest and Know .....	140

## LIST OF ABBREVIATIONS

---

- MSP = multi-sided platform
- UGC = user-generated content
- OLX = Online Exchange, an online marketplace allowing users to buy and sell their (second hand) goods - in this research mainly refers to OLX India specifically;
- PV = page view or page load
- DAU = daily active user
- B = regression coefficient
- SE = standard error
- p = probability value of statistical test

# 1 INTRODUCTION

---

*In this first chapter, the basic elements and the cause of the research are introduced. The current challenges for online marketplaces and scientific knowledge gaps concerning gamification are revealed, leading to the problem statement which is the subject of this research. Following the problem outline, the foundations of this study are described in terms of research goal, research questions and the research approach. Also, the scientific and societal relevance of this research are depicted. Lastly, sub chapter 1.3 explains the structure of this report.*

---

## 1.1 Problem Outline

### 1.1.1 Online Marketplaces in Society

In recent years, several traditional economic models are being challenged by start-ups with disruptive new business models, where supply and demand are connected in a flexible and direct way, through an online platform. Consumers interact with each other and form both sides of the market. Airbnb is a commonly known example: a website where home owners can offer their empty rooms or apartments for rent and people looking for accommodation can find a place to stay (when they are on holiday for instance). Airbnb asks a commission of the rent as intermediary between the supply and demand, but leaves enough profitability for suppliers and is thereby very successful. This completely differs from the traditional hotel market and gives individuals (who can be both tenant and landlord) a sense of control and a fair competitive price. Other examples are Kickstarter (for crowd sourced investments), Peerby (to borrow items from neighbours), Craigslist (for buyers and sellers of new and used goods), SnappCar (to rent a private car from an individual), 3D Hubs (to find and use private 3D printers), Magnet.me (a recruitment marketplace for job seekers and employers) and eBay (an auction platform for consumers). These initiatives are amongst the fastest-growing businesses of the past years and are part of the so-called 'sharing economy' or 'collaborative economy' (Owyang, Tran, & Silva, 2013; Tanz, 2014).

The markets in which these examples are active are "characterized by the presence of two distinct sides whose ultimate benefit stems from interacting through a common platform" (J.-C. Rochet & Tirole, 2003, p. 990). These common platforms are mostly called 'two-sided' or 'multi-sided' platforms, which are defined as "technologies, products or services that create value primarily by enabling direct interactions between two or more customer or participant groups" (Hagiu, 2014, p. 71). The sides are the sellers and the buyers, in other words the supply-side users and the demand-side users. The main value of these platforms for society is that the search and/or transaction costs for participants are considerably reduced (Hagiu, 2014; Seamans & Zhu, 2014). A fair portion of multi-sided platforms relies on online content provided by the supply side of their users. This content is generally called user-generated content (UGC) (Albuquerque et al., 2012). The platforms serve only as a means for communication where both sides of the market (e.g. buyers and sellers) are brought together. The basis for this communication is the UGC. Revenues for these platforms are mostly obtained from transaction commissions or through online advertisement (Albuquerque et al., 2012; Bakos & Katsamakas, 2008; Eisenmann et al., 2006). The examples mentioned above fall into this category, but other multi-sided platforms – such as Google's Android operating system, Sony's PlayStation, American Express and PayPal – do not. They do not only consist of an online marketplace which



relies on online UGC, but concern more tangible items such as game consoles, smartphones and credit cards. Albuquerque et al. call this type of multi-sided platforms ‘online two-sided market of user-generated content’, which they describe as an “[online] intermediary that maximizes its own objectives by bringing together content creators, consumers, and in some cases advertisers” (2012, p. 407). Many other comparable names and definitions can be found, but they roughly describe the same examples, phenomena and markets: two-sided platform, two-sided market, two-sided network, multi-sided network, utilitarian peer-to-peer trading service, electronic marketplace, intermediation service providers (Armstrong, 2006; Eisenmann et al., 2006; Goldsmith et al., 2013; Hagiu, 2014; Hamari, 2013; Peña-López, 2010; J. C. Rochet & Tirole, 2006; J.-C. Rochet & Tirole, 2003; Seamans & Zhu, 2014). In this research, ‘online two-sided market of user-generated content’ will be referred to as online marketplace, which indicates that it concerns an online platform that purely facilitates communication of two user sides through online content created by these users. There could be more sides to a marketplace, but mostly it has a supply and a demand side.

### 1.1.2 Challenges for Online Marketplaces

Characteristic for two-sided platforms and more specifically online marketplaces is the ‘chicken-egg’ problem that occurs, often in an early stage. Caillaud and Jullien describe the chicken-egg problem as follows: “to attract buyers, an intermediary should have a large base of registered sellers, but these will be willing to register only if they expect many buyers to show up” (2003, p. 310). The value of the platform for one side of its users depends on the number of users from the other side that are active and vice versa. This can hinder platform growth, while the platform might have enough potential users to be successful. A practical example of this chicken-egg problem can be given with Uber, a two-sided platform which where individuals can find and offer taxi rides. If there are very few users on Uber that offer taxi services, a consumer will not use Uber to get from A to B, because the possibilities are limited, waiting times are long and prices are high. On the other hand, if there are no users requesting rides, there will be fewer people offering their driving services, because they will not be able to make money. Even now, when it is a large and very successful platform, there is still an imbalance: using Uber will never result in an instant ride (not enough supply to meet demand) nor result in an instant customer (not enough demand to meet supply). On the long term, after overcoming the initial growth phase and chicken-egg problem, there is a continuous imbalance between the two sides, which needs to be resolved in order to grow. These challenges for online marketplaces are commonly acknowledged within literature (Armstrong, 2006; Bakos & Katsamakas, 2008; Cambini et al., 2011; Eisenmann et al., 2006; Goldsmith et al., 2013; J.-C. Rochet & Tirole, 2003). Hagiu even states: “[o]vercoming the chicken-and-egg problem is one of the most difficult challenges for many MSP’s [Multi-Sided Platforms]” (2014, p. 72).

The underlying phenomena for the chicken-egg problem are network effects, meaning that the supply and demand side are always interdependent and can both limit (negative network effects) and stimulate (positive network effects) each other to grow. Network effects were first described by Katz and Shapiro: “The utility that a given user derives from the good depends on the number of other users who are in the same ‘network’ as he or she” (1985, p. 424). Positive network effects grow as the network size grows, because of positive feedback loops. This is also referred to as the ‘bandwagon effect’ (Rohlfs, 2003). The effects of buyers on buyers or sellers on sellers are called ‘same-side network effects’. The effects sellers on buyers and vice versa are called ‘cross-side network effects’. For most online marketplaces, the cross-side network effects are positive. After all: more active sellers implies more product diversity, thus choice for consumers and also implies more competition for sellers thus lower prices for buyers. The other way around: more buyers decreases selling time and increases demand thus market price, which is a positive development for sellers (Bakos & Katsamakas, 2008; Eisenmann et al., 2006; Peña-López, 2010; Seamans & Zhu, 2014). The same-side network effects on online marketplaces are predominantly negative, but positive same-side network effects can also be discerned (Eisenmann et al., 2006). Airbnb is used to illustrate this. First,

if user X wants to rent a home for a holiday weekend in Amsterdam, the number of other demand side users on Airbnb can negatively affect the value of Airbnb for user X. Namely, more demand for Amsterdam Airbnb homes drives up the prices and decreases the availability of the homes. However, the fact that many people want to rent and especially have already rented a home through Airbnb in Amsterdam can result in more reviews and ratings of sellers and their homes, increasing transparency and quality for buyers. Also, for Amsterdam home owners, it can be useful when there is a large Airbnb offer in Amsterdam, because it makes it easier to determine a good price homes. On the other hand, this also means the price needs to be competitive and the quality needs to be good, in order to get buyers to contact you. As can be discerned from these examples, the negative same-side network effects usually outweigh the positive same-side network effects.

Online marketplaces heavily rely on (positive) cross-side network effects for growth, where the supply-side user usually provides the bottleneck. Sellers have to supply user-generated content, for which they have to be able and willing to do so. Generally this is a larger effort than browsing the existing content, which is the activity that demand-side users perform on an online marketplace. For online marketplaces, the main challenge is to stimulate users to create as much content as possible, because this is the motor of the entire platform. It is vital for marketplaces to target the user side that is lacking in numbers, which mostly is the supply-side (the sellers) (Armstrong, 2006; Caillaud & Jullien, 2003; Eisenmann et al., 2006; Hagiu, 2014; Jordan & Hariharan, 2015; J. C. Rochet & Tirole, 2006). The network interaction between the creators and final users of UGC indicates the general performance of online marketplaces (Albuquerque et al., 2012).

Adding to the challenges posed above, online marketplace users are increasingly becoming aware of the data they do or not wish to share online (as they become increasingly dependent on internet services) and have different attitudes towards different types of data. The way companies and online platforms collect their data also influences their decisions about these platforms and their reliability (Khatibloo, 2011). Because of numerous privacy scandals in the past couple of years, the attitude of online users towards supplying their data to online platforms has become more and more conservative. Also, because of the complexity in this field (endless combinations of users, platforms, data and contexts), there is no single best practice to engage users, but rather a vast paradigm (Roe, 2012).

### 1.1.3 Gamification and its Potential

Over the last couple of years, 'gamification' has gained attention. Gartner, a large information technology research company, predicted that by this year, seventy percent of Global organisations will have at least one gamified application (Petty & van der Meulen, 2014). Both in management and scientific literature, more and more is published, with the discussion and progress mainly focused within game studies, human-computer interaction and social sciences (Huotari & Hamari, 2012). Gamification does not have to concern a full technical or computer-based game, but rather elements of it, which are used in another context. A simple and illustrative definition of gamification is: "*the use of game design elements in non-game contexts*" (Deterding et al., 2011, p. 7). Where serious games combine all types of gaming elements into a whole, in order to provide an immersive and 'other world' experience, gamification uses the strategies of persuasive technology to affect behaviour during a task, thus serving a facilitating role in a service a product, rather than the game elements being the core of it (Deterding et al., 2011). The main potential and added value of gamification is the increased user engagement because of the elements of play, fun and competition, which can be added by gaming, creating a 'gameful experience'. Increased engagement can result in an increased or transformed motivation to, depending on the system where it is implemented, participate, use, learn, have social interaction or perform tasks (Deterding et al., 2011; Groh, 2012; Hamari et al., 2014; Hense et al., 2014; Huotari & Hamari, 2012; Nicholson, 2012; Thiebes et al., 2014; Werbach & Hunter, 2012).

Within scientific literature, gamification has mainly been applied in online services (Hamari, 2013), workplaces (Oprescu, Jones, & Katsikitis, 2014), commercial standardised work (for instance in line assembly work) (Hense et al., 2014) and education (Borys & Laskowski, 2013). Gamification does not necessarily have to be implemented in an online system, but this is true in most of the cases and especially in non-scientific literature, where the most gamification cases are to be found. These commercial gamification applications, described in blogs, websites and the like, have online marketing and employee motivation as main application fields. Gamification is used to increase revenue by stimulating the activity of users on online services or employee productivity. Successful examples are profile completeness progress bars (for instance on LinkedIn and Facebook) and services as Foursquare, Codecademy, Waze and Nike+ Running. Usually, a software layer is implemented which incorporates game elements as points, achievements and rewards into a service or product (Deterding, 2014a; Groh, 2012).

When looking back at the main challenge for online marketplaces, the potential of gamification to aid in this challenge can easily be discerned. UGC is driven by user activity, where the decision to place content is not only dependent on the demand for content (the network effect) but also on the level of engagement of a user with the online marketplace. The users themselves must supply UGC and be intrinsically willing to do this. Users should be engaged with the platform, which often relates to the fact that there is something for them to receive or gain, in return of their data supply (Forrester Consulting, 2014). This can be translated as the experience that someone has when visiting the online marketplace, which is a vital component when making a decision to place content (Albuquerque et al., 2012). If gamification can add elements of fun, play and competition to this experience, it will presumably have a positive influence on the amount of UGC that is placed on an online marketplace. This statement is supported by the only previous study on the effect of gamification on an online marketplace that has been executed. In this study, Juho Hamari (2013) examined if adding badges, as a form of achievements and social comparison, could increase the user activity on Sharetribe. Sharetribe is an online marketplace that enables individuals that live close to each other to have non-monetary transactions, such as carpooling and borrowing each other's goods. Hamari concluded: "This study was able to confirm that users who had actively exposed themselves to badges in Sharetribe were also significantly more likely to actively use the service, list their goods for trade, comment on listings and complete transactions" (2013, p. 243).

#### **1.1.4 Lack of Valid Scientific Gamification Studies**

The main potential and added value of gamification fit well into the main challenge of online marketplaces: engaging users with a gameful experience, in order to stimulate their activity and post their content on the platform. However, one empirical gamification effect study on an online marketplace is not enough. This is broadly acknowledged by Hamari (2013) in his conclusions, where he emphasises the need for more comparable effect studies. There are many industry examples of successful gamification applications on online marketplaces and online platforms in general, where the goal usually is to stimulate users to post more content. However, proper documentation of these studies is rarely available, making the true value of these applications hard to acknowledge from a scientific point of view.

Moreover, the gamification field contains strongly divided opinions, definitions and movements (Deterding, 2014a; Groh, 2012). Some see it as a re-invented marketing tool to trick and exploit customers in order to increase profits, while others speak of a true enhancement of user engagement by creating valuable gameful experiences (Farrington, 2011; Hamari, 2013). Currently, gamification is in a commercial niche, preventing its scientific evaluation. The gamification field is "littered with shallow interpretations and implementations" (Deterding, 2014a, p. 306). Management literature mentions both extreme positive and negative effects of gamification, but these have not been appropriately tested and proven and are "largely based on anecdotal and intuitive presumptions" (Hamari et al., 2014, p. 3025). To illustrate: in October 2014 Google search results for 'gamification'

counted over 1.8 million, with trends continuously rising since the first searches in 2010 (Google, 2014). Gamification Corp. (2014) states there are 116 gamification vendors, offering commercial consulting and software implementation services (for example: PunchTab, Badgeville and BunchBall). Looking at scientific literature, a recent literature review found only 24 scientifically relevant papers on empirical gamification studies. Within these studies, many methodological shortcomings could still be identified. However, the majority of the 24 studies did report partially positive effects of gamification in terms of motivation, enjoyment or other behaviour (Hamari et al., 2014), which means that the effects are worth looking into. Scientific gamification studies not only lack in empirical evaluation, but also in design of gamification: “there are currently no established, let alone empirically tested methods for the design of gameful systems” (Deterding, Björk, Nacke, Dixon, & Lawley, 2013, p. 2). Rather, there is a plethora of design principles, foundations, frameworks and perspectives derived from game design theories. More fundamentally, the definition of gamification presented above (Deterding et al., 2011) has been challenged and amended by a number of other researchers, resulting in the fact that multiple definitions are used (Burke, 2014a; Huotari & Hamari, 2012; Zichermann & Cunningham, 2011).

The lack of scientific literature, the lack of research results and its young age indicate that gamification is currently still in its infancy. The design principles are scattered and diverse, there are multiple definitions, the number of valid evaluation studies is scarce, there is a rampant growth of self-proclaimed gamification experts and some have deemed gamification as a marketing buzzword. This is also commonly recognised within the main gamification literature, where the need for more empirical research on gamification design and evaluation is expressed (Deterding et al., 2011; Deterding, 2014c; Groh, 2012; Koivisto & Hamari, 2014; Thiebes et al., 2014). Exemplary, Sebastian Deterding – who is considered one of the most influential gamification researchers – states it as follows: “Today, the main challenge has become to work against the grain of existing preconceptions of gamification (be they apocalyptic or utopian), established by evangelists, critics, industry practices, and mass media reporting. Many have rightfully questioned whether gamification is anything more than a marketing ruse to sell the next digital snake oil” (2014a, p. 306).

### 1.1.5 Problem Statement

The problem which has been identified above and which will be the subject of this research can be summarised as follows.

*Online marketplaces need user-generated content in order to grow and ultimately survive, which implicates that the engagement of their users with the marketplace is vital. Gamification might be able to engage online marketplace users and stimulate the creation of user-generated content, but this effect has not been scientifically proven. Also, there is no generally accepted design approach to gamify online marketplaces or comparable online platforms in general.*

Concluding: the potential combination of gamification and online marketplaces, in the context of the network effect and forthcoming chicken-egg problem, has not been explicitly identified before. It is worth investigating, as it is an interesting combination of a relevant societal subject and a promising but disputed method.



## 1.2 Research Foundations

### 1.2.1 Goal

The goal is to find the applicability of gamification to increase the amount of UGC on online marketplaces, in order to aid online marketplaces with the important challenges they face. This will be done by designing a gamification treatment for a case online marketplace, according to a pre-defined gamification design method and evaluating the effect of this treatment. The evaluation should be methodologically valid and a structured game design method should be followed, following the future research recommendations from existing gamification literature (Deterding, 2014a; Hamari et al., 2014). This will allow reproduction and aid generalisation of the research results. Therefore, the goal of this research is:

*To structurally design and evaluate the gamification of an online marketplace, aimed at the increase of user-generated content, while avoiding the known pitfalls of existing gamification studies.*

### 1.2.2 Research Questions

The main research question that follows the research problem and corresponding research goal is:

*What is the suitability and capability of gamification to increase the amount of user-generated content on online marketplaces?*

The following sub questions need to be answered in order to answer the main research question:

1. What is the origin and definition of gamification and how can it be delineated from similar fields?
2. What lessons can be learned from gamification criticism and previous gamification studies?
3. Which design method can be used to structurally gamify online marketplaces?
4. Through application of the chosen design method: what is the most promising treatment to gamify an online marketplace in order to increase the amount of user-generated content?
5. What is the effect of the gamification treatment on the amount of content generated by users?
6. What is the suitability of the design method for the way of working of online marketplaces?

The sub questions form the prerequisites that are needed to answer the main research question. The first is an insight into the definition and origin of gamification, so that one perspective can be adopted in the rest of the research. Second, the amount of criticism that gamification has received might be worth looking into, in order to extract some good practices for this research. As discussed in the problem outline, there is a vast number of design principles and methods for game design, but there is no established, generally practiced and reviewed gamification design method. A structured design approach is needed in order to answer the research question, so the third prerequisite is a design method that needs to be found in and/or constructed from literature. The fourth prerequisite is a gamification treatment, constructed with the chosen design method, which has the most potential to increase the amount of UGC. The fifth prerequisite is to determine the quantitative effect of the developed design, for which an evaluation method needs to be used. The last prerequisite is a more qualitative assessment of the design method performance, to assess the suitability of application to online marketplaces general.

### 1.2.3 Research Approach & Methods

#### General Approach

The use of gamification to increase the amount of user-generated content on online marketplaces will be evaluated in two different ways, based on its suitability and its capability:

1. Suitability. A qualitative evaluation, focusing on the suitability of gamification to online marketplaces. A design method is evaluated by application to a case, which also results in recommendations to improve the design method. The suitability considers the way of working of online marketplaces and the known criticism on gamification design that has been mentioned in literature.
2. Capability. A quantitative evaluation, focusing on the capability of gamification to increase user-generated content on online marketplaces. The actual effect of a gamification treatment (which was developed with the design method) on the behaviour of online marketplace users is evaluated. This is done on one specific online marketplace that serves as a case in the research.

The second evaluation is the most important in this study, because it is more objective and can be thoroughly grounded in existing literature, while the first evaluation is mainly based on qualitative experiences of participants in this research.

#### Methods to be Used

The first step is to conduct a literature review and desk research to provide a theoretical foundation, in which both management and scientific literature are addressed, given the current state of art of gamification. With this, sub question 1 and 2 can be answered.

Next, a gamification design method is needed. Very recently, Deterding (2014c) wrote an article in which he reviews relevant game and gamification design methods from scientific and management literature, lists gamification critique and lessons to be learned and synthesises all into a new and comprehensive 'gameful design' method. This design method will be used as reference point for gamification design in this research for three reasons. First of all, it is the only gamification design method available. Secondly, it incorporates most of the theory and gamification criticism that (would) been used here in order to develop a new gamification design method, if Deterding's method would not have been available. In fact, a gamification design method was partially developed for this research, until the method by Deterding was found in literature. The resemblance between these methods is significant, indicating that the method by Deterding has been grounded on the same ideas. Lastly, the perspective of Deterding (2014a) on a definition of gamification is adapted in chapter 2.1, so the choice of his corresponding design method seems logical. The method will be explained in more detail in chapter 2.4. Because of this delineation, the research is limited to a certain perspective on gamification. However, the gamification field as a whole is evaluated by summarising relevant literature in chapter 2 and by generalisation of the design method and case results to gamification theory where possible, at end of chapter 5 and in chapter 6.

Rather than examining the gamification design method and the problem set theoretically, adding an empirical study seems much more valuable. Especially with regard to sub research question 5, where an effect of the gamification design method needs to be examined, and the practical problem of online marketplaces, an empirical research is suitable. Given this, an online marketplace case is needed for which a gamification treatment can be developed and in which the design can be empirically evaluated. The case should resemble the general online marketplace and UGC problem as much as possible and allow for a general gamification design and evaluation method to be applied. The research approach is general to the online marketplace and UGC problem described and not specific to this research. So the design and evaluation method can be used in other cases also, based on the reflections that are given at the end of this report. Moreover, it allows for a possible theoretical generalisation – to a certain extent – of the empirical results in the case.

When the case and design method are clear, the design method can actually be executed, which will generate the most promising treatment and thus the answer to sub question 4. Methods to be used here are semi-structured interviews, desk research, analysis of historical data and workshops. These are needed to dive into the case.

To evaluate the design an online, double blind, randomised controlled experiment (or 'A/B test', in online jargon) (Ron Kohavi et al., 2008) will be used, because is used in most scientific gamification evaluation studies. In these studies, the authors reflect on this method as successful, scientifically valid (if preconditions such as sample size are met) and relative simple to apply (Denny, 2013; Farzan et al., 2008a, 2008b; Hamari, 2013). Webb even states: "Ideally, gamification would be applied and measured with A/B testing, [...] in which some users work with the gamified system while others use the non-gamified system" (2013, p. 611). Appendix A contains a review of five methodologically valid empirical gamification studies according to Hamari et al. (2014), where this is concluded. The exact interpretation of the randomised controlled experiment for the research and the setup that was used are described in chapter 4.2. However, important to note here is that evaluation of gamification design is an iterative process and not a single event. A first evaluation does not only generate data on the effect of the design, but also serves as input for small design changes and further evaluation of these small changes, etc. (Deterding, 2014c; Ferrara, 2012; Hamari et al., 2014). It might provide very valuable insights to look into various user segments for which the design was or was not effective (Ron Kohavi et al., 2008). In this research it is only possibly to perform a first evaluation step, due to time constraints. The data as result of the experiment will be analysed, after which the effects of the implemented design can be presented, answering sub question 5. Given the available theory, set out in the beginning of this chapter, a positive effect of gamification on the amount of user-generated content on an online marketplace is expected. The experiment is used to assess the magnitude and significance of this effect.

## 1.2.4 Relevance

### Scientific Relevance

The scientific relevance of this study can easily be distinguished. For online marketplaces it is interesting to see the main results of this research, namely the effects of gamification on the amount of UGC that is produced. If so, gamification can be a new solution to their main challenge (Hagiu, 2014). Also, thorough documentation of the evaluation and design method used in this research will also allow easy reproduction within another case. Gamification literature demands for more empirical research and case studies; especially research that is executed with a valid evaluation methodology (Deterding, 2014a; Farzan & Brusilovsky, 2011; Groh, 2012; Hamari et al., 2014; Nicholson, 2012; Thiebes et al., 2014). "Further studies should especially try to avoid these [methodological] pitfalls in order to refine the research on gamification" (Hamari et al., 2014, p. 3029). This need is also confirmed by the novelty of the gamification field in general and the rampant growth of anecdotal experiences in common web sources. Also, using and describing a specific method to design gamification will be of added value, since it has not yet been done within scientific literature (Deterding et al., 2013; Deterding, 2014c).

### Societal Relevance

In general, for society, it is vital for new online marketplaces in promising markets to overcome the chicken-egg problem. Namely, this problem does not necessarily indicate that there is no user base interested in the marketplace, but rather an interdependent stall of both user sides in adapting the online marketplace (J. C. Rochet & Tirole, 2006). While online marketplaces can have a very positive impact on society and the possibilities of individuals, as a part of the sharing economy and illustrated with successful examples as Airbnb, Uber and Craigslist (Seamans & Zhu, 2014). Many of the fastest growing business of the past decade have been online marketplaces and their popularity can be

translated to the fact that their users can lower their transaction and search costs and save time (Hagiu, 2014). Craigslist, the largest classifieds website in the U.S.A., was estimated to have saved classifieds advertisement consumers around five billion dollars between the years 2000 and 2007, compared to the traditional situation with only paid newspaper advertisements (Seamans & Zhu, 2014). Gamification could potentially aid online marketplaces in using the network effect to their benefit, overcoming initial start-up barriers and settling imbalance issues between the supply and demand sides.

### 1.3 Report Structure

The research starts with a literature review and desk research. This is described in the theoretical founding in chapter 2, containing the origin, definition, perspectives and criticism of gamification. Also, chapter 2 contains underlying theory to the adopted perspective and a description of the gamification design method by Deterding (2014c). Chapter 3 concerns the case, the gamification treatment for the case and a description of the implementation of this treatment. Chapter 4 describes the experiment setup for the quantitative evaluation of the design and the results of the experiment. Chapter 5 contains the discussion of the experiment results, limitations of this study and recommendations for future research, both for both for online marketplace managers and gamification researchers. Lastly, chapter 6 contains the conclusions and reflections. Where possible, the sub research questions are answered provisionally throughout the chapters, but will be definitively answered in chapter 6, along with the main research question. The reflections encompass the more practical issues regarding the design, evaluation and research project in general. See Figure 1 below for an overview.

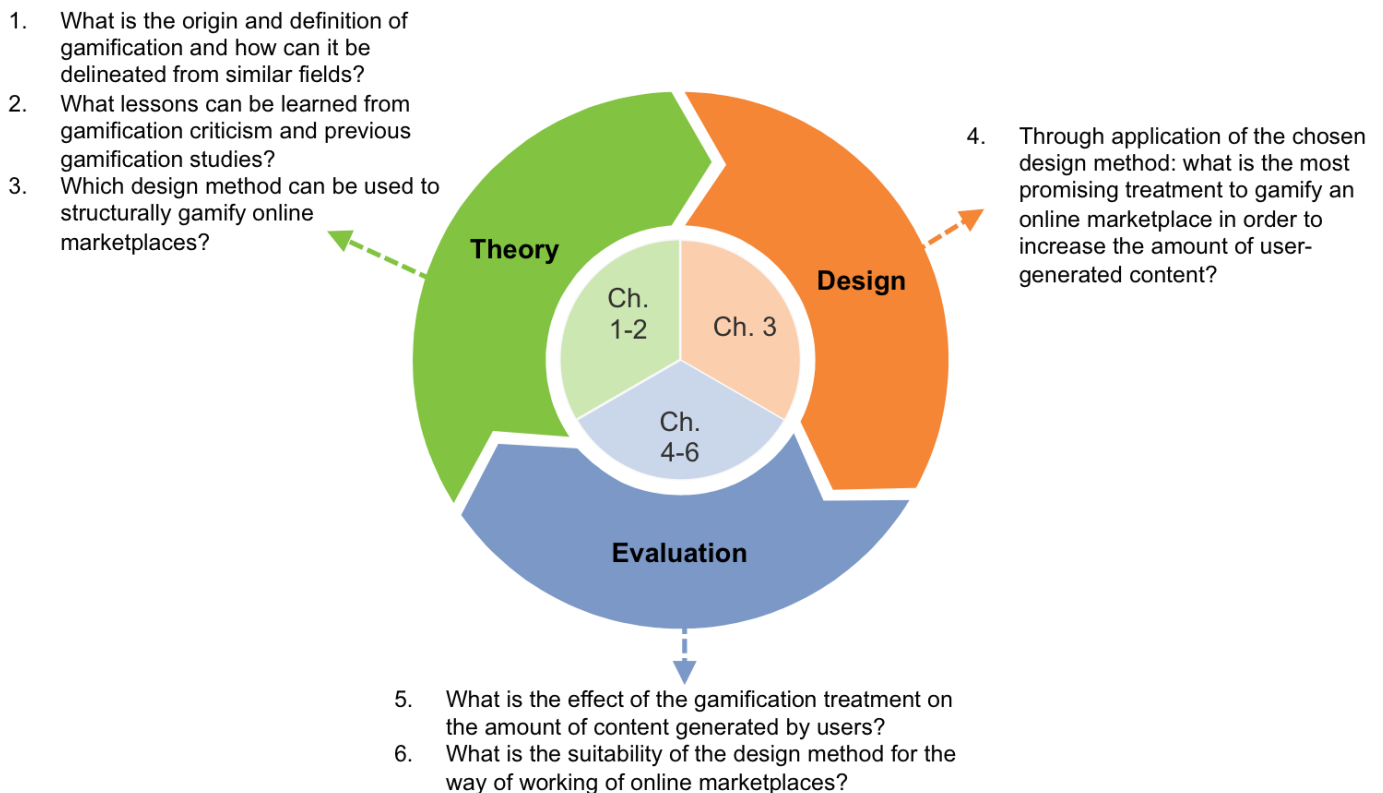


Figure 1: Report Structure

## 2 THEORETICAL PERSPECTIVE ON GAMIFICATION

---

*In this chapter a theoretic founding for the research is provided, from management and scientific literature, answering first sub question of this research. Only literature that actually mentions 'gamification' was explored. There are many related fields that also provide relevant content (Hamari et al., 2014), but for the sake of time and simplicity the literature was delineated this way. Based on the research goal, the research questions and the subject matter in the case, the theoretical fields that need to be examined can be distinguished. Since gamification has multiple definitions and not one has yet been commonly adopted, it is necessary to define gamification for this research. Also, as indicated in the first chapter, the gamification criticism cannot be ignored and might provide interesting lessons for this research. Other lessons can be learned from user engagement and motivation theory, on which gamification is partly founded. Once this has been done, the chosen gamification design method is explored and adapted for this research. The following sub research questions are provisionally answered in this chapter:*

- 1. What is the origin and definition of gamification and how can it be delineated from similar fields?*
  - 2. What lessons can be learned from gamification criticism and previous gamification studies?*
  - 3. Which method can be used to structurally design gamification for online marketplaces?*
- 

### 2.1 Gamification Defined

First of all, a look into the origin, definitions and delineation of gamification is needed, in order to provide background and pave the road for the perspective adopted on this research.

#### 2.1.1 Origin

Research on gaming with a serious purpose started in the 1980s, when Malone and Bowman pinpointed the attractive elements of computer games, in order to use these elements in an educational setting. This way, the engagement and motivation of students could be improved (Borys & Laskowski, 2013). 'Gamification' as a term originated in the digital media industry and was proposed for the first time around a decade ago. It received public attention after 2010, when several conferences covered the subject and some large corporations had implemented it in their business strategy. A very influential and often viewed presentation by Jesse Schell (2010) is often recognised as one of the kick-starters. From there, gamification as a term was commonly adopted. It has been implemented in news and entertainment media, interaction design, health, productivity and sustainability and has been recognised as a new branch in the gaming field (Deterding et al., 2011) (Groh, 2012)

#### 2.1.2 Definitions

Because of the novelty and diversity of the gamification field, it is very relevant and even necessary to define gamification. This way the clarity, validity and reproducibility of this research can be better assured.

## Management Literature

Deterding et al. listed some definitions by managers, consultants and vendors on the web, which are mostly very practical and focused on the benefits for their potential clients (2011, p. 10):

- 'the adoption of game technology and game design methods outside of the games industry'
- 'the process of using game thinking and game mechanics to solve problems and engage users'
- 'integrating game dynamics into your site, service, community, content or campaign, in order to drive participation'

Other industry definitions are to be found in several management books that have been published on gamification (Burke, 2014a; Paharia, 2013; Werbach, 2014b; Zichermann & Linder, 2013). A leading book in this field, in terms of scientific citations, was written by Gabe Zichermann and Christopher Cunningham, who gave gamification the following definition: "The process of game-thinking and game mechanics to engage users and solve problems" (2011, p. xiv). The discussion is still on going, with for instance a recent redefinition of gamification by Brian Burke in a Gartner blog post, where he claimed that gamification is often defined too loosely, causing confusion, implementation failures and too high expectations. He redefined gamification as: "the use of game mechanics and experience design to digitally engage and motivate people to achieve their goals" (Burke, 2014b, p. 1), accompanied by his recent book on gamification (Burke, 2014a). This in turn resulted in a new wave of blog posts and reactions by other 'industry experts'.

## Two Perspectives within Scientific Literature

Within scientific literature, Huotari and Hamari defined gamification based on service marketing, halfway 2011. They compared games with core services and game elements with enhancing services, resulting in the following definition: "Gamification is a form of service packaging where a core service is enhanced by a rules-based service system that provides feedback and interaction mechanisms to the user with an aim to facilitate and support the users' overall value creation" (Huotari & Hamari, 2011, p. 13). For instance: where a LinkedIn profile is the core service, the progress bar for the number of personal details filled is the enhancing service. Or: a café is the core service and its mayorship competition is the enhancing service.

At the end of that year, Deterding, Dixon, Khaled and Nacke (2011) provided a new definition, while recognising that until that moment, the previously mentioned paper by Huotari and Hamari was the only scientific attention paid to gamification yet. It was however too broad (a vending machine with a display is also a 'rules-based system that provides feedback and interaction mechanisms to the user with an aim to facilitate and support the users') and not within the right context (gamification in the sense of game elements can also be the core service, not necessarily a service packaging), they argued. Their new definition of gamification is: "the use of game design elements in non-game contexts" (Deterding et al., 2011, p. 10).

With this, they defined game elements as "elements that are characteristic to games" (Deterding et al., 2011, p. 12). These elements are found in most of the games (digital and physical) but not automatically in all, they are related to games and they are determinant for gameplay. The game element list was quite broad, as can be seen in Figure 2. These elements are usually linked to the MDA model (Hunicke, Leblanc, & Zubek, 2004), which states that game elements can be divided over three different categories, in order of hierarchy: aesthetics (fun experiences), which are caused by dynamics (strategies, behaviours and interactions of players/users), which are made possible by mechanics (the rules and most basic game components). This model has been used to pinpoint different elements in serious and entertainment games and is quite useful. But, when it comes to gamification, there is much confusion and diversity in terminology concerning these different game elements. The variety and quantity of lists of game elements in literature are enormous. For instance, Thiebes (2014) constructed a comprehensive list of game elements which is compromised of many



other lists of game elements and rated by the effect that they have shown in different studies. Only the game elements that have actually proven to be successful in evaluation studies have been included. Still, the list theoretically not finite and the elements are not mutually exclusive.

The non-gaming context is not to be bound, because there is no need to. For the contexts in which gamification can be of added value have not all been explored or determined, but the potential is very broad (Deterding et al., 2011). Deterding et al. (2011) recognise that their definition is still quite broad and debatable. Kevin Werbach (2012) adapts the definition and states that non-game contexts are defined by the fact that the objective for their user or actor lies outside of a game. So gamification involves game elements and game design, for a purpose other than a game (which is mostly 'to have fun').

Level	Description	Example
<i>Game interface design patterns</i>	Common, successful interaction design components and design solutions for a known problem in a context, including prototypical implementations	Badge, leaderboard, level
<i>Game design patterns and mechanics</i>	Commonly reoccurring parts of the design of a game that concern gameplay	Time constraint, limited resources, turns
<i>Game design principles and heuristics</i>	Evaluative guidelines to approach a design problem or analyze a given design solution	Enduring play, clear goals, variety of game styles
<i>Game models</i>	Conceptual models of the components of games or game experience	MDA; challenge, fantasy, curiosity; game design atoms; CECE
<i>Game design methods</i>	Game design-specific practices and processes	Playtesting, playcentric design, value conscious game design

Figure 2: Game elements as defined by Deterding et al. (2011, p.12)

Late 2012, Huotari and Hamari responded to the definition of Deterding et al. (2011) and amended their first definition from 2011. They once again compare gaming with service marketing and state that a game is defined by both systemic and experiential conditions, not just systemic conditions. The experiential condition is defined as having a 'gameful experience', or a "hedonic, challenging and suspenseful experience[s] for the player(s)" (Huotari & Hamari, 2012, p. 19). Therefore, the definition of a game depends on each individual, because every person can have a different experience with every game, and the customer can only determine the added value of gamification. Thus, rather than looking at the content, methods and system, gamification should be defined by the goal it tries to achieve: giving customers a gameful experience and thereby potentially raise their engagement level. Furthermore, Huotari and Hamari (2012) argue that the set of game elements has not clearly been defined and can also be found in non-game contexts (which was also recognised by Deterding et al. (2011)). Moreover, the elements do not automatically create gameful experiences. They refer to game elements as 'affordances', where an affordance is a quality of the service that contributes to the development of gameful experiences. These affordances, i.e. the gamification, can be provided by either (1) the core service provider, (2) a third party service provider, (3) the user/customer of the core service him-/herself or (4) another user/customer of the core service provider (Huotari & Hamari, 2012). See the figure below for illustrative examples.

Core service	Enhancing service	Gamified service	Gamification provider
Clothing store	Loyalty program offered through Facebook deals [8]	Customers who check in regularly using Facebook Places are offered reductions.	Clothing store (core service provider) and Facebook
Restaurant (e.g. Starbucks)	Local Badges in Foursquare	Customers who check in at least three times a week to a same location using Foursquare get a badge.	Foursquare (a third party)
Sports bar	Drinking game [34]	Deciding to incorporate a drinking game to watching hockey, for example.	Customer himself/herself
Coffee house	Tip offered through Foursquare [9]	Adding a quest-like tip to other customers while they are waiting coffee.	Another customer and Foursquare

**Figure 3: Gamification examples from a service marketing perspective (Huotari & Hamari, 2012, p. 20)**

Huotari and Hamari proposed the following new definition of gamification: “a process of enhancing a service with affordances for gameful experiences in order to support user's overall value creation” (2012, p. 19). Note that this definition does not specify the nature of the service being enhanced, thereby acknowledging that not only non-game, but also game contexts can be gamified. Value creation refers not to direct economic benefits but to the fact that the added affordances (or game elements) transform usage motivations and intentions from extrinsic to intrinsic and from more utilitarian to hedonic, by providing a gameful experience (Huotari & Hamari, 2012). The economic benefit is then more for the core service provider, if the gameful experience results in increased user engagement (implying more conversions, page views, user-generated content, etc.).

As one can see, there are two perspectives on gamification: an experiential perspective led by Hamari and a systemic perspective led by Deterding. On October 8<sup>th</sup> 2014, searches on Scopus and Google Scholar were done for papers with both ‘gamification’ and ‘defin\*’ in the title, the abstract or the keywords. In both searches, Deterding et al. (2011) came back as most cited result and Huotari & Hamari (2011) as the second most cited result, respectively with 114 and 23 citations on Scopus and with 494 and 96 citations on Google Scholar. The definition by Deterding et al. (2011) is most widely adopted in the literature that was examined for this research and its relatively high adoption level is also confirmed within the individual papers (Borys & Laskowski, 2013; Groh, 2012; Hamari et al., 2014; Hense et al., 2014; Nicholson, 2012; Oprescu et al., 2014; Thiebes et al., 2014; Volkova, 2013).

Kevin Werbach recently published an article on a new definition of gamification from a process point of view, by definition of “the process of making activities more game-like” (Werbach, 2014a, p. 271). He depicts the definition by Deterding et al. as ‘the elemental definition’ and the definition by Huotari & Hamari as ‘the service marketing definition’ and uses a progress bar example to illustrate that both of these definitions do not fully meet their purpose. LinkedIn uses a progress bar to give feedback to users as to how far they are in completing their online profile on the platform. Apparently, it engages users and stimulates them to fill in more of their details and is therefore often recalled as a successful example of gamification. Werbach (2014a) argues that Microsoft (amongst many others) also uses progress bars, but then to indicate the progress of software installation. According to the elemental definition, both of the examples would be called gamification, since a progress bar is a game element and both contexts are non-game contexts. But from logical reasoning, the Microsoft example is not

actually gamification, so the elemental definition is too broad. The service marketing definition would apply only if the progress bar enhances the core service that a user is entailed in. So is installing a piece of software the core service of Windows? And is being able to fill in your profile details the core service of LinkedIn? This might be the starting point for an interesting discussion, but is not the main question as to whether the progress bar can be defined as gamification or not. Another prerequisite for gamification from the service marketing definition is if the user's overall value creation is supported, with a side note that this user value is inherently subjective and different for each individual. So a user has to feel something like 'pleasure', 'suspense', 'mastery' or 'gamefulness', in order to call the service he/interacts with gamified (Huotari & Hamari, 2012). This is different for each individual. The point Werbach tries to make is that the elemental definition depends on the definition of game elements. He states: "there is no universal list of game elements. This inherent uncertainty is problematic" (2014a, p. 267). The service marketing definition depends on what a core service is and the value individual users allocate to a service enhancement. Both are imperfect and most furthermore they do not take into account the design process and the intentions of the designer, which are at least equally as important. What matters is not which game elements, concepts, patterns or principles are selected, but how they are selected, implemented and integrated (Thiebes et al., 2014; Werbach, 2014a).

### Concluding on a Gamification Definition

Sebastian Deterding published work in which he recognises the arguments of Huotari and Hamari. He states: "I suggest expanding the remit of gamification from the structuring of objects to the framing of contexts, and from game design elements to motivational affordances" (2014a, p. 307). He proposes to combine both perspectives into a more socio-technical view, which recognises motivational affordances and also the context, the design process and motivational factors: "Understood as such – a unified whole of restructuration and reframing – gamification is a holistic socio-technical systems design practice [...] one that understands humans interacting with technology as assemblages, activity systems, or ecologies of heterogeneous and intertwined actors" (Deterding, 2014a, pp. 312–313). This view is very closely aligned with Werbach's process definition and is adopted throughout the rest of this research, both in the gamification design method that was chosen (Deterding, 2014c) and the theory that was reviewed in the rest of this chapter.

### 2.1.3 Context & Delineation

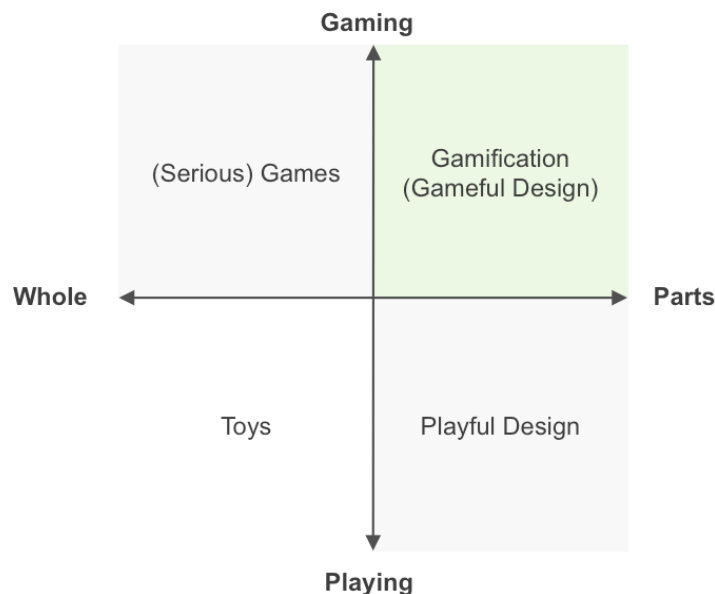
Now that the different definitions and perspectives within gamification have been identified, it is useful to demarcate it from comparable gaming fields. This was done using the primary gamification literature and the literature connected to the perspective as adopted in the previous paragraph (Deterding, 2014a, 2014c; Groh, 2012; Hamari et al., 2014; Werbach, 2014a; Zhang, 2008).

Many parallel terms have been and are being introduced, such as 'productivity games', 'surveillance entertainment', 'funware', 'playful design', 'gameful design', 'behavioural games', 'game layer' or 'applied gaming'. Furthermore, there are closely related fields such as serious games, entertainment games, pervasive games and playful interaction. These are all part of a larger phenomenon described as 'the ludification of culture' (Deterding et al., 2011; Groh, 2012; Hamari et al., 2014). The fields and practices within the ludification of culture are very similar and overlap in multiple ways (application, origin, design practice, contents, etc.). Moreover, many terms actually refer to the same definitions. Therefore it is hard and partially unnecessary to pinpoint their exact differences and relations. In principle, serious games and gamification seem. However, some delineations are important to make.

There are two major differences to identify, when delineating gamification from the most related fields. The first is the difference between gaming and playing. Playing, or 'paida', is the primary form of spontaneity, joy and improvisation and occurs without pre-defined rules; whereas gaming, or 'ludus', is bounded by rules and arbitrary obstacles (Caillois, 1961; Warmelink, 2011). Gaming entails that involved actors competitively try to reach certain goals (Deterding et al., 2011). Therefore gameful

interaction entails a purpose or goal towards which an interaction steers, while playful interactions are more focused on creativity and a free interaction environment (Werbach, 2014b).

The second distinguishment is between whole games and partial games. Gamification never entails a whole standalone game, but is always a partial game or game parts. This relates to the non-gaming context that is often mentioned when speaking of gamification. Serious games and gamification seem hard to distinguish, especially because the design goals are very comparable. Both have desired instrumental outcomes (explicit targets, often in a business context) and desired experiential outcomes (certain experiences that are attempted to be invoked with the users or participants) (Deterding et al., 2013). But there is a distinguishable difference between serious games and gamification. Serious games, just as entertainment games, are standalone fully functioning games that do not necessarily directly relate to another (non-game) context. As Hamari and Koivisto state: “Gamification refers to adding ‘gamefulness’ to existing systems rather than building an entirely new game as is done with ‘serious games’” (2013, p. 2). Deterding formulates this as: “the designed [gameful] system itself has a hybrid nature, being neither ‘pure’ functional software nor a ‘full-fledged’ game” (2013, p. 2). Entertainment games just serve the purpose of fun and enjoyment, where serious games are meant to stimulate explicit or implicit learning for its users. However, for this learning, there has to be a transition phase from the game to the real world, which is often done in a de-briefing. In gamification, this transition is not so apparent, because the partial game or game parts are directly connected to the real world (Carron, Kordon, Labat, Mounier, & Yessad, 2013; Groh, 2012; Hamari et al., 2014; Wenzler, 2008). When plotting these differences on two axes, the fields can be clearly distinguished, see Figure 4 below.



**Figure 4: Differentiating gaming from playing, and full games from gamification. Adapted from (Deterding et al., 2011, p. 13)**

However, not all delineations are covered in this figure. Another important difference is that between pervasive games and gamification. Pervasive games are normal games that are extended out of the ‘magic circle’, i.e. the virtual gaming world, into the real world (Deterding et al., 2011). Examples are location-based games, live action role-playing games or augmented reality games. Also, there is the concept of persuasive technologies. Persuasive technologies are meant to affect behaviour and attitude directly, whereas motivational affordances used in gamification are aimed at changing motivations. Gamification tries to stimulate feelings of flow, mastery and autonomy (explained later on in this chapter), which originate from games, in a non-game context. It is not aimed at audio-visual

simulation (user interface design) or monetary rewards (loyalty marketing) to influence behaviour (Hamari & Koivisto, 2013).

## 2.2 Lessons from Past Gamification Studies

As stated in the introduction of this report, gamification has received a lot of criticism in the past couple of years, partly due to its non-scientific commercial niche and poor gamification applications in this field. Also, there are some relevant scientific case studies documented and summarised in literature reviews. In this sub chapter, the most relevant criticism and case studies will be looked into and summarised, in order to extract useful lessons for this research. The basis for this is provided by the recent work of Deterding (2014a, 2014c) and a literature review by Hamari, Koivisto and Sarsa (2014) and their sources.

### 2.2.1 Gamification Criticism & Potential Dangers

#### Criticism

There is a certain movement of researchers that argues against the term and concept of ‘gamification’ per se. This includes John Ferrara, Ian Bogost and Margaret Robertson. For example, Ian Bogost – who is a game designer and researcher – stated in a blog post, titled ‘Gamification is bullshit’: “Gamification is bullshit. I’m not being flip or glib or provocative. I’m speaking philosophically. More specifically, gamification is marketing bullshit, invented by consultants as a means to capture the wild, coveted beast that is videogames and to domesticate it for use in the grey, hopeless wasteland of big business, where bullshit already reigns anyway. [...] I’ve suggested the term “exploitationware” as a more accurate name for gamification’s true purpose, for those of us still interested in truth.” (2011) . Also, Ferrara has argued in the preface on his book on playful design: “Then, a graceless and overly memorable buzzword crashed into the culture: gamification. The name itself betrays the conceptual flaw of this fad, implying an experience that is by its nature something other than a game but dressed up to resemble one. [...] the technology doesn’t deliver what they thought it would, and the hype collapses into the trough of disillusionment.” (2012, pp. xv & xvi). Later on, he comes down to the argument that gamification contains a part of the word ‘game’, but often does not resemble the same as the work of true game designers. Game designers focus only on the player experience, as this experience is the only reason for people to play games. He argues there cannot be a trade-off between external objectives and enjoyable gameplay, because this will inevitably lower the overall quality. Therefore, gamification designers should also be focused on player experience (Ferrara, 2012). Margaret Robertson, another well-known game designer, has critiqued gamification on the basis that it takes the least essential aspects of games and presents them as the most essential. She stated in a blog post: “Gamification, as it stands, should actually be called pointsification, and is a bad thing because it’s a misleading title for a misunderstood process, although pointsification, in and of itself, is a perfectly valid and valuable concept” (Robertson, 2010).

The quotes above are taken from personal writings, from game designers, a few years ago. They were elaborated upon here because they indicate the general standpoint of game designers against gamification as a hype or marketing tool – which it was regarded in the first few years of its existence. However, with help of the work of Deterding, Hamari and many others, a more scientific and design-oriented view on gamification has been developed, which is the basis of this research. As Andreas Lieberoth recently stated: “Because just positing gamification as a lie like Ferrara or bombarding interested practitioners with proposed use cases is uninteresting and counterproductive, researchers and practitioners alike are realizing the need to knuckle down and register effect data—not just from subjective evaluations but also observable behavior and net gains in the intended setting.” (2014, p. 16). The critics did provide valuable feedback, not only in blog posts, but also in scientific



publications. The criticism on gamification can be summarised and categorised into four sectors (Deterding, 2013, 2014a; Nicholson, 2012; Werbach, 2014b):

1. Not systemic: certain system elements are addressed rather than looking at the general system qualities, the dynamic interaction of users with the system as a whole and the user experiences as result of that interaction.
2. Reward-oriented: extrinsic rather than intrinsic motivation (the difference will be addressed later on in this chapter) is invoked, mainly by focussing on rewards.
3. Not user-centric: a very instrumental perspective, based on business goals of the system owner, is used and the users' goals are ignored or even impaired.
4. Pattern-bound: this is the most heard cluster of criticism: only a small set of game elements are used (mostly points, badges and leader boards), instead of expanding the view and focussing on structural game qualities that need to be invoked, regardless of the specific game elements that are used.

### Potential Dangers

Apart from criticism based on the concept, application and content of gamification, potential dangers of gamification have also been identified. The first is the danger of exploiting users to perform activities that they actually do not wish to perform. Gamification should not be used in a manipulative context, where motivational affordances are deployed only because of their known stimulating effect, not because of a gameful and enjoyable experience they might offer. There have been cases in an enterprise setting, where gamification was used to positively stimulate employees to perform better by showing them internal leader boards and KPI's. However, the effect was the other way around and employees felt they were constantly monitored, pressured to work harder and the social cohesion between employees was negatively affected (Werbach, 2014b). Another potential danger of gamification is so-called 'overjustification'. When rewards are implemented as a part of gamification (can be tangible and intangible) to stimulate the performance of a certain task, the motivation to perform the task can shift from being mostly intrinsic to extrinsic (Groh, 2012). In other words, the user stops doing the task because he/she likes to, but does it because of the reward. A study by Lepper et al. first demonstrated this phenomenon, in which children would draw more pictures, but of less quality, if they were given a monetary reward to draw pictures. Moreover, the children indicated they did not like drawing anymore when the payments stopped, even though they did like drawing before they got paid (Lepper, Greene, & Nisbett, 1973).

### 2.2.2 Relevant Case Studies

With the general advices from gamification critics set out, an in-depth look into relevant scientific case studies is needed. Because from there, it might be possible to identify best practices to overcome the mentioned criticism and potential dangers, so they can be incorporated into this research.

The reference point here is a number of existing gamification studies and more specifically: the most recent literature review of empirical gamification studies by Hamari, Koivisto and Sarsa (2014). They consider 5 of the 24 empirical gamification studies they reviewed as well executed, due to proper experiments or proper psychometric measurements and due to sufficient sample sizes. Also, they state: "As the research on gamification progresses, care should be taken to ensure that future results are more comparable. This can partly be ensured if future studies will build upon the previously well executed inferential studies" (Hamari et al., 2014, p. 3030). The relevant cases were thoroughly examined and are described in appendix A, including the evaluation and design methods that were used.



The most important lessons learned from these cases are the following (Hamari et al., 2014):

- Most of the studies concluded that gamification only works in part of the hypothesised relationships between game elements and outcomes. The context being gamified and the qualities of the users were identified as the main underlying confounding factors. This is recognised by Deterding (2014a).
- No structured game or gamification design methods were used, or at least not set out in the papers and reports. Game elements were selected and implemented on a logical (or seemingly random) basis, following the nature of the product or service to enhance.
- Social elements and social interaction are a very important factor for gamification success (Hamari & Koivisto, 2013).
- When gamifying a utilitarian system, where the goal is to increase the quality and quantity of the user-generated content, providing an explicit challenge with feedback on the performance in regard to the challenge, stimulates the users in such a way that their performance increases. Performance feedback and social identification are also beneficial for content quality and quantity. However, a challenge without feedback can be disadvantageous for such goals (Jung, Schneider, & Valacich, 2010).
- An experimental setup with users randomly assigned to one or more experimental groups and one control group is quite common, with a set of hypotheses defined prior to the experiment: controlled experiment. Especially 2x2 setups are useful when testing two different game elements. Gamification is evaluated based on a difference in values of one or more KPI's (or dependent variables) for the experimental and control group, before and after the gamification. As Farzan et al. stated: "running controlled experiments [...] can be an effective method for determining the strengths and weakness of different [motivational] incentives, which can aid designers in deploying the most effective mechanism for their community" (2008b, p. 572).

### 2.2.3 Designing an Experiment

In order to be able to design the setup of the experiment, which will be used to assess the effect of a gamification treatment on the behaviour of the users in a case online marketplace, some lessons and pitfalls from existing research are set out here.

#### Pitfalls of Existing Empirical Studies

In their recent literature review, Hamari, Koivisto and Sarsa (2014) identified the following methodological limitations in existing empirical gamification studies, which are to be avoided in this and future research:

1. Too small sample sizes, with N=20;
2. If user experiences and attitudes were surveyed ex post, no validated psychometric measurements were used;
3. Lack of control groups;
4. Multiple game elements were implemented as a whole, so that the individual effect of each element could not be measured;
5. Only descriptive statistics were presented, without mentioning relationships between constructs;
6. Very short timeframes for experiments, causing the novelty of an implemented game mechanic itself to be a potential factor, which was not taken into account;
7. Unclear reporting of results;
8. No study used multi-level measurement models, including the game mechanics, game dynamics / psychological outcomes and behavioural outcomes

These pitfalls will be used to develop the experiment setup, described in chapter 4.2. Also, in chapter 5.1.2, they are used to reflect on the limitations and the methodological validity of the experiment in hindsight.

Also, Hamari states in his literature review that methodologically, it is important to keep the first treatment designs as simple as possible in terms of incorporated motivational affordances, because this way the effect of individual motivational affordances or game elements can actually be measured. If successful, more affordances to enhance the design can be added later (Hamari et al., 2014). However, the perspective and design method adopted see gamification as a holistic design practice, where the user context and behaviour are intertwined with the gamified system and motivational stimuli emerge from this as a whole. Changes in behaviour can therefore not be linked to a specific game element or motivational affordance. Nor can possible conclusions on effects of individual affordances be generalised to other cases, because there the context and users are again completely different.

### Behaviourist Versus Cognitivist View

Deterding states that as long as each researcher has a different operationalisation of enjoyment or performance, the only findings that will be reported are that ‘what people call gamification can have positive effects’ (Deterding, 2014b). According to him, gamification research needs to unpack the black box of the human psyche and use psychological measurements in empirical studies. Psychological outcomes should be measured as mediation variable between the effect of motivational affordances on user behaviour (Hamari et al., 2014). This stems from the cognitivist psychological view, whereas the behaviourist view seeks to explain loops of action, feedback and response by user behaviour only. Not having psychological measurements during an experiment reduces bias, as participants can still be unaware of participating. Also, with a behaviour-oriented experiment, large numbers of respondents can more easily be used. The big disadvantage of this approach is that the ‘why’ question can not be filled in (Werbach, 2014b).

Even if psychometrics are used in the experiment in this research, the link between a psychological outcome and a motivational affordance is very difficult to construct, because it conflicts with the holistic gamification perspective that was adopted. Also, the double blind nature and large respondent base are seen as important characteristics to evaluate the effect of a gamification treatment. Moreover, using psychometrics is not possible due to case restrictions. Therefore, the behaviourist view is used for the experiment setup (see chapter 4.2) and only behavioural data is measured.

## 2.3 Underlying Theory of Gamification Perspective

Given the adopted perspective of gamification as socio-technical view, which recognises the gaming elements, but also the context and motivational factors as important building blocks (Deterding, 2014a), it is useful to look into some theories that are the components of this view, in order to create a better general understanding.

### 2.3.1 Motivation

#### Self-Determination Theory

To have motivation for something means to be energised and activated to do something, which can highly differ in level for each individual (Ryan & Deci, 2000). Also, the type (or orientation) of motivation can differ. In their self-determination theory, Ryan & Deci (2000) characterise two main types: extrinsic and intrinsic motivation. An extrinsic motivation consists of a certain desirable outcome, often an incentive from one’s environment (Zhang, 2008), while an intrinsic motivation is purely fuelled by enjoyment. A common practice to find out someone’s intrinsic motivation for a certain activity is to conduct a so-called laddering interview, whereby each answer is followed by another ‘why’ question, until the most basic reason is given (Deterding, 2014c).

According to self-determination theory, intrinsic motivation is based on three main principles, for which all people have an innate psychological need (Groh, 2012):

1. *Competence*: experiencing the feeling of mastery, becoming good at something, overcoming challenges;
2. *Relatedness*: interact with and be connected to others, do something that is meaningful to others;
3. *Autonomy*: controlling your own life, voluntarily participating in activities.

According to many gamification researchers, another innate psychological need, which fuels intrinsic motivation, is fun (Deterding, 2014a; Huotari & Hamari, 2012; Richards, Thompson, & Graham, 2014; Werbach, 2014b). Unfortunately, fun is very subjective and hard to define.

### Motivational Affordances

What are motivational affordances and how do they relate to game elements? According to Deterding, affordances are “not an objective feature of a design element, but a relational quality of both object and subject” (2014a, p. 217). They emerge from a complete system, rather than a single stimulus or design element. Thus a game element can afford a certain feeling or motivational fulfilment, based on the system, the user and his/her specific state and goal. The experiential perspective of Hamari is largely based on the concept of motivational affordances, which was originally conceived by Zhang, who used them to illustrate that IT systems can support people’s motivational needs, but its design is dependent on users and their context (Zhang, 2008).

## 2.3.2 Meaningful and User-Centered Gamification

One of the basic principles of the gamification perspective of Deterding and his design method (2014a, 2014b) is meaningful gamification, by focusing on the users and their context. By acknowledging this principle, much of the gamification criticism is theoretically overcome. According to this principle, the needs and goals of the users should be placed above the needs of the organisation, which wants to gamify their system. User and business goals should at least be aligned, even though they might seem disjointed in first instance. A positive and meaningful experience for users will be more beneficial for these organisations in the long term (Nicholson, 2012). In other words: focusing on connecting to intrinsic motivations of users rather than using rewards as an extrinsic motivation. This automatically involves a bottom-up construction of a holistic gamified system, rather than a top-down application of turnkey game elements. In order to do this, a gamification designer cannot proceed without knowing the user context as much as possible (Richards et al., 2014).

## 2.4 Gamification Design Method

Here, the gamification design method that was chosen as a reference point for design in this research in chapter 1 will be elaborated upon. The gameful design (or gamification design) method ‘Lens of Intrinsic Skill Atom’ by Deterding (2014c).

### 2.4.1 Gameful Design by Lens of Intrinsic Skill Atoms

Taking into account and recognising the current state of gamification, its pitfalls, the misuse, the evolving definitions and all the general theory, Sebastian Deterding mentions the foundations of a gamification design method: “if the re-envisioned scope of gamification are socio-technical systems, if its re-envisioned goal is motivational experiences, and if motivational experiences are systemic, emergent affordances, then a promising re-envisioned gamification design method would entail formalising desired motivational experiences in the form of design lenses, using these lenses to analyse target activities, and then engage in iterative experiential prototyping until the total prototyped

socio-technical system affords the targeted motivational experiences” (2014a, p. 320). Based on this view, he developed the gameful design method and ‘Lens of Intrinsic Skill Atoms’, which contains the steps as displayed below. The gameful design method consists of an innovating mode and an evaluating mode. The innovating mode is where the designer tries to create a new system around the needs of target users, whereas the evaluation mode is where skill atom components (or game mechanics) are used to enhance an existing system (Deterding, 2014c).

For this research, the evaluating mode (rather than the innovating mode) is useful, because the case concerns an already existing online marketplace to which gamification will be applied, in order to improve the platform. The method is used to extract the skill atom(s) in the system and, by applying different design lenses, issues in this skill atom are depicted and possible solutions are developed (Deterding, 2014c). “The innovating mode of gameful design serves to create a new system around a target user need, whereas the evaluating mode serves to analyse and improve an already existing system. In both, the approach is to identify intrinsic challenges of the target activity and then ideate and iteratively prototype and test (new or amended) skill atom designs that structure those challenges in a motivating, enjoyable form, involving additional design lenses that focus qualities of either the total skill atom or individual components.” (Deterding, 2014c, p. 31)

The evaluating mode consists of the following steps (Deterding, 2014c):

1. Strategy
  - a. Define target outcome and metrics
  - b. Define target users, context, activities
  - c. Identify constraints and requirements
2. Research
  - a. Translate user activities into behaviour chains (optional)
  - b. Identify user needs, motivations, hurdles
  - c. Determine gameful design fit
3. Synthesis
  - a. Identify skill atoms of existing system for opportune activities/behaviours
4. Ideation
  - a. Brainstorm ideas using design lenses
  - b. Prioritise ideas
  - c. Storyboard concepts
  - d. Evaluate and refine concept using design lenses (optional)
5. Iterative Prototyping
  - a. Build prototype
  - b. Playtest
  - c. Analyse playtest results
  - d. Ideate promising design changes

*Repeat steps a-d until desired outcome is achieved. Increase prototype fidelity as playtest results approach desired outcome.*

Most steps are quite self-explanatory or will be explained later on in this chapter. The concept of skill atoms and their ‘lens of intrinsic skill atoms’, which form the core of the gameful design method will be elaborated upon in this chapter (Deterding, 2014c).

### 2.4.2 In-depth Look at Skill Atoms

The general form of a skill atom can be seen in Figure 5 below. The skill atom and its lens are the base of the gameful design method. Deterding describes a skill atom as follows: “the smallest self-contained system: as a system or ‘atom’, it consists of smaller reoccurring particles, yet it cannot be broken into these without losing its systemic ‘gaminess’” (Deterding, 2014c, p. 28).

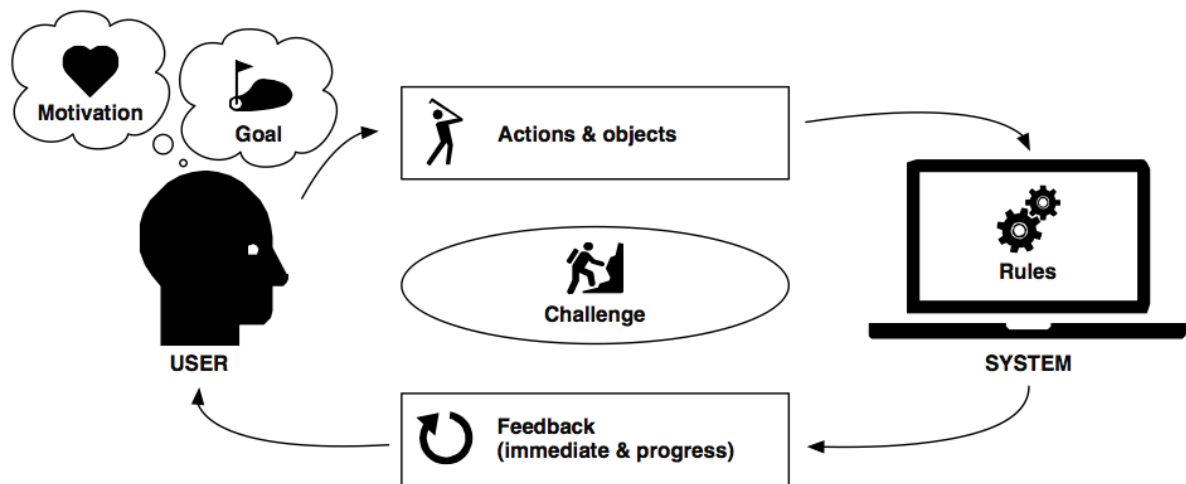


Figure 5: General form of a skill atom (Deterding, 2014c)

The elements of a skill atom are explained in detail in chapter 3.2.3, where the design method is actually applied to the case of this research. In general, a user takes an action, which is the input to and must comply with the rules of the system. The system gives feedback to the user. By interacting with the system, a user can achieve his/her goal and become more competent to solve his/her challenges regarding the system (Deterding, 2014c). When simplified, this can be described as an activity loop containing motivation, action and feedback (Werbach, 2014b)

### 2.4.3 Suitability for this Research

The gameful design method by Deterding (2014b) is very suitable to be used in this research, but a few amendments were made based on previously discussed theories in chapter 2 or on general observations and conclusions that can be defended on a logical basis. The amendments are clarified in this sub chapter.

#### Step 1: Strategy

In step 1b, it is not only important to identify target users and their activities, but also to depict the context in which they act and visit the online marketplace, as well as the cultural, political, economical or other influences that are at play. This increases the ability to think as the online marketplace user during the design, which is vital for any form of user-centred design (Richards et al., 2014).

Step 1c from the original method only focuses on technical, legal and resource (time, budget, people) constraints and requirements. In the list of requirements, it seems of great added value to include the general positive qualities that users associate with the online marketplace (Richards et al., 2014). These qualities can be obtained by user interviews, historical data analysis, observation, desk research, etc. An example can be that users find the online marketplace better than competing online marketplaces because it has a certain feature. If this is the case, the gamification treatment should not jeopardise this feature, because this will probably negatively influence users' activities on the platform, regardless of the quality of the gamification treatment.

#### Step 4: Ideation & Step 5: Iterative Prototyping

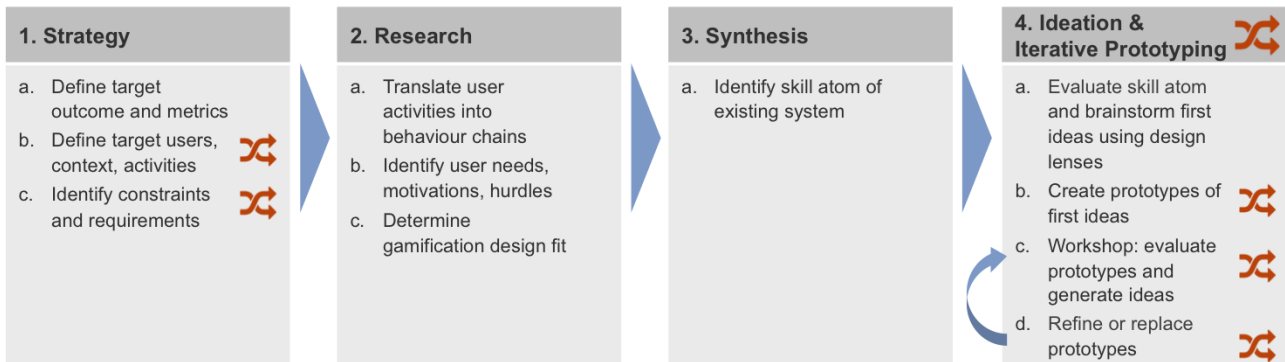
The last amendment concerns the method with which ideas are generated. Ideation will partly be done by the designer and partly by other participants, in workshop form. For the purpose of user-centred design these participants should be online marketplace users, or resemble online marketplace users as much as possible. This fits with theories on meaningful gamification and user-centred design, thus incorporating the advices and critique of gamification design as identified in chapter 2.2.1 (Deterding, 2014a; Nicholson, 2012). Deterding concluded, by putting his gameful design method into practice, that asking questions like 'how might a game around goals look?' were

still too abstract for participants without game design experience (Deterding, 2014c). Therefore it is wise to make the input for participants in ideation workshops as concrete as possible. A solution is to already provide prototypes or mock-ups of some of the first ideas. This way, the participants can propose amendments for the prototypes but also generate new ideas. Using prototypes to receive qualitative feedback and make quick adjustments to early designs is also recognised and recommended as a suitable method in literature, which can be complementary to a controlled experiment (Ron Kohavi et al., 2008; Ries, 2011).

Following the above, it seems logical to combine step 4 and 5 (ideation and iterative prototyping). This combined step starts with the generation of first ideas by the game designer. The following iterative prototyping cycles will contain prototyping ideas, evaluating the prototypes in a workshop and adjusting the prototypes based on the workshop. Also, new ideas can be generated in a workshop and prototyped for the next cycle (which is also inherent to iterative prototyping and playtesting (Ferrara, 2012)). The cycles can be repeated as often as needed.

### Final Design Method

An extensive description of the steps in the method, as it was applied in this research including the amendments, is featured in appendix B. An overview of these steps can be seen below in Figure 6. The orange arrows are located at steps in which a small amendment to the method by Deterding (2014b) has been made (corresponding with appendix B and the above).



**Figure 6: The gamification design method (Deterding, 2014b) with amendments, as applied in this research**



## 3 GAMIFICATION DESIGN

---

*In this chapter the application of the gamification design method, as laid out in the previous chapter of this report, to a case online marketplace is described. First an introduction to the case is given and its suitability to help answer the research questions is tested. A summary of the design process describes the road towards the final implemented treatment, 'the Selling Assistant', as result of this design process. This iterative process, given by the theoretical gamification design method that was chosen, is documented in more detail and in chronological order in appendix C. The following sub research question will be provisionally answered in this chapter:*

4. *Through application of the chosen design method: what is the most promising treatment to gamify an online marketplace in order to increase the amount of user-generated content?*
- 

### 3.1 Case Introduction: OLX

Before critically looking at and adjusting the design method by Deterding (2014c), 'Lens of Intrinsic Skill Atoms' and applying it, this sub chapter introduces the case. This includes case data to understand the context and an explanation of the suitability of the case for this research.

#### 3.1.1 Introduction & Delineation

OLX (abbreviation of 'online exchange') is a global online platform (owned by digital and media company Naspers), which is active in over 100 countries in more than 50 languages (Naspers, 2014). It is active in countries in South America, Africa, Eastern Europe and Asia. Here, most online marketplaces are in a growth stage and the countries can be described as emerging markets (Lowe, 2014).

This research focuses on the OLX India region, because of practicalities (available resources to conduct an experiment, in terms of time, support and user base. OLX India (<http://www.olx.in>) is an online marketplace. More specifically, OLX India (from here on 'OLX') is a classifieds website, where users can sell and buy unused second hand items, such as clothing, cars and electronic equipment. One side of the market concerns the sellers. Users can post content in the form of a listing (or 'ad'), which is essentially an advertisement of something they have for sale. This can be anything from a new mobile phone to an old pair of jeans to a baby-sitting service to an apartment for rent. The listings are classified into several categories and sub categories, hence the name 'classifieds'. The other side of the market is the buyers: users can reply to the listings that have been posted, by contacting the seller with a price offer, questions or something else. Users can be either buyers, sellers or buyers + sellers, depending on their activities on the website.

OLX only facilitates the connection of these two sides of the market. There is no monetary transaction possible, so this needs to be arranged between buyers. All functionalities for OLX users are currently free. It is possible to create an account, from which you can edit or delete own listings and save favourite listings from others. However, an account is not needed for any functionality of the website. OLX operates on three platforms: a desktop website, a mobile website and a mobile application (iOS + Android).

In this phase, OLX is focused on growth of its user base (Griffith, 2014). In order to enable this, both the buyer and seller side of this online marketplace need to increase, resulting in the network effect

and typical chicken-egg problem as described in the introduction of this report. If the offer of listings is not large and diverse, buyers who visit OLX will be less likely to return. Vice versa, if OLX does not receive a lot of visiting buyers, its reputation as a website where you can sell fast and for a good price will decrease, leading to a smaller number of listings posted. 'Marktplaats' is an online marketplace and is the leading Dutch classifieds website, which operates almost as a monopolist and is over 15 years old. It has 1.3 million daily visitors, out of ~17M inhabitants, who post around 350,000 listings, which is one listing for fewer than every 4 visitors (eBay International AG, 2015). OLX is market leader in India, but has not reached such high levels of user activity nor user penetration (Batra, 2014). Therefore, one of the goals of OLX lies in increasing the number of listings that are posted. OLX does this with an extensive marketing campaign and optimisation of their mobile application and desktop website (Griffith, 2014).

OLX India is the 42<sup>nd</sup> largest website in India, in terms of visitors. It received 16.5M visitors in March 2015, who spent 9.5 minutes viewing more than 11 pages on the site on average (SimilarWeb, 2015). The time spent indicates that visitors are not only there to consume information, but also provide input (user-generated content). As one can imagine, placing a listing takes more time than reading the average web page.

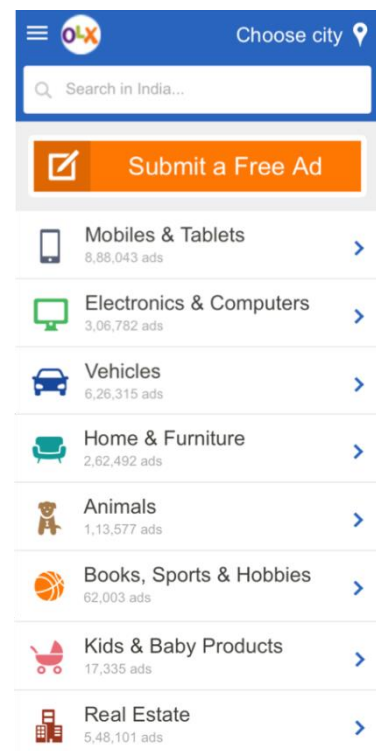


Figure 7: OLX India mobile website homepage (OLX, 2015a)

The focus of the gamification design and implementation will be on the OLX mobile website<sup>1</sup>. In the past 3.5 years OLX India traffic has grown 150 times and mobile traffic is predicted to grow even more (now 80% of all traffic) (PTI, 2014). Also, on online marketplaces in general, mobile traffic is becoming more and more important, due to the accessibility, location detection possibility and embedded photography options (Jordan & Hariharan, 2015). There is much potential for the OLX mobile website. The user base increases at rapid speeds, but the number of users who posts a listing could still increase. Also, the number and frequency with which users post a listings can be improved, when comparing the OLX mobile website to comparable online marketplaces inside and outside of India. Moreover, not every visitor to the mobile website continues to view more than one page (so does probably not find what he/she is looking for). Finally a relatively high share of listings is not of optimal quality. By improving prices, descriptions and pictures included in listings, the overall quality of the marketplace improves, which makes the platform more valuable for buying visitors<sup>1</sup>.

A screenshot of the OLX mobile website can be seen in Figure 7.

### 3.1.2 Suitability of OLX as Case

An online marketplace case is needed for which a gamification treatment can be developed and in which the design can be empirically evaluated. The case should resemble the general online marketplace and UGC problem as much as possible and allow for a general gamification design and evaluation method to be applied.

All of listings on OLX are created only by the users themselves and thus are UGC. The chicken-egg problem is applicable in the sense that OLX is trying to become the largest and thus only classifieds in India and for now, growing their user base is extremely important. Given the network effect,

<sup>1</sup> Personal communication with CTO (Naspers Classifieds), Business Improvement Leader (Naspers Classifieds),

increasing one side of users (sellers) will have effect on the other side (buyers). Therefore, the challenges of OLX as described very closely resemble the more general online marketplace problems that are mentioned in literature (Cambini et al., 2011; Hagi, 2014; Seamans & Zhu, 2014).

The mobile website is specifically suitable because it has a high number of daily active users, but they post a relatively low number of listings. Identifying the specific reasons for and the possibilities to overcome this problem can be done with the gamification design method. The digital nature of the mobile website allows for iterative prototyping and playtesting (Deterding, 2014c). As for evaluation, it's a digital online platform, so a controlled experiment with real users can easily be set up. This method is often applied to websites, when testing the effect of different versions on user behaviour, with predefined hypotheses and metrics. In those cases it is often called 'split-testing' or 'A/B testing', which is the same as a randomised controlled experiment, but a term used in online jargon (Crook, Frasca, Kohavi, & Longbotham, 2009; Ron Kohavi et al., 2012, 2008).

## 3.2 Applying the Gamification Design Method

The gamification design method by Deterding (2014b), with the amendments as described chapter 2.4.3, was applied to the OLX mobile website, in order to arrive at the most promising gamification treatment to increase the amount of UGC on OLX. Here, the outcomes of this process are described, but not in chronological order. For instance, information on certain requirements that were incorporated into the final gamification treatment may have been gathered during the workshops (which is a later step according to the design method used). The final outcomes are described, with the sub chapter numbers and titles following the structure of the steps in the design method (as depicted in Figure 6 in chapter 2.4.3 and as extensively described in appendix B).

### 3.2.1 Strategy

#### a. Define Target Outcome and Metrics<sup>2</sup>

The target outcome of the gamification treatment is to increase the number of new listings, measured by the average number of listings posted per user. The number of listings should increase relative to the number of users, because the number of users is an external variable that is not part of the research, but does influence the number of new listings. Therefore increasing the average number of new listings posted per user is the actual target outcome.

#### b. Define Target Users, Context, Activities

##### Target activity

The target activity for users that the gamification treatments need to stimulate is to post a listing that meets the basic OLX requirements to be featured on the website. This is the only activity that will directly increase the number of listings per user. Posting a listing can be stimulated by resolving motivational issues and/or adding fun, as will be explained later on in this chapter.

##### Theoretical user context

As noted in chapter 2, gamification literature clearly indicates the importance of contextual factors as a part of user-centred design. Three theoretical perspectives can partly indicate these contextual factors (Hamari et al., 2014):

1. The voluntariness of people who use system: this indicates the nature of behaviour of users and their general attitude towards the system. OLX is a voluntary free service, which users

---

<sup>2</sup> Personal communication with Senior Product Manager Mobile (OLX India) & CTO (Naspers Classifieds). September & October 2014.

choose to use themselves. There is no external organisation or institution that obliges or pushes for OLX use.

2. The system's purpose: utilitarian or hedonic? In other words: is it used functionally or for fun? OLX is in principle a utilitarian system to buy and sell used goods, but can be used for fun (for example by users who just browse listings in their spare time).
3. Are users involved in the system on a cognitive or affective base? Both, users visit OLX because they have certain positive feelings towards it (affective) and also because they know its function and its possibilities (cognitive)<sup>3</sup>.

#### Practical user context

Apart from this, there is more specific contextual knowledge, which has partly determined the design process in some stage. This knowledge was gathered through a workshop with five OLX users at Delft University of Technology (see appendix C.3 for details) and through an internal report by OLX (2014a)<sup>4</sup>.

- The internet connection in India is generally slow (EDGE).
- Many users have a feature phone with limited computational capability and no guaranteed support of latest mobile website development frameworks.
- Selling items online is not embedded into Indian culture. People rather sell their used goods to a local Kabadiwala (people who go around the homes to buy second hand items and pay the price of the item materials) as garbage or to local shops. This way they are offered a price instead of having to suggest one by themselves. However, when selling to a local shop or Kabadiwala, the price they receive is generally lower than the price they could get on OLX.
- Residents of smaller cities are good to target for OLX, because they do not have the degree of access to all types of second hand shops that big cities do provide.
- Almost everyone knows the OLX brand and slogan 'OLX pe bech de' ('sell it on OLX') (Batra, 2014), but not everyone is aware of what OLX actually is.
- If users have not experienced a successful OLX deal, the potential reward of selling something is not enough to convince them.
- Any changes to the mobile website should look like they are integrated into the OLX site and associated with OLX, otherwise people will think it's third-party advertising. This especially holds for pop-ups.
- Most mobile users download the OLX app and use that instead of the mobile website. This is not true for new or infrequently returning users, or for users with a feature phone rather than a smartphone.

#### c. Identify Constraints & Requirements

The user context, the time frame of the research and the technical possibilities determine the constraints and requirements for the design process. This concerns everything but the user behaviour that the treatments need to stimulate, because this is already considered in the metrics.

#### Requirements

The requirements describe aspects, qualities or forms with which the gamification treatments *should* comply.

- Simple to view, navigate and interact with on both feature phones and smartphones;
- Fast in load times (so as few images, videos, etc. as possible);
- Free to use;
- Fit seamlessly into the current OLX mobile website.

<sup>3</sup> Personal communication with CTO (Naspers Classifieds) & Business Improvement Leader (Naspers Classifieds). September & October 2014.

<sup>4</sup> Several OLX team members and ~40 OLX users were interviewed on their experiences with OLX and field research was done in order to examine user contexts.

### Constraints

The constraint describes aspects, qualities or forms with which the gamification treatments *must* comply.

- Implementable with HTML5 and Javascript within a few days (this is based on time and resources within the research).

### General OLX qualities

Qualities associated with OLX (or Indian classifieds in general), according to users (OLX, 2014b)<sup>5</sup> and workshop participants are:

- Simplicity of usage
- Free
- Quality, original content
- Speed of transaction

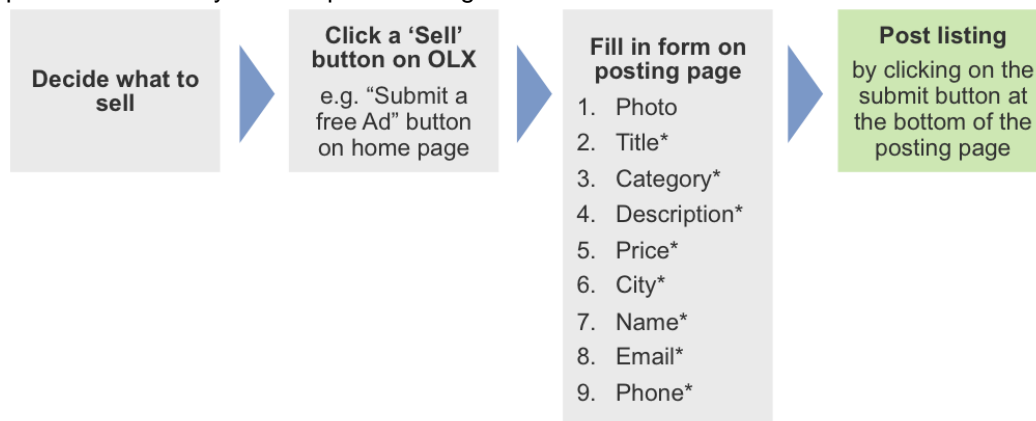
These can be recognised as general values of OLX and mobile sites to which users are attracted and which must not be harmed by a gamification treatment. Even though the gamification treatments are primarily focused on increasing the number of new listings, they must still consider these general values.

A general research under users of mobile websites (Keynote Systems, 2012) reveals the main frustrations that are encountered when visiting a mobile website and reasons not to visit this site again. The top reasons include slow page loads, sites that are not optimised for phone screens (too big, too much text), difficult interaction and difficult navigation.

## 3.2.2 Research

### a. Translate User Activities into Behaviour Chains

The target activity for users is to post a listing. Figure 8 below shows the activities OLX users must perform when they want to post a listing.



**Figure 8: Behaviour chain for OLX users who want to post a listing**

### b. Identify User Needs, Motivations, Hurdles

The following information was gathered from multiple sources: interviews with OLX employees<sup>6</sup>, workshops with users and an OLX expert (see appendix C.3 and C.5), from OLX internal research,

<sup>5</sup> A questionnaire conducted amongst many (potential) users.

<sup>6</sup> Personal communication with Senior Product Manager Mobile (OLX India), Business Improvement Leader (Naspers Classifieds), UX Design Manager (OLX India) & CTO (Naspers Classifieds). September 2014 – February 2015.

also mentioned before (2014a, 2014b, 2014c)<sup>689</sup>, and from external sources (Batra, 2014; Sathe, 2015). Definitions of needs, motivations and hurdles stem from Deterding's gameful design method (2014c).

### Need

A need is an underlying reason for OLX users to perform the target activity, thus to post a listing, so it has to do with the most basic of intrinsic motivations, which are not only applicable to OLX. This need can be one or a combination of the following:

- do interesting, exciting, new things (fun);
- earn money;
- have social interaction.

### Motivation

The motivations for OLX users are the more direct reasons for them to actually post a listing or start the posting process. There is only one real motivation:

- users want to sell an item.

Different contexts in which this motivation can occur are:

- to stimulate recycling of products;
- to give something back to the community;
- an item is upgraded to a new version, so the old version needs to go somewhere;
- to clean up the house;
- an item is never used;
- an item was bought but is not suitable for own use.

### Hurdles

The practical challenges that OLX users face when posting a listing can be translated as hurdles:

- Users find it hard to estimate which items are suitable for selling and might be of value to others.
- Users have difficulty selecting a fitting price for an item.
- Users find it hard to write an attractive title and description for a listing.

### Discouragements

Next to user needs, motivations and hurdles, also the inverse motivations (discouragements) need to be identified, because these are essentially the reasons to not post a listing. In other words, these are discouragements that need to be taken away for users to post a listing. Some of these might be able to change with the influence of a gamification treatment:

- Don't have anything to sell.
- Never considered to sell something online.
- Prefer to donate, i.e. don't like the concept of selling used goods.
- Don't see why other people would value my old stuff.
- Did not receive enough/expected/any response the first time.
- Too much hassle.
- Resale value of an item is too low.
- Too much emotional attachment to an item.

## **c. Determine Gamification Design Fit**

The following questions need to be answered in order to assess the suitability of the gamification design method as a possible way to intervene in the case (Deterding, 2014c, pp. 34–35; Werbach & Hunter, 2012, p. 49). At the beginning of this chapter, the suitability of the OLX case to answer the research questions was evaluated. This concerns the suitability of the chosen design method by Deterding (2014c), given the case information that was unfolded in the previous paragraphs.



1. *Does the activity connect to an actual user need?*

Yes, because posting a listing will directly influence most users' needs. Posting a listing can be fun, new and exciting. Also, selling an item to someone, including digital and physical contact, involves social interaction. Moreover, selling an item generates income.

2. *Is lacking motivation a central issue or opportunity (and not e.g. poor usability)?*

Yes, one of the main reasons for users to not post a listing is the fact that they do not know what to sell (they assume they do not have any items of value for other people).

3. *Does the target activity involve an inherent challenge with a learnable skill?*

Yes, posting a listing gets easier every time. If one item to sell can be identified and sold, users will learn that they can sell almost everything on OLX. Also, writing an attractive text for the listing, making the right photos and choosing the right price are learnable skills.

4. *Is affording experiences of competence an effective and efficient way of improving motivation (and not e.g. defusing fears)?*

Yes, because the main user needs that underlie the motivations to post a listing can all be experienced once a successful item sale through OLX has been done. This generates an experience of competence. However, one could also identify this as 'defusing fears'.

### 3.2.3 Synthesis

#### a. Identify Skill Atom of Existing System

This step was simplified by creating just one skill atom, which considers the most influential and opportune behaviour that users can perform in order to change the metric(s). OLX already identified their metrics and desired outcome, so only the relevant part of the platform needs to be translated into a skill atom; rather than translating the entire system into multiple skill atoms and thereupon finding room for improvement. Posting a listing on the OLX mobile site can be fitted into a skill atom as shown below in Figure 9. Note that only the most important elements from the previous design steps have been translated into this skill atom, based on the internal and external sources mentioned in the previous parts of 3.3. Also note that the user motivations have been put into the 'Goal' box and the user needs have been put into the 'Motivation' box. In the current design method these terms do not naturally connect. This inconsistency will be further discussed at the end of this report.

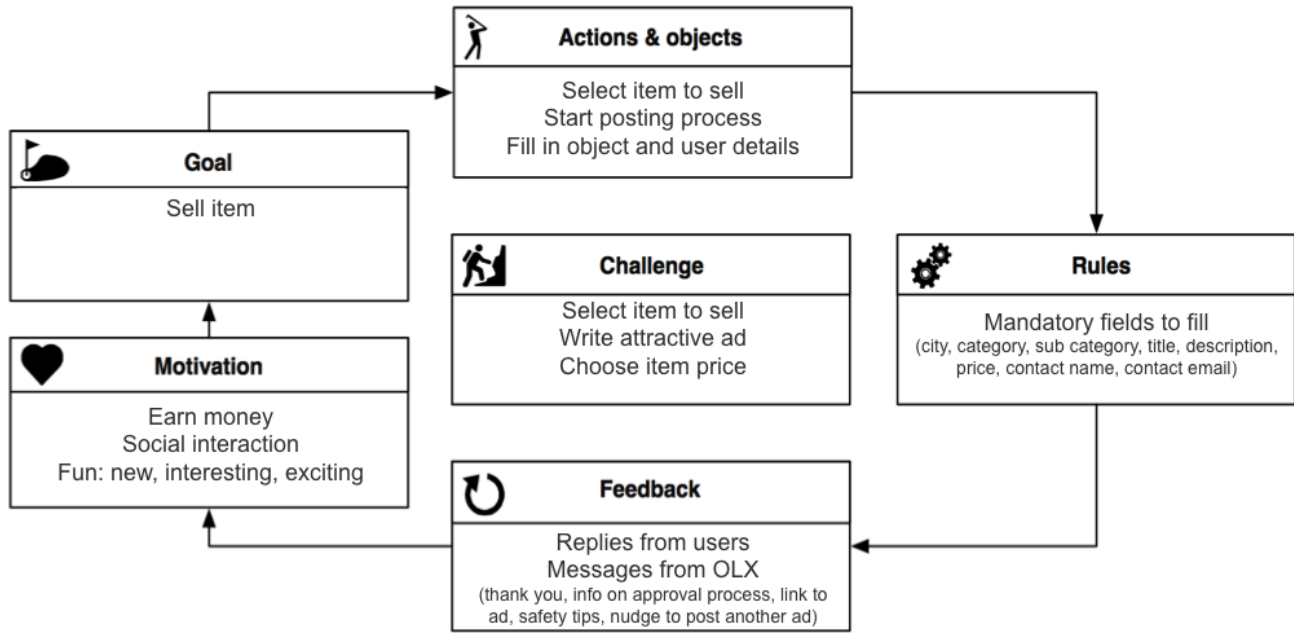


Figure 9: Skill atom for posting a listing on OLX mobile website

Below, general meanings of the elements of the skill atom are explained (deducted from the gameful design method (Deterding, 2014c)) and the specific interpretation of these elements for OLX is clarified:

- **Motivation:** the intrinsic psychological need that is satisfied by interacting with the system. By satisfying this need, the user is driven to continue to engage with the system. Such a definition relates more to the *needs* than the *motivations* from step 2b before and thus the *needs* from step 2b were filled in here.
- **Goals:** the specific state that is suggested by the system ('call to action') and that is to be actively pursued by the user. Clearly, this relates to the *motivation* from step 2b: selling an item. In terms of OLX system state, this is the posting of a listing, because the actual selling is done outside of OLX.
- **Actions & Objects:** the system's objects with which users interact and the actions with which they perform these interactions. I.e.: "What does the user do with what to achieve the goal(s)?" (Deterding, 2014c, p. 39). In the OLX case, users have to select an item they wish to sell (an offline process, not directly connected to OLX website objects). Then they have to start the posting process by entering the posting page, through one of the 'sell' buttons on various pages of the site. Once on the posting page, users have to fill in the item (price, description, photo's, (sub)category, title, etc.) and their own (name, location, email, etc.) details.
- **Rules:** the specific rules regarding the user actions in the system. On the OLX mobile website, these rules are the mandatory fields that need to be filled in when posting a listing (see the skill atom above). They are required in order to perform the action of posting a listing.
- **Feedback:** information aimed at informing the user of the system changes as result of their action(s). OLX gives feedback in the form of a confirmation page with a thank you message, which confirms the successful posting and informs the user about the posting approval process ("All set! Your listing has been submitted for approval. You will see it on OLX within the next 2 hours. You can view your listing here <link>."). The same information is sent to the user via an email. Also, a button and text nudge the user to post another listing right away. Other feedback that users receive (once the listing is approved) is in the form of feedback from other users: replies to their listing. Where the OLX system feedback is aimed

at the posting of a listing, replies from others inform the user of the chance of success of reaching his/her actual goal: selling an item.

- Challenge: this concerns the challenge users can perceive when looking at the gap between the normal state of the system and the goal they wish to achieve. Mostly, challenges connect to learnable skills, making the challenge smaller. This definition connects to the hurdles and partly to the *discouragements* from step 2b of the design method. The most important ones, according to OLX internal research (2014c) are featured in the skill atom.

### 3.2.4 Ideation & Iterative Prototyping

With all the knowledge gathered in the design steps 1, 2 and 3 above, the actual design of gamification treatments could start in step 4 of the gamification design method: ideation & prototyping. As explained in sub chapter 2.4.3, the ideation and iterative prototyping step is quite comprehensive. It consists of the following parts, which can be repeated as often as needed:

- a. Evaluate skill atom and brainstorm first ideas using design lenses
- b. Create prototypes of first ideas
- c. Workshop: evaluate prototypes and generate ideas
- d. Refine or replace prototypes

Appendix C elaborates on the actual chronological process of the ideation and prototyping that was conducted in this research. This includes the iterative cycles in terms of ideas, prototypes/mock-ups, workshops and comments, refinements, etc. Appendix C gives a full understanding of the complete development process and context of the final gamification treatment that was selected, implemented and evaluated. To give a general overview of these iterative cycles a diagram has been constructed, see Figure 10 below. Ideas are transformed into prototypes and go through various workshops and discussions, after which they are either dismissed or refined. The names of the parts in Figure 10 correspond with the chapter titles of appendix C. When placing the parts in the context of the parts *a* through *d* as mentioned above, they correspond as follows: *a* = 1; *b* = 2; *c* = 3 & 5; *d* = 4, 6 & 7.

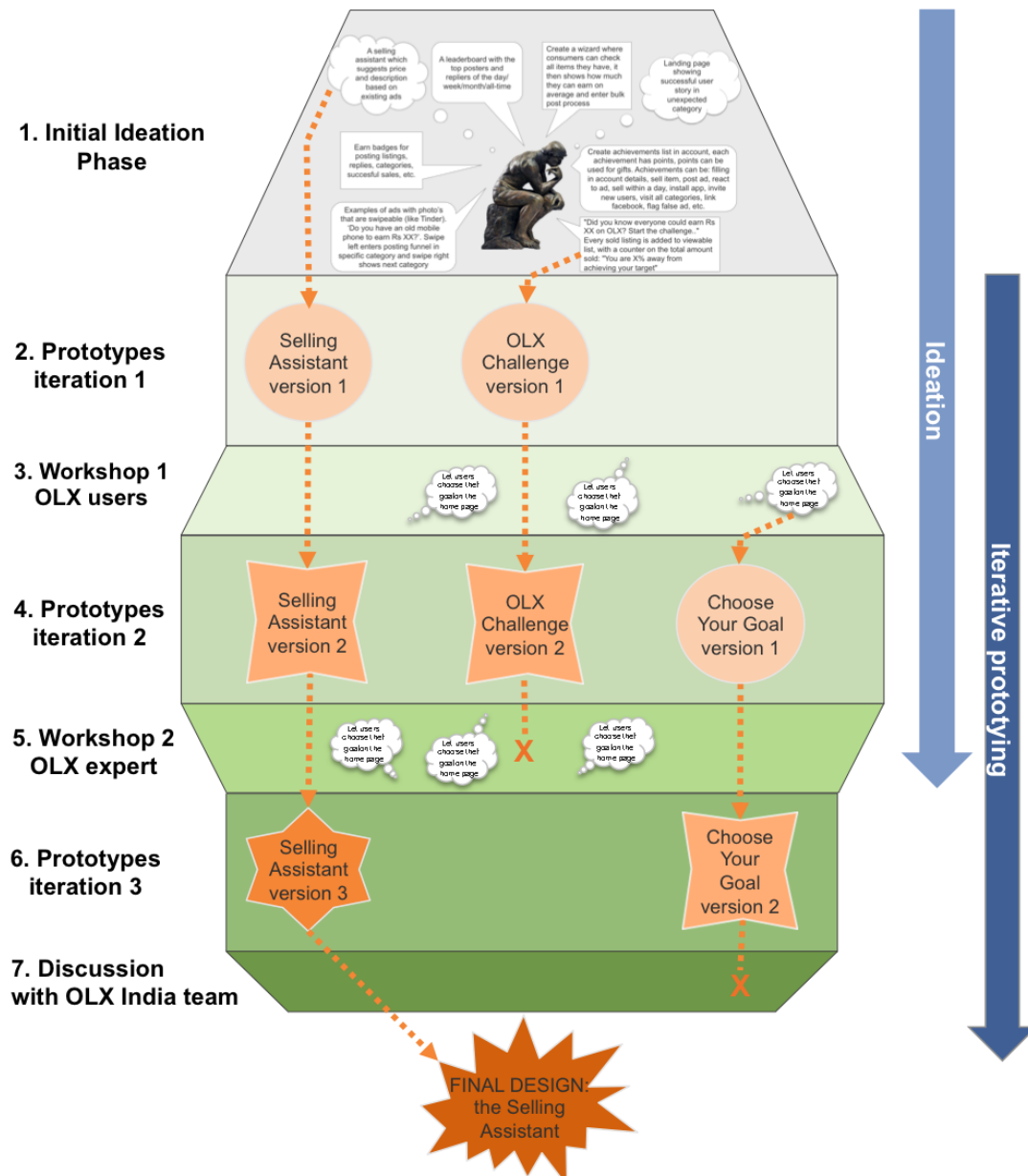


Figure 10: Ideation and iterative prototyping, step 4 of the gamification design method

The ideas in the icons in the figure above are not necessarily readable, but please note that the ideas from the initial ideation phase are also featured in Figure 22 in appendix C.1. Of these ideas, two were turned into a presentable mock-up or prototype. These were evaluated and refined through two workshops, in which respectively five OLX users (Indian students at Delft University of Technology) and one OLX expert (Business Improvement Leader at Naspers Classifieds) were involved. Based on an idea in the first workshop, a third prototype was created. See the photos in Figure 11 below and appendix C.3 and C.5 for an extensive description of the two workshops.



Figure 11: Workshop with OLX expert (bottom right photo) and workshop with OLX users (other photos)

The goal was to have a shortlist of two or three treatment mock-ups/prototypes after the workshop iteration cycles, which could be discussed with the OLX team. Two treatments passed through all the workshops, matched the constraints and requirements and were included in this shortlist: 'Choose Your Goal' and 'the Selling Assistant'. The shortlist was presented to the OLX India team and refined again, based on the comments. The treatment that was selected, implemented and evaluated is 'the Selling Assistant'.

### 3.3 Final Gamification Treatment: The Selling Assistant

Having introduced the case, explained the gamification design method and described the process of completing the design steps, this sub chapter elaborates on the final gamification treatment and thereby answers sub question 4 of this study. This concept has gone through the most number of iterations, because it was developed during the initial ideation step and already evaluated as a prototype in the first workshop with Indian students. Through the two workshops and through discussions with the OLX India team, the treatment design was refined time and time again. The responsible team at OLX India selected the 'Selling Assistant' treatment to be implemented and tested, because they thought it had the highest chance to affect the number of listings. Also, in both workshops, the idea of this treatment was generally regarded as the one with the most potential.

#### 3.3.1 Concept of 'the Selling Assistant'

A way to increase the number of posted listings is to get more users to enter the posting funnel. One of the biggest challenges for users, which prevents them from starting the posting process, is that they do not know what to sell; or think they don't have anything to sell because of the low resale value

of an item. The idea is that appealing explicitly to this challenge with a call to action (for instance: 'I don't know what to sell') will engage the users who cope with the described challenge. This was confirmed in the user workshop. Breaking the challenge into smaller steps by providing guidance and a limited number of choices for items to sell originates from the game designs lenses used. The next best action is to select which product to sell from a small list, rather than 'sell something'.

Information is needed to convince users of the views and reactions they will get from other users with their listing. The Selling Assistant shows users categories with a low number of listings (thus containing products that are currently not obvious for users to be sold on OLX), but a relatively high number of page views (thus containing items that relatively many people want to have). This can be translated into a number of 'average views per listing', which indicates the need for a new listing. By showing this number to users, they might be convinced that also these products are eligible for sale on OLX. On the long term, this mechanism can be used by OLX to drive the number of listings specific categories, which allow them to effectively and dynamically match supply and demand. Also, by showing only categories with items that a large share of users has laying around unused in their house, the chance of stimulating users to post such an item is higher.

12 categories were included into two different versions of the Selling Assistant, based on two criteria. The first is that the category should account for at least 1% of the total number of listings on OLX. This way, an increase in the number of new listings in a category would actually have a significant effect on the platform in general. If only very small categories are featured, the total potential of the treatment in terms of extra new listings per user is still relatively low. The categories that fulfilled this first criterion were ordered based on their number of page views per listing. The 12 highest were chosen.

### 3.3.2 Workflow & Design

The Selling Assistant consists of three pages, which consecutively link to each other, in the following order:

1. the home page;
2. the product page;
3. the posting page.

The home page is the original mobile home page, but a button is added that leads to the product page. The product page features 6 categories and the average number of views per listing (ad) in those categories. Selecting a category leads to the posting page. If the category selected is specific enough, it is prefilled in the form in the posting page. Apart from this, the posting page is exactly the same as in the original mobile website. See Figure 12 below for screenshots of the three pages.

From literature on online randomised controlled experiments, it is known that a change in colour or wording for a button could have large effects in terms of user behaviour (Deng, Li, & Guo, 2014; Hynninen & Kauppinen, 2014; R Kohavi, Deng, Longbotham, & Xu, 2014). So for the home page, two version with different buttons were created, in order to detect the difference and to thoroughly check the effect of the Selling Assistant:

1. Suggest: the button on the home page features the text: 'Suggest me what to sell';
2. Know: the button on the home page features the text: 'I don't know what to sell'.

In this case, it would be good for the validation of the treatment if both of these versions yield similar or at least comparable results, because then the effect of the general gamification design is greater than the effect of the change in layout.

Also, users were supposed to get the idea that the product page and the data on views per listing were generated dynamically with each page load, while in reality, the product page is a static page with the same content loaded every time. Therefore, if users click the Selling Assistant button on the



home page, they are randomly assigned to version A or B of the product page, implying that the product page is dynamic (because users see different versions if they visit the page multiple times). The product page has two versions, each with 6 different categories that are featured (a split of on the 12 categories that were chosen in total, see previous sub chapter 3.3.1):

1. Version A: features the categories Sofa's, Home & Kitchen Appliances, Mobile Phones, TV - Video – Audio, Decor & Furnishings, Motorcycles;
2. Version B: features the categories Pets, Cars, Video Games & Consoles, Laptops, Fridge - AC -Washing Machine, Mobile Accessories.

To give a general idea, the Suggest home page, Version A of the product page and the posting page (with category Sofa's from the product page selected and prefilled) are displayed below. Complete overviews of the workflows and designs of the original mobile website and the Selling Assistant can be seen in appendix D.2 in Figure 35.

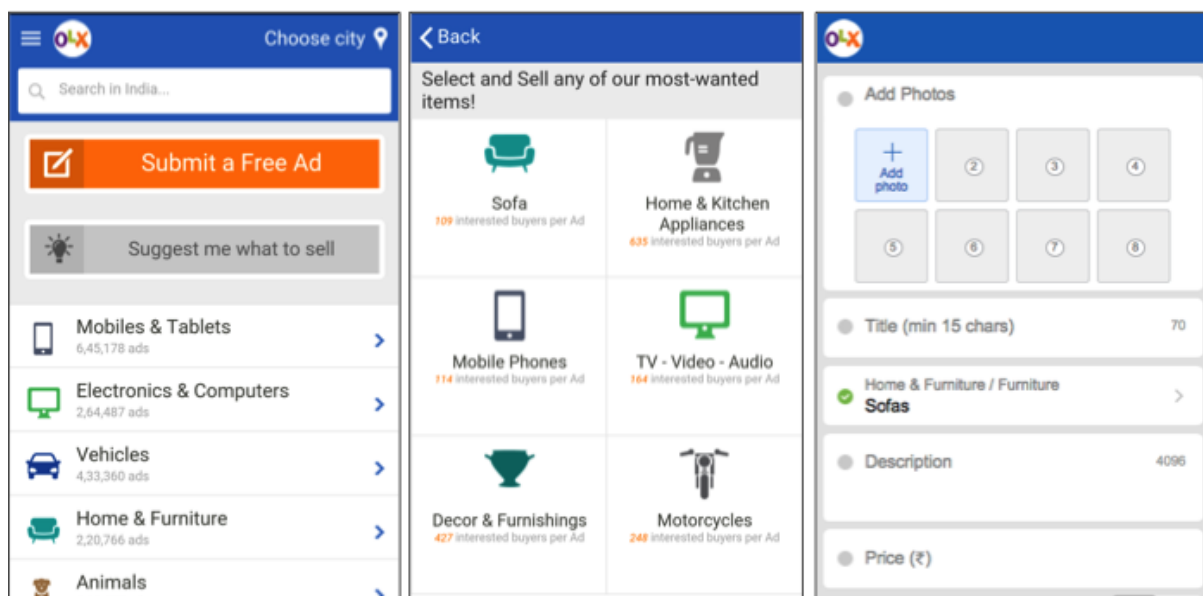


Figure 12: Final design of gamification treatment 'The Selling Assistant', with home page (suggest version), product page (version A) and posting page, from left to right.

### 3.3.3 Game Design Lenses Used

During the workshops held (see appendix C), most of the game design lenses from the design method were used (Deterding, 2014c), either plenary or by individual participants during their ideation process. The following game design lenses (see appendix D.1 for details) are still identifiable in the final design of the Selling Assistant and have been applied in some stage of its development:

- Scaffolded complexity: the concept of posting a listing is broken down to a less complex task, which is also the first task to perform when posting a listing – selecting an item to sell. So instead of entering the posting funnel, the user is guided into the first logical step of posting a listing: selecting an item to sell.
- Appeal to motivations: a motivation for users to sell their items on OLX is to earn money and have interactions with others. By showing users the demand for listing in different categories, they can indirectly derive whether they will get many replies and a good price for their item if they post it.
- Limited choice: Normally users who want start posting have to select one from 11 main categories, then from 105 sub categories and then 198 sub-sub categories. This means three choices with respectively 11, 18 and 11 options to choose from on average. Limiting the choice to only 6 categories and using sub categories instead of main categories might make

the choice less difficult. Moreover, the main categories are created by OLX and may not be self-evident for each user, while a sub category certainly is recognisable.

- Templates: providing a constrained set of products as starting point, including demand indications, might partly take away the fear or inability for users to start from scratch with posting a listing. Also, the sub category that is selected in the Selling Assistant will be pre-filled into the form on the posting form when they continue.
- Onboarding: the button on the homepage, which leads to the 'Selling Assistant', says 'I don't know what to sell' or 'Suggest me what to sell'. This might create a strong want in the user to start, because not knowing which item to sell is one of the biggest challenges for OLX users.
- Interim goals: this is closely related to 'limited choice'. The main challenge of not knowing what to sell is broken down in to multiple small steps, structuring the path to posting a listing. First, a user makes his/her need explicit, then her/she chooses a category and then the details of the listing need to be filled in (but with the choice for a category already explicitly made).

### 3.3.4 Final Requirements & Constraints Check

As a final check, the Selling Assistant examined on the requirements, constraints and general OLX qualities from chapter 3.2.1 (step 1c of the design method).

#### Requirements

- ✓ Fast in load times (so as few images, videos, etc. as possible) → no images, video's or big files are included, so the load time of the home page or the product page of the Selling Assistant is not compromised in comparison to the original mobile website.
- ✓ Free to use → there are no costs associated with using the Selling Assistant.
- ✓ Fit seamlessly into the current OLX mobile website → the Selling Assistant was redesigned by the OLX user experience team in order to comply with the OXL website design.

#### Constraints

- ± Implementable with HTML5 and Javascript within a few days (this is based on time and resources within the research) → the implementation took more time than expected, mainly due to the automatic category prefill option in the posting page, when a category is selected through the Selling Assistant.

#### General OLX qualities

- ✓ Simplicity of usage → the Selling Assistant is quite straightforward and its simplicity was optimised by the OLX team and during the iterations between the workshops.
- ✓ Free → there are no costs associated with using the Selling Assistant.
- ✓ Quality, original content → not influenced by the Selling Assistant.
- ✓ Speed of transaction → not influenced by the Selling Assistant.

## 4 EXPERIMENT WITH GAMIFICATION TREATMENT

---

*Chapter four will lay out the exact application of a randomised controlled experiment for the evaluation of the designed gamification treatment: the Selling Assistant. The hypotheses, experiment setup and data collection are described, as well as the steps performed to prepare and validate the collected data. Then, the data that was collected during experiment is analysed and the most important findings are depicted, lead by the hypotheses that were formed and the findings and remarks that naturally follow the results. The following sub research question will be provisionally answered:*

- 5. What is the effect of the gamification treatment on the amount of content generated by users?*
- 

### 4.1 Expected Effects of the Selling Assistant

Based on the description of the Selling Assistant in the previous chapter and the premises to choose this gamification treatment (e.g. the target outcomes as described in 3.2.1), a deeper dive can be taken into the actual effects that it is expected to generate. The hypotheses described here will serve as the alternative hypotheses for the statistical tests, where the associated null-hypotheses are always in the form that there is no difference or no effect cause by the Selling Assistant. According to the available gamification theory, there should be a positive effect of the Selling Assistant on the number of new listings posted by new users on OLX, as set out in the previous chapters of this report. Therefore, the hypotheses are one-sided.

This sub chapter contains only the shortlist of the most important and interesting hypotheses that were actually validated based on the experiment that was done to evaluate the Selling Assistant. Due to the resources available, the way the experiment was set up and the data collection possibilities, not all hypotheses on expected differences could be tested. The excluded hypotheses and the specific reasons why they were excluded can be seen in appendix F.

In the hypotheses described here, 'seeing' the Selling Assistant or original mobile OLX website means visiting one of these pages and 'interacting' means clicking one of the links on a page (and not exiting to another website or closing the browser). A user is someone who visits OLX within a certain defined period (in this case, during the experiment). A new listing is a gross new listing, thus a listing that has been posted by a user, but has not necessarily passed OLX approval regulations. A lister is a user who posts 1 or more listings.

#### 4.1.1 Main Hypotheses

The Selling Assistant was designed to stimulate the number of new listings that each user posts. It is assumed to do this both directly, through the product page with suggestions, and indirectly, by stimulating users to interact more with the website and giving them inspiration for items to post in general. The general share of users that will post something is expected to be different, because users are inspired and encouraged by the Selling Assistant to sell something on OLX, whereas they might otherwise have not thought of anything to sell at all. Moreover, the average number of listings that a lister posts is expected to be different, because of similar reasons. The Selling Assistant was designed to lower the efforts it takes to post a listing. Furthermore, the indications of the number of views each category receives (on the product page) might be even more appealing to experienced

users than new users, which could positively influence the number of listings per lister. These expectations are translated into hypotheses below.

Hypothesis 1 relates to the main metric and target of the case and experiment. This is the main hypothesis to support, since it indicates the overall effect and whether gamification has the capability to increase the amount of UGC on an online marketplace.

*Hypothesis 1. Users who have seen the Selling Assistant will post more new listings on average than users who have seen the original mobile OLX website.*

Hypothesis 2 concerns the actual effect of the Selling Assistant in terms of the influence it has on users who have interacted with it, rather than all users who have seen the button on the home page.

*Hypothesis 2. Users who have clicked the Selling Assistant button (and seen the product page) will post more new listings on average than users who have not clicked the Selling Assistant button.*

Hypotheses 3 and 4 are used to further explore where the possible difference in the overall number of new listings comes from. They can be used to pinpoint the effect of the Selling Assistant.

*Hypothesis 3. The proportion of users that is a lister will be higher for users who have seen the Selling Assistant than for users who have not seen the Selling Assistant.*

*Hypothesis 4. The number of listings per lister will be higher for users who have seen the Selling Assistant than for users who have not seen the Selling Assistant.*

Hypotheses 1 and 2 are tested with the experiment results in chapter 4.4.1; hypotheses 3 and 4 are tested in chapter 4.4.2

### 4.1.2 User Segments to Explore

The effect of the Selling Assistant on the number of listings per active user will probably not be the same for all the OLX users, but stronger in groups with specific factors and weaker in groups with other factors. It might provide very valuable insights to look into various user segments for which the Selling Assistant was or was not effective (Ron Kohavi et al., 2008). However, not all expected user segment differences that stem from the workshops could be tested. These are featured in appendix F. Some are still interesting enough to mention here, as they are explored in the results in sub chapter 4.4.

#### Home and Product Page Variants

As mentioned in sub chapter 3.3.2, multiple versions of the Selling Assistant's home and product pages were implemented in the experiment, because the OLX team expected they would yield different results. Naturally, it is interesting to see if the different home and product page variants yield different results. For the home page variants (Suggest and Know), the only causally expected difference could be in the clicks on the home page button, since the difference between Suggest and Know is in fact only the text in the button. Other than this, the variants are the same, so there should be no significant differences in number of listings only, because these cannot be rationally explained. This difference is best measured in the proportion of users that clicks the button, thus is in some way interested, engaged and/or motivated by the text. For the product page variants (Version A and B), a difference in number of new listings posted can be expected, because each variant has its own 6 categories that are featured and displayed to users. Some categories might be more appealing for users than others.

## New and Returning Visits

OLX has identified various user segmentations that have proven to distinguish users with different behaviour. An OLX report (2014c) shows that significant differences in behaviour can be attributed to new and returning users, in all sorts of contexts. This can also be defended on a logical basis, for example: a new home page feature focused on explaining the working of the OLX platform in more detail might result in more listings posted by new users, but will only be an extra nuisance with likely useless information for experienced (returning) users. The workshops that were held during this research (see appendix C.3 & C.5) also confirmed the expected differences between new and returning users.

In the experiment, the following two segments are distinguished:

1. New user (has not visited the OLX site before during a pre-defined period, e.g. during the experiment);
2. Repeat (has visited the OLX site before during a pre-defined period, e.g. during the experiment).

Note that the classifications of new and returning for users are limited to the experiment time, due to the unavailability of historical user data. In other words: a new user in the Original variant might be someone who has been using OLX for over a year and has sold and bought many items through the platform. He or she is classified as new in the experiment during the first visit and as returning for the second visit and possible visits afterwards. So for the Original variant, the new and returning classifications are not really different from one another. For the Know and Suggest variants, the classifications do have a meaning, because no user has seen either of these variants before, so novelty effects could be tested (R Kohavi et al., 2014).

The expected differences between new and returning user behaviour - originated from the workshops and OLX internal research - are based on the overall characteristic of a user being new or returning and are not limited to the experiment time. Also, this variable could not have a user as experimental unit, but a page visit (or page load), because users are assigned as new and returning during multiple visits. The expected differences could therefore not be translated directly to testable hypotheses. However, exploration of the differences between new and returning visits in the experiment will still be done later on in this chapter. Furthermore, see appendix F for the excluded hypotheses on the specific new and returning segment differences.

The differences between the users segments of Selling Assistant variants and new and returning visits are explored with the experiment results in respectively chapters 4.4.3 and 4.4.4.

## 4.2 Experiment Setup

In this sub chapter, the setup of the experiment is described, which was used to evaluate the effect of the Selling Assistant on the number of listings that were posted.

### 4.2.1 Double-Blind Randomised Controlled Experiment

A double-blind randomised controlled experiment (or: randomised controlled trial) was used. A fixed percentage of people that visited the OLX mobile homepage were randomly assigned to one of three variants: 'Original', 'Suggest' or 'Know' (see sub chapter 3.3.2). Original is the control variant, whereas Suggest and Know are both treatment variants. The number of users in each variant is roughly the same and the variants differ only in one factor: the website version that is shown to users (with or without the Selling Assistant). The website variant that each user experiences is the same during the entire experiment, even for multiple visits on different days. The experiment is double blind, because for the nature of the research it is necessary that both researcher and participants are unaware of who is participating in the experiment. The users do not know if they are included into the

experiment nor to which variant they have been assigned. Double blind is chosen to prevent researcher and user bias (for users, knowing that they are in the experiment could alter their behaviour) and to ensure user privacy (no personal data is used). The experimental unit over which calculations are made and for which the random sampling is applied is a unique OLX user who visits the OLX mobile website.

#### 4.2.2 Timeline, Population, Sampling & Variants

The experiment was live for exactly 7 days, from Monday February 23<sup>d</sup>, 16:00h (CET) until Monday March 2<sup>nd</sup>, 16:00h (CET). This way, each day of the week was represented. During this period, each OLX mobile website user was checked for two conditions:

1. Entry on the OLX mobile home page (e.g. the first page on the OLX website they saw has the URL <http://olx.in/i2/> or <http://www.olx.in/i2/>);
2. Located in India (based on stored browser cookies and IP address).

If one of the conditions was not met, the user was excluded from the experiment. If both conditions were met, the random sampling method was applied. For every user who met the conditions, a fixed percentage of users was excluded. This percentage was chosen by the OLX team, based on the number of users that they were willing to expose to the Selling Assistant for this research. The rest was included and assigned to one of the three variants:

- 33% to the 'Original' variant, which is the same as the normal mobile website;
- 33% to the 'Suggest' variant, which includes the Suggest version of the Selling Assistant home page (see sub chapter 3.3.2);
- 33% to the 'Know' variant, which includes the Know version of the Selling Assistant home page (see sub chapter 3.3.2);

To provide clarity: an OLX mobile website user is someone who uses a mobile phone internet browser to go the OLX website. Based on the mobile phone browser, he or she is automatically redirected to the mobile website (rather than the desktop website). Another possibility is that it is someone who is not necessarily browsing on a mobile phone, but specifically chose to visit the mobile version of the OLX site by going to a specific mobile URL or selecting the mobile version at the bottom of an OLX web page.

The users in the Suggest variant who clicked the Selling Assistant ('Suggest me what to sell') button on the home page and the users in the Know variant who clicked the Selling Assistant ('I don't know what to sell') button on the home page were grouped in a virtual bucket. From this virtual bucket users were randomly selected and assigned to one of the two Selling Assistant product pages:

- 50% to the 'Version A' product page;
- 50% to the 'Version B' product page.

Two variants of the Selling Assistant product page were created in order to suggest to users that the product page had dynamic content while it was actually static (see sub chapter 3.3.2). However, every user sees the same product page during the experiment, not an alternation between the two, undoing the reasoning above. This does allow for an exploration of the difference in posted new listings between the two product page variants (version A and version B).

Figure 35 in appendix D.2 shows a detailed workflow overview of the experiment and its variants, including screenshots of the web pages.

Optimizely (D Siroker, Koomen, Kim, & Siroker, 2014; Dan Siroker & Koomen, 2013) is a third-party web-based platform that allows relatively easy setup of online double-blind randomised controlled experiments (or A/B tests, in their jargon). It was used to conduct the experiment, because OLX uses this platform as a standard. Optimizely is one of the most used platforms for A/B tests, even though some limitations are known (Borden, 2014; Optimizely, 2015). See sub chapter 6.2.2 in the reflection



for more details on this. Naturally, since Optimizely was used, their random sampling method and user allocation scheme (see Figure 36 in appendix E.1) were applied in the experiment.

### 4.2.3 Participant Privacy

No personal user data was used in this research or will be made public through this research. During the experiment, some personal data was collected in the form of IP addresses of OLX users that were included in the experiment. Optimizely uses this, in combination with cookies and browser behaviour, to determine if a user is a unique user and assigns a unique and random user identification number. Only the user identification number was used in this research, IP addresses were removed from the dataset.

### 4.2.4 Data Collection

The data that was collected is partly standard user behaviour that is always logged by Optimizely and partly data that was specifically set to be measured, based on the hypotheses that were defined in sub chapter 4.1.

Optimizely logs each page load (or page view) of each user in the experiment variants. For each of these page loads, the following variables were collected by default:

- Timestamp: the date and time at which the page was loaded
- User identification number: unique random number for each user (see 4.2.3).
- Description: the URL of the page that was loaded. Some page load URLs are recoded to specific names, if this was set as a manual rule in the experiment setup. For instance, in this experiment, if the button leading to the Selling Assistant was clicked, the URL description of that page load was recoded to 'click\_on\_suggest\_button'.
- Browser: internet browser used to access the website (e.g. Safari, Google Chrome, etc.).
- Source: how a user reached the website (a referral link from another site, a search engine, direct visit or an online marketing campaign).

Also, these variables were defined specifically for this experiment:

- Home page variant: to which home page variant the user was assigned (Original, Suggest or Know).
- Product page variant: to which product page variant the user was assigned (Version A or B). Naturally, this is only registered for users who actually visited the Selling Assistant product page, by clicking on the Selling Assistant button in the Know or Suggest variant home page.
- Location: from which location the user was accessing the website. New Delhi, Mumbai, Bangalore, Chandigarh, Chennai, Coimbatore, Jaipur and Pune were logged. If the location was detectable but none of these 8, 'other' was logged. If location was not detectable, 'unknown' was logged.
- New or Returning visitor: if the user was visiting the OLX website for the first time within the experiment, he/she was logged as new. If the browser or web page was closed and the user returned on the OLX mobile website during the experiment time, he/she was logged as returning.

Javascript code, on the details of how this specific data was collected, can be seen in appendix E.2.

### 4.3 Data Preparation, Validation & Statistical Methods

In this sub chapter, the steps that were performed to validate, merge and filter the data are described. After these steps the data was ready to be analysed. Also, the methods that were used to do this, including statistical tests are described. For the data validation and analysis the programs R Studio (version 0.98), IBM SPSS statistics 22, Microsoft Excel 2011 and G\*Power (Faul, Erdfelder, Buchner, & Lang, 2009) were used. R packages 'pscl', 'vcd', 'ggplot2', 'gplot', 'psych', 'data.table' were implemented. R Studio and all packages used are available from CRAN at <http://CRAN.R-project.org/>. The same software was used for sub chapter 4.4.

#### 4.3.1 Data Preparation

Before validation and further analysis was possible, the raw datasets needed to be prepared. This was done in the three steps shown in Figure 13 below, which are elaborately described in appendix G, including the R scripts used.

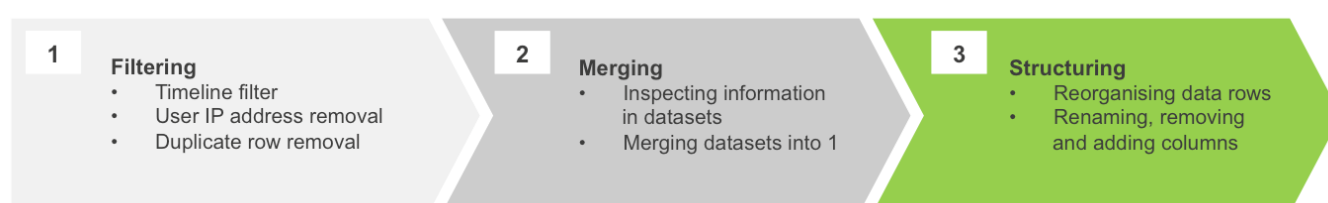


Figure 13: Data preparation steps performed

The most vital things that needed to happen were the removal of all the duplicate data logs, the merging of the different raw datasets into one and the transformation of the data in such a way that each case is not a page load, but a user. This was needed because a user is the actual experimental unit on which the hypotheses are based. Because the values for the variables were collected per page load, some users had different values assigned to them for different page loads. For example, some users visited the site from both Mumbai and New Delhi or with both Chrome and Safari as browser during consecutive visits. For these users, the most occurring values were assigned to them. More on this explained in the next section. Also, based on the raw data, some extra variables for each user could be derived, such as the number of listings posted (derived from the URL description) and the number of page views (derived from the number of page load data logs per user). Appendix G.5 features data samples of the dataset organised per user and per page load.

#### 4.3.2 Data Exploration & Validation

Before testing actual hypotheses, the data was explored and validated, to check if it corresponds with the experiment as implemented. The validation steps undertaken are described in detail in appendix H. The most important findings per step are summarised here.

##### 1. User Segment Overlaps and Missing Values

In total, 51103 users were included in the experiment. 16744 of them were in the Original variant, 17156 were in the Suggest variant and 17203 were in the Know variant. 765 users clicked through to the product page, of which 350 were assigned to Version A and 349 were assigned to Version B. The remaining 66 users did not continue after seeing the product page (did not post a listing) and therefore Optimizely did not have the chance to log their product page allocation in a later page load. The exact user behaviour and the difference between the Suggest and Know variants in terms of product page views is discussed in the results chapter. No users were included in multiple variants, corresponding with the experiment setup.

For other segmenting variables, multiple values were assigned to each user, in different page view logs. For each of these variables, the most occurring value for each user was allocated to each user (in step 3 of the data preparation as described above). 1.3% of the users could not be allocated to a browser, 49% of the users could not be allocated to a location and all users were allocated to a source. For the 'new or returning' variable, no single value could be allocated to each individual user. The reason for this is that many users were defined both as new and later as returning user, during the experiment, based on their consecutive visits. Therefore, this variable could not be included in the general analysis as a predicting variable for the number of new listings that were posted. However, when widening the definition of the experimental unit to visits rather than users, the visits can be assigned as new or returning. Based on this, some analysis is done later on in chapter 4.4.4.

## 2. Extreme & Duplicate Values

For the extreme value check, the page views per user variable showed no large outliers. The new listings variable did: the maximum is 6 listings, posted by one user. However, in the Know variant, there is a user who did not interact with the product page, but posted 29 listings, of which 2 were duplicate listings. Because of this discovery, duplicate values for posted listings needed to be deleted manually (duplicate values in the data preparation were removed automatically, see appendix G.1.3, but apparently not thorough). This manual deletion process is described in detail in the appendix H.2. The 27 remaining outlier listings were genuine and correct, but the extreme number could heavily (and possibly incorrectly) influence the results. The outlier listings are taken into account in some analyses, but in general, they were changed to 6, which corresponds to the other maximum number of listings. This way, the high number is still taken into account, but with a less severe impact.

## 3. Fit of Control Group to Population

The fit of the results (in terms of mean listings per user) of the users in the Original variant to the population (all the mobile web users during the same period) was not very accurate. The population average was 58% higher. However, the comparison cannot be made one-to-one, because the users in the experiment were selected based on the criterion that they landed directly on the homepage, while the users in the population are all the users in general. Also, the population data provided by OLX comes from two different databases and its accuracy cannot be guaranteed. Therefore, there is no definitive reason to doubt the controlled randomised experiment, nor the validity of the relative comparison between the Original, Suggest and Know variant, which is the main subject of this research. This is described in more detail in appendix H.3.

## **4.3.3 Statistical Testing Methods**

### New-Listings-Per-User Data

For hypotheses 1, 2, and 4 'new-listings-per-user data' was used. Meaning: the variable of interest (dependent variable) is the number of new listings each user posted during the experiment. Each case is a user, thus there are 51103 cases. This data can be generalised as 'count data', where the minimum value is 0 and all value are integers. The distribution of the new-listings-per-user data as result of the experiment was explored in appendix I.1. There, it was concluded that normality, or a normality approximation (following the central limit theorem), was not appropriate. This is partly due to the fact that the experiment setup closely follows a Poisson process and the mean of this theoretical Poisson distribution is not high enough to assume normality (Nussbaum, Elsadat, & Khago, 2008).

However, a Poisson distribution does not exactly fit to the new-listings-per-user data, because of over-dispersion (the variance is greater than the mean) and because of the fact that many zeros occur (many users did not post a listing at all). This problem is common for count data and a negative binomial model can best be used to solve this, because it corrects for over-dispersion and is a less restrictive model for the same type of data (Cameron & Trivedi, 1998; Gardner, Mulvey, & Shaw, 1995; Nussbaum et al., 2008; Zeileis, Kleiber, & Jackman, 2008). See appendix I.1 for more

information on this topic. Negative binomial models used to test hypothesis one were cross-referenced with Poisson regression. The differences in regressor coefficients were marginal, but the negative binomial models were improvements of the Poisson models in terms of residual deviance and log-likelihood. Moreover, SPSS frequently showed that the validity of the Poisson models fit was uncertain.

Also, as a first check, Mann-Whitney U tests (or Wilcoxon rank-sum test) are used. This is the nonparametric alternative to the independent-samples t-test. It is mainly useful for data that is non-normal with independent treatment and control groups, with a categorical independent variable that influences an ordinal or ratio dependent variable. This test cannot control for other variables, but can give a general indication as to where it is expected that there is a significant difference between two groups in the distributions of a variable. If so, a regression model is needed to confirm this. If not, doing a regression analysis is only needed if the actual predictor coefficients are of interest. This is not really the case, since the hypotheses are aimed at detecting general effects of the Selling Assistant, rather than quantitative values for its influence. Also, the Mann-Whitney U test is used in a very similar gamification study on an online marketplace (Hamari, 2013).

### Comparing Two Proportions

For some parts of the analysis, two proportion values need to be compared, in order to check if they differ significantly. A few examples of proportional values to be compared between Original, Suggest and Know:

- Proportion of Selling Assistant home page visits with an SA button click;
- Proportion of home page visits with an interaction (bounce rate);
- Proportion of unique users with a posted listing ( $\geq 1$ ).

These are two sample comparisons, similar to T-tests, but with proportions rather than means. For this, the Z-test for proportions of Independent samples can be used and normality can be assumed according to the central limit theorem. More information on this test and its applicability can be found in appendix I.2.

## 4.4 Results

This sub chapter will provide a walkthrough of the most relevant experiment results, guided by (but not limited to) the hypotheses and the expected segment differences as set out in sub chapter 4.1. For each hypothesis the relevant experiment results are shown, followed by an appropriate test for statistical significance in order to verify the hypothesis in question. For all hypothesis tests, a significance level of 0.05 was used. Appendix I features extensive information on the statistical distributions of the data used to test the hypotheses and the corresponding statistical tests that were used, based on these distributions. To support the understanding of the results as explained in this chapter, apart from the figures featured in the chapter, it is also useful to take a look at Figure 35 in appendix D.2, where the full experiment workflow of the Selling Assistant design is featured.

### 4.4.1 Effect on the Number of New Listings per User

What is the effect of the Selling Assistant (gamification treatment) on the number of new listings (user-generated content) that the OLX users generate? The general effect of the Selling Assistant presence is evaluated with hypothesis 1, while the effect of the Selling Assistant on the users who actually interacted with it is evaluated with hypothesis 2. These are the most important hypotheses for this research; therefore the results of the statistical tests are displayed in more detail.

## Hypothesis 1: Effect of Selling Assistant Presence in General

### General Results

Table 1 below shows the number of users, new listings and average number of new listings per user for the Original, Suggest and Know variants. Also, the Know variant with the outlier listings is shown in grey (see appendix H.2 for more info on this). It indicates the relative difference of Suggest and Know, compared to Original. As one can see, the number of new listings is quite small, compared to the number of users. Therefore, the absolute differences in new listings per user look negligible, while the relative differences are substantial. This could indicate, at first sight, that both the Suggest and Know variants were successful in stimulating users to generate more listings. More descriptive data can be found in appendix I.1.1.

**Table 1: Experiment results of Original, Suggest and Know in terms of average new listings per user**

	Users		New Listings		New Listings per User	
<b>Original</b>	16,744		143		0.0085	
<b>Suggest</b>	17,156	+ 2.5%	169	+ 18.2%	0.0099	+ 15.3%
<b>Know</b>	17,203	+ 2.7%	170	+ 18.9%	0.0099	+ 15.7%
<b>Know (with outlier)</b>	17,203	+ 2.7%	191	+ 33.6%	0.0111	+ 30.0%

### Two-Sample Comparison

To verify if the overall difference is statistically significant and check hypothesis 1, unpaired two-sided Mann-Whitney U tests were conducted. For Know, the outlier was kept in, because it has no influence on the results of the test (it compares and ranks each number of new listings per user in two variants, thus 6 and 27 posted listings are both considered as highest value). Because the shapes of the distributions are roughly the same (see Figure 42 in appendix I.1.1), the Mann-Whitney U test can be used to see if there is a significant difference in the number of listings per user of two groups. The results of the tests can be seen in Table 2 below.

**Table 2: Results of Mann-Whitney U tests to verify hypothesis 1**

Test	W-value	P-value	Total Comparisons	Share Original Higher
Original versus Suggest	143,479,788	0.278	287,260,064	49.948%
Original versus Know	143,859,574	0.239	288,047,032	49.943%

The W-value is the number of times that the number of new listings of a user in Original was ranked higher than the number of new listings of a user in Suggest or Know. The total number of comparisons is the product of the number of users in the experiment variants. The share value indicates the share of the W-value in the total number of comparisons done in the test. Looking at these shares, there clearly is no large difference between the number of listings per user in Suggest and Know versus Original. The P-values ( $> 0.05$ ) confirm this. The main reason for this is that there are still many users who did not post something at all. All these users are ranked equally in the test, making the majority of the comparisons equal. So on the total number of users in the experiment, the number of extra listings as result of the Selling Assistant is very small.

### Regression Models

To further investigate hypothesis 1, negative binomial regression models were fitted to the data. The data without the outlying user was used, because the outlier is so extreme. Please note that the coefficients (B-values) are log-transformed, therefore also the exponentiated coefficients are shown. These indicate the relative increase of the number of new listings compared to the intercept. So an

exp(B) value higher than 1.0 means an increase in the dependent variable (new listings) and a value lower than 1.0 is a decrease. The full SPSS output of the regression models can be seen in appendix J.1.

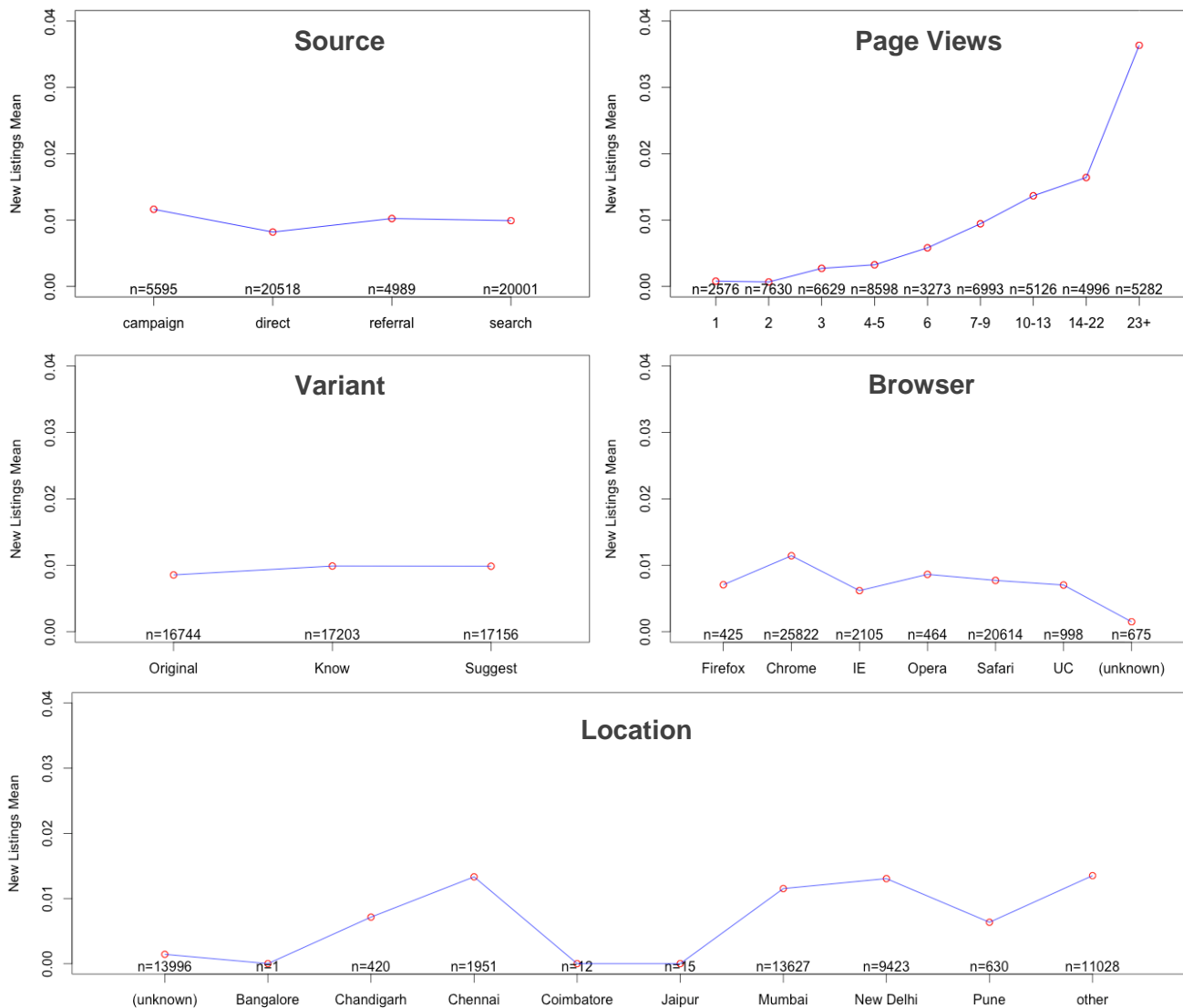
First of all, the effect of the three different variants (Original, Suggest, Know) on the number of new listings was tested. With the Original as intercept ( $B = -4.763$ ,  $\exp(B) = 0.009$ ,  $SE = 0.084$ ,  $p = 0.000$ ), both Suggest ( $B = 0.143$ ,  $\exp(B) = 1.154$ ,  $SE = 0.114$ ,  $p = 0.211$ ) and Know ( $B = 0.146$ ,  $\exp(B) = 1.157$ ,  $SE = 0.114$ ,  $p = 0.201$ ) were no significant predictors for the variance in the number of new listings. Moreover, the overall model performed very badly (likelihood ratio Chi-Square = 2.12,  $df = 2$ ,  $p = 0.347$ ). Therefore, more predictor variables needed to be added.

From the experiment data, the following predictor variables could be included:

- *Browser*: the browser is not expected to have a causal effect, but could explain variance, as the browser also indicates what type of phone users have, which can say something about the technology adoption of a user and his/her income.
- *Source*: the source of a user indicates whether he/she reached OLX directly, through a search, by clicking on an external referral link to OLX or by clicking on an external OLX advertisement. Again, no causal effect is expected, but it could explain variance.
- *Location*: an expectation that originated from the user workshop (see appendix C.3) was that users from small towns would be more susceptible to the Selling Assistant than users from large cities. Not enough data from small town users was collected to test this (see appendix F.2.2), but 8 different cities could be allocated to users.
- *Number of Page Views*: there is an expected causal effect here, as users who are more active on OLX generally post more listings (OLX, 2014c). Also, the interaction effect between the variants and the number of page views is included. This is because using the Selling Assistant inherently implies visiting pages and the mean of the number of page views per user is slightly higher for Suggest and Know (11.31; 11.32) compared to Original (11.13). The page views could possibly be a mediator between the effect of the home page variant and the number of posted listings. This was tested in a small ordinary OLS regression model with standardised variable values (to slightly correct for non-normality), using the PROCESS macro for SPSS (Hayes, 2013). The indirect mediation effect was not significant, so not further taken into account. See appendix J.1.4 for the output of this test.

The role of the predictor variables in explaining extra variance in the number of new listings is supported by the plots in Figure 14 below, which show the difference in means of new listings per user, for each value or grouped values (in case of page views) of predictor variables. Rather than showing the actual data points the mean gives more information, given the relatively small differences and the discrete nature of the new listings per user data. The y-axes all have the same scale, indicating that the experiment variant explains a very small amount of the listing variance.





**Figure 14: Visual inspection of the effect of predictor variables on the mean listings per user**

The second model, with the extra predictor variables included, performed much better (likelihood ratio Chi-Square = 312.91, df = 23,  $p = 0.000$ ). The residual deviance decreased with 16% and the log-likelihood increased, compared to the first model. Original serves as intercept for the experiment variant predictor and for the other categorical predictors random values were chosen because their specific coefficients are not really of interest. Now, controlling for the extra variables, Suggest ( $B = 0.055$ ,  $\exp(B) = 1.056$ ,  $SE = 0.121$ ,  $p = 0.650$ ) and Know ( $B = 0.037$ ,  $\exp(B) = 1.037$ ,  $SE = 0.1209$ ,  $p = 0.762$ ) are even less influential predictors for the number of new listings. Controlling for the other predictor variables and compared to Original, it is 95% certain that the number of new listings differed between -17% and +34% for users in Suggest and between -18% and +32% for users in Know.

Although improved, the second model is not correctly specified (deviance/df = 0.077). This is logical, because many variables that could explain posting behaviour of OLX users are not included in the experiment and it is not possible to include all of them and their interaction effects (not even theoretically, given irrational behaviour of people). However, to correct for the possible misfit, the scale was adjusted by the deviance in a third model. This does not change the coefficients, but does change the standard errors. Even in this third model, with artificially lowered standard errors and thus

generally smaller confidence intervals for parameter coefficients, the experiment variant was not a significant predictor variable.

As can be seen in Figure 14, the average number of listings per user varies greatly per location. More specifically, this is a difference between tier 1 (extremely large) and tier 2 (large) cities. But because only 1,107 users were logged as tier 2 city users, of whom only 7 posted a listing, the difference is not significant in the regression model.

### Conclusion

There is no statistically significant difference between the new listings per user in the Original, Suggest or Know variants. Therefore, hypothesis 1 is not supported. It cannot be concluded that the general presence of the Selling Assistant had a significant effect on the number of listings that users posted. However, because the relative difference in the average number of listings is quite big for both Suggest and Know (>15%, see Table 1), it is worth looking into only the users who actually posted a listing. If all the users who did not post are left out, the results might change. However, this is a different type of question, which will be explored later on in this chapter. First, hypothesis 2 will be verified, which is essentially a more specific version of hypothesis 1.

### Hypothesis 2: Effect of Interacting with the Selling Assistant

The main expected influence of the Selling Assistant does not come from the presence of the button on the home page, but from the product page with the category suggestions. Users who clicked the button on the home page leading to the product page are the only users who have experienced the full Selling Assistant. Examining the difference between their behaviour and the behaviour of users who did not click the button seems like a sound method to estimate the true effect of the Selling Assistant. Naturally, this effect can only be measured within the Suggest and Know variants.

### General Results

Within the Suggest and Know variants, respectively 2.1% and 2.4% of the users clicked the Selling Assistant button on the home page. The Know button was a bit more popular and is featured in Figure 15 on the right. The users who did click the Selling Assistant button, ended up on the product page with 6 suggested categories. The effect of this page can be explored by looking at the average number of listings per user and segmenting users in those who did visit the product page and those who did not visit the product page. This is displayed in Figure 16 below. Please note that the user who posted the outlier listings in the Know variant did not visit the product page.

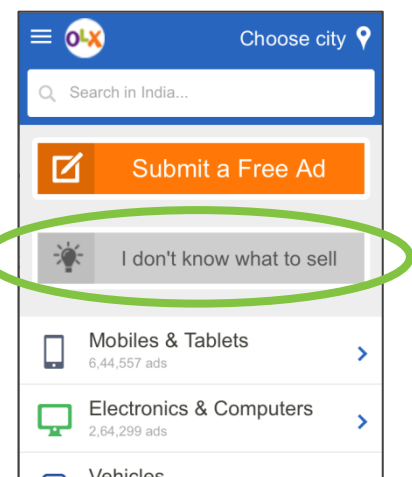
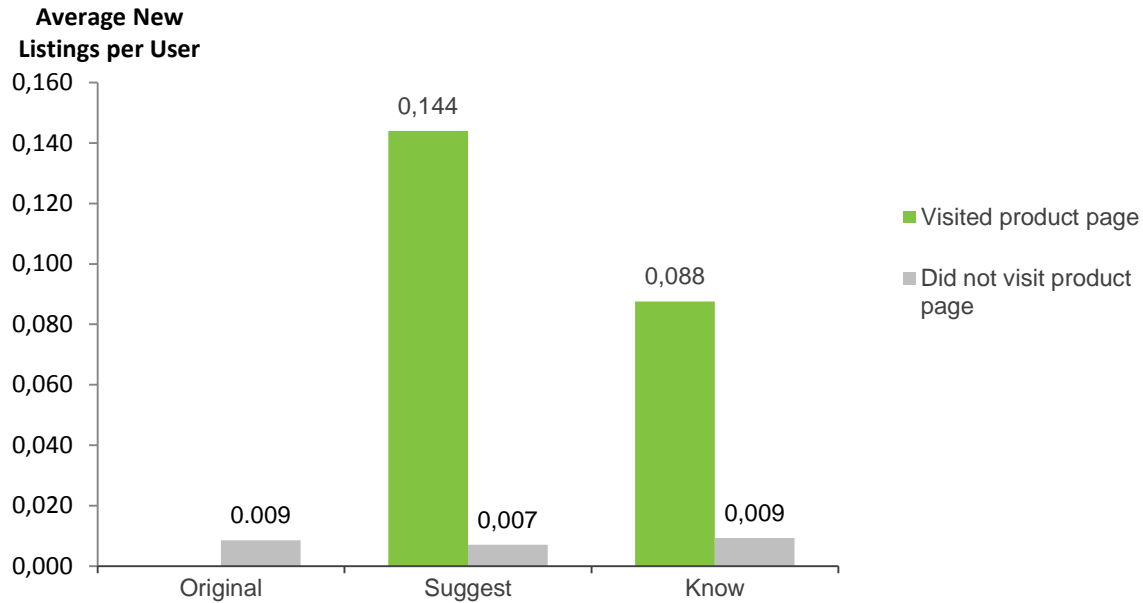


Figure 15: Home Page Know Variant



**Figure 16: The number of new listings per user per home page variant, segmented in users who did and did not visit the product page**

The differences are vast. The percentage of users who visited the product page in both Know and Suggest was quite small in percentage (2.4% and 2.1%). In absolute numbers this were 354 and 411 users for Suggest and Know respectively. These users posted 51 and 36 listings of the total 169 and 191 in Suggest and Know respectively. The average number of listings per user for users who visited the product page is also higher than in the population (users not in the experiment) during the same period. The fact that the users who saw the product page posted much more listings than users who did not, could confirm hypothesis 2.

#### Two-Sample Comparison

In order to verify the results from above, two Mann-Whitney U tests were conducted. In these tests, the users with and without a product page view within Suggest and within Know were compared. Theoretically, the users with and without a product page visit could also be compared combined over the Suggest and Know variants. But in this case, it is wise to segment as much as possible, to control for independent variables. The results of the Mann-Whitney U tests are displayed in Table 3 below.

**Table 3: Results of Mann-Whitney U tests to verify hypothesis 2**

Test	W-value	P-value	Total Comparisons	Share Product Page Higher
Within Suggest, visited product page versus not visited product page	3,257,398	0.000	5,947,908	54.8%
Within Know, visited product page versus not visited product page	3,653,106	0.000	6,901,512	52.9%

The number of comparisons for Suggest is the product of the 354 who have seen the product page and the 16802 users who have not seen the product page. For Know, these are 411 and 16792 users respectively. The results of the tests clearly indicate that there is a statistically significant difference in the number of new listings per user for users who have and have not visited the product page. However, in these tests, there was no controlling for the other predictor variables, so more regression models need to be made.

### Regression Models

Negative binomial models were constructed for both subsets of data from Suggest and Know users. Of course, in these models the experiment variant predictor variable was not included. Instead, the binary product page variable (which is 'yes' or 'no') was included as main predictor, as well as its interaction effect with the page views (for the same reason as in hypothesis 1). In order for the negative binomial model to converge, location values Jaipur and Coimbatore were taken out (because of their small  $n$ ), as well as users with an unknown browser. For both Suggest and Know, the negative binomial models had a better fit than identical Poisson models. No click to product page was set as intercept value and again random values were selected as intercept for the other categorical predictors (since their coefficients are not of interest). The SPSS output of the models can be seen in appendix J.2.

For Suggest users, the regression model had a good fit (likelihood ratio Chi-Square = 304.36,  $df = 17$ ,  $p = 0.000$ ) and showed that clicking to the product page is the main predictor variable for the variance in the number of new listings. Users who clicked to the product page significantly posted more listings than users who did not ( $B = 2.599$ ,  $\exp(B) = 13.445$ ,  $SE = 0.205$ ,  $p = 0.000$ ). Controlling for their browser, location, source and number of page views, users in the Suggest variant who clicked on the button leading to the product page of the Selling Assistant posted between 9 and 20 times as many new listings (with 95% confidence) as users who did not click the button leading to the product page. The same model, but with the scale controlled for the deviance (to correct for the misfit, in the same way as with hypothesis 1), resulted in an even smaller confidence interval, because of corrected standard errors.

For Know users, the same model was applied and had a significant fit (likelihood ratio Chi-Square = 192.89,  $df = 17$ ,  $p = 0.000$ ). Also here, the product page was the most influential predictor variable for the variance in the new listings and users who visited it significantly posted more listings than users who did not ( $B = 2.269$ ,  $\exp(B) = 9.673$ ,  $SE = 0.218$ ,  $p = 0.000$ ). Controlling for their browser, location, source and number of page views, users in the Know variant who clicked on the button leading to the product page of the Selling Assistant posted between 6 and 15 times as many new listings (with 95% confidence) as users who did not click the button leading to the product page. Again, controlling the scale for deviance gave a smaller standard error, thus smaller confidence interval (approximating towards the  $\exp(B)$  value of 9.7).

An ordinary OLS regression model with the PROCESS macro for SPSS (Hayes, 2013) and standardised variables was used, in order to confirm the effect recognised above. Clicking to the product page (yes or no) was tested as moderator for the effect of being in the Suggest or Know variant (yes or no, independent variable) on the number of posted listings (dependent variable), controlling for the number of page views (co-variate). The moderator was a highly significant predictor. See appendix J.2.3 for more on this.

### Conclusion

Hypothesis 2 can be supported with 95% confidence: users who visited the product page of the Selling Assistant posted a higher number of new listings than users who did not.

#### 4.4.2 Diving into the Actual Effect of the Selling Assistant

It is yet unclear on what specific type of users the Selling Assistant had an effect. Did it convince more users to post a listing in the first place, or did it mainly stimulate users who were going to post a listing anyway and thereby increased the number of listings per lister?

##### Hypothesis 3: the Proportion of Users that Posted a Listing

It is suspected that the proportion (or share) of users that posted a listing in the Suggest and Know variants was higher than in the Original variant. If this is true, it provides more insight into the actual effect of the Selling Assistant.

##### General Results

Figure 17 on the right displays the proportion of users that posted 1 or more listings during the experiment (thus is a lister), including 95% confidence intervals. The outlier listings in Know are included here, since the number of listings per user is not relevant. As can be seen from the figure, there is a difference. The Suggest and Know variants both have a higher proportion of listers.

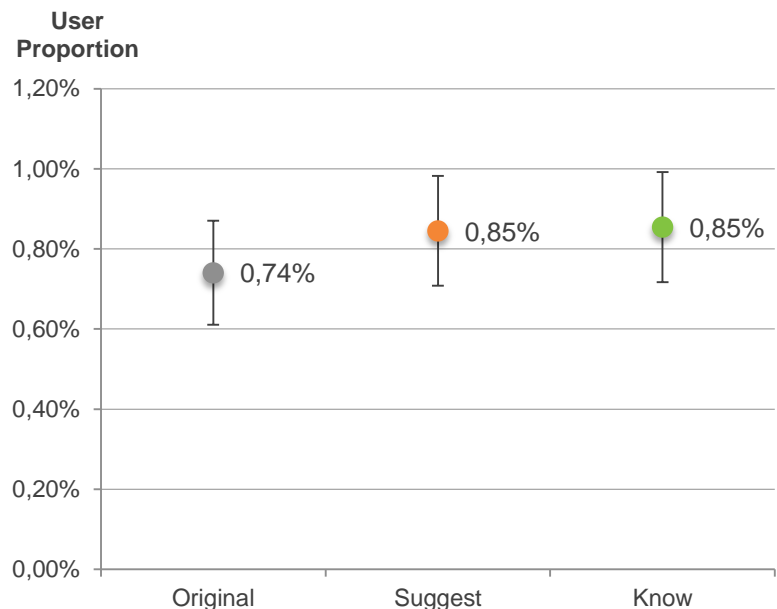


Figure 17: The proportion of users that is a lister, per experiment variant

##### Two-Sample Comparison

To verify if hypothesis 3 can be supported, a Z-test for proportions of independent samples was used. Table 4 below shows the results. The P-values ( $>0.05$ ) indicate that a statistically significant difference between the proportion of users that is a lister in Original and Suggest or Know variant cannot be supported with 95% confidence.

Table 4: Results of Z-tests for proportions to verify hypothesis 3

Test	Relative Difference in Lister Proportion	Test Z-Score	P-Value
Original versus Suggest	14.1%	1.085	0.28
Original versus Know	15.4%	1.179	0.24

##### Conclusion

Hypothesis 3 is not supported. A logistic regression model could be applied with as dependent variable a new variable 'post?' (binary, values 1 and 0) which indicates if users posted a listing, including the predictor variables from regression models above. However, controlling for more variables will only decrease the variance in the proportion of users that is a lister that is explained by the experiment variant. So the likelihood of the variant being a significant predictor in the proportion of listers will be lower. In the Z-test, it was already non-significant, so there is no point here in creating a logistic regression model.

### Hypothesis 4: the Number of Listings per Lister

Seeing as there is no significant difference in the proportion of users who was a lister, the extra number of listings in Suggest and Know might be caused by the fact that the listers in these variants posted more listings on average.

#### General Results

A first check is the difference in the mean listings per user, split for each of the predictor variables that were used for hypothesis 1. In this case, all experiment data was used, but the users who did not post a listing were filtered out. Figure 18 below shows their mean number of listing for each variable, including the variant they were in. Now, confidence interval bars ( $p = 0.95$ ) were added, in order to provide more information on the variance (and because no regression analysis was performed, so the visual inspection is leading). Again, all y-axes have the same scales. As can be seen, the experiment variant has the least influence on the variance in the average number of listings per lister, so it is not likely that it is a significant predictor.

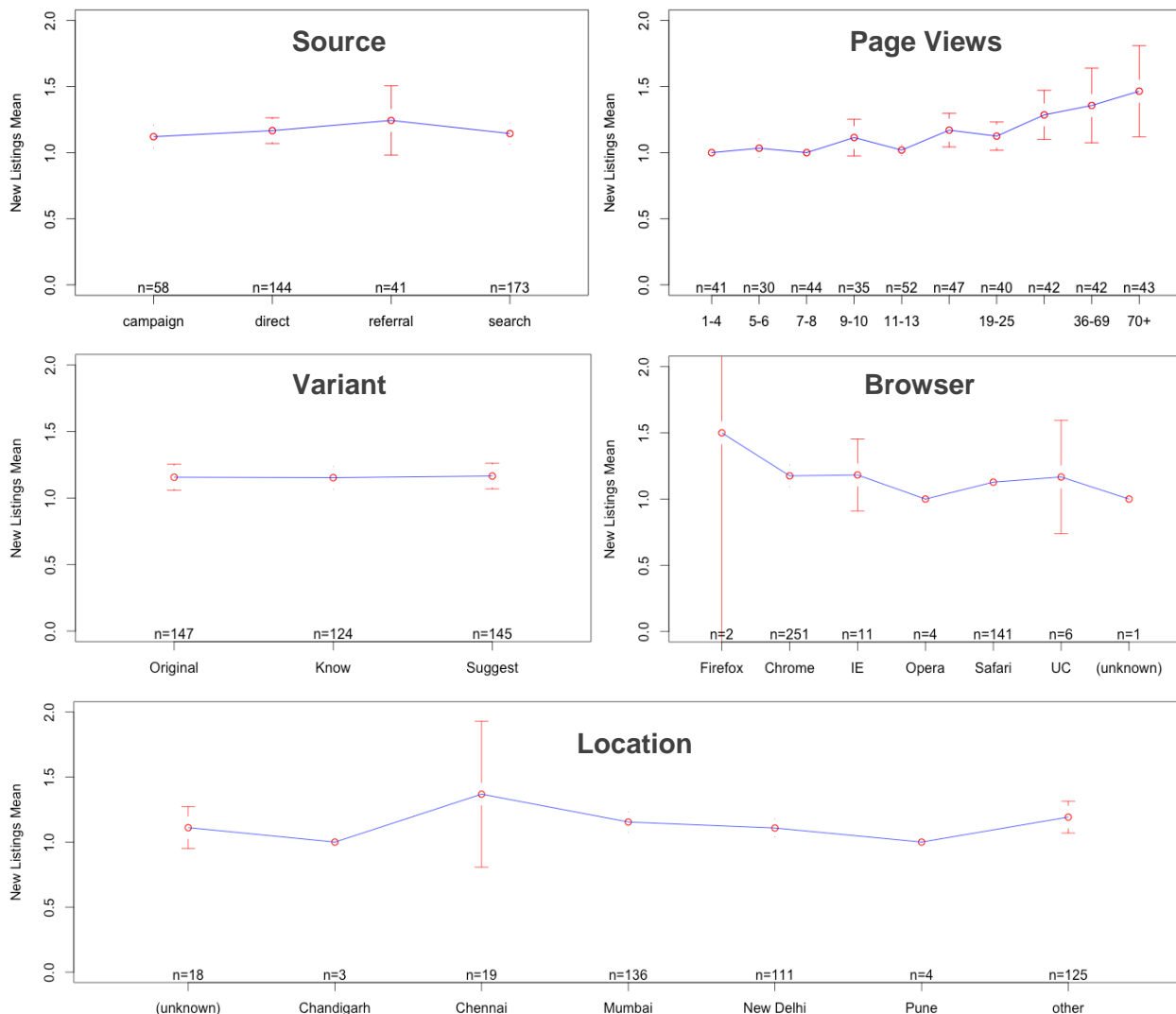


Figure 18: Visual inspection of the effect of predictor variables on the mean number of listings per lister

#### Two-Sample Comparison

In order to check if there is a difference between the number of listings per lister, comparing Original with Suggest and Know, two independent sample Mann-Whitney U tests were conducted, of which



the results are displayed in Table 5 below. The p-values indicate there is no significant difference, even without controlling for other variables, which confirm the suggestions from the graphs in Figure 18.

**Table 5: Results of Mann-Whitney U tests to verify hypothesis 4**

Test	W-value	P-value	Total Comparisons	Share Original Higher
Listers in Original versus listers in Suggest	8965.5	0.945	17,980	49.86%
Listers in Original versus listers in Know	8957.5	0.646	18,228	49.14%

### Conclusions

Hypothesis 4 cannot be supported. There is no significant difference in the average number of listings that listers posted, when comparing listers who have seen the Selling Assistant (were in Suggest or Know variants) and users who have not seen the Selling Assistant (were in the Original variant).

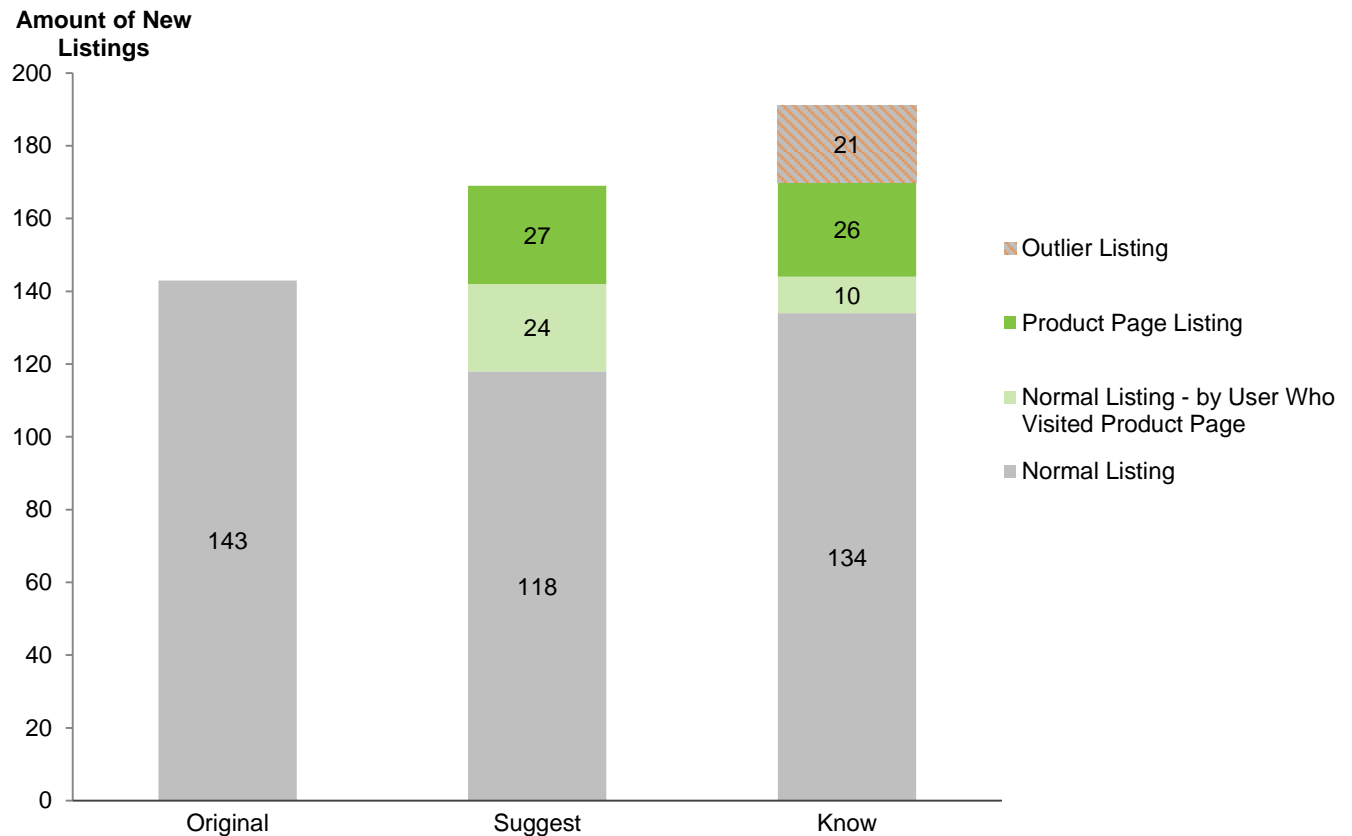
### 4.4.3 Exploring Selling Assistant Variants

As indicated in chapter 4.1, apart from the specific hypotheses, there are some interesting user segments to look into, which might answer some of the remaining questions regarding the effects of the Selling Assistant on OLX users. What type of listings did the users who visited the product page post? Which version of the product page yielded more listings? How come that the users who visited the product page in Suggest posted more listings on average than users who visited the product page in Know (recall Figure 16)? Is it due to the different text in the home page button (Suggest versus Know) or due to the different categories that were recommended (Version A versus Version B)? These questions will be explored here.

#### Suggest Versus Know

In the Original variant, users posted only one type of listings, namely 'normal' listings, which are posted through the normal posting page, which is reached by pressing on of the many 'sell' buttons throughout the OLX mobile website, such as the orange sell button on the home page (see Figure 15). In the Suggest and Know variants, many normal listings were also posted. However, next to normal listings, the users who clicked through to the Selling Assistant product page posted listings directly through this product page: by selecting a category on it, proceeding to the posting page and filling in the listing details. These listings will be called product page listings.

The number of posted listings per variant (Original, Suggest, Know) per type is displayed below in Figure 19. Naturally, Original only contains users who did not visit the product page and who posted normal listings. For Suggest and Know, it is interesting to see what the distribution of the types of posted listings is. Looking at Suggest, a substantial number of listings (51 out of 169, thus 30%) is posted by users who visited the product page, of which around half (24) is a normal listing and another half (27) is a product page listing. In the Know variant, the share of listings posted by users who visited the product page is lower than in the Suggest variant, namely 36 out of 191 listings, thus around 19%. Of these listings, the majority (26 out of 36) are product page listings. The Know variant contains the 21 outlying listings (see appendix H.2 on this), which would all fall into the category normal listings. Not taking these outlier listings into account and adding up all the normal listings (by both users who did and did not visit the product page), would result in an almost equal number of normal listings for the Original, Suggest and Know variants: 143, 142 and 144 respectively. Given this fact and looking the graph below, it seems that due to the Selling Assistant, there were 27 extra listings in the Suggest and 26 extra listings in the Know variant (which are all product page listings). Taking Original as reference point, the Suggest variant yielded 19% more listings during the experiment and the Know variant yielded 18% more listings during the experiment.

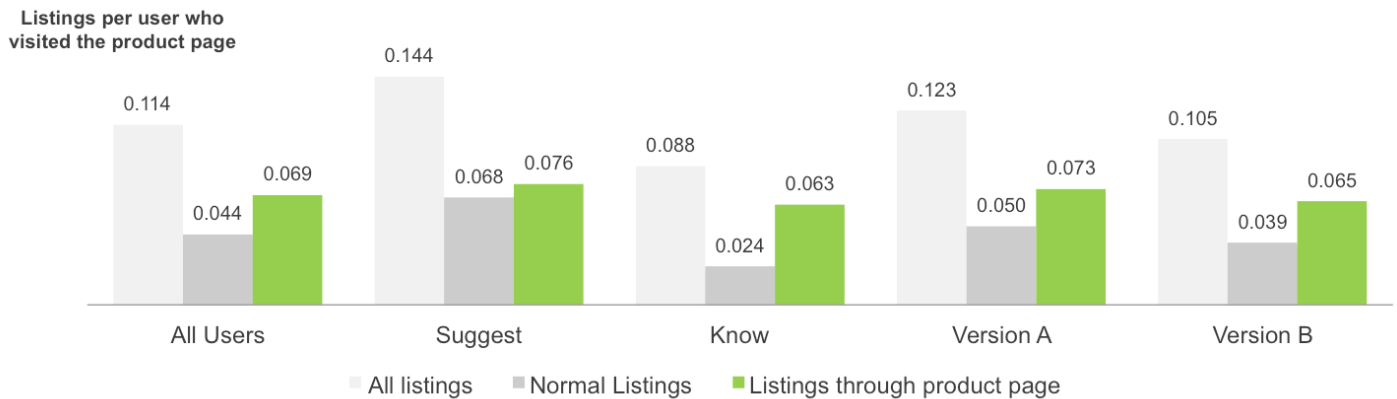


**Figure 19: The number of listings per home page variant, segmented in different listing types**

The similar results of Suggest and Know support the general validity of the experiment. After all, apart from the difference in the number of clicks on the home page button leading to the product page, there should be no large differences in the results between Suggest and Know. This is because the only difference in the user experience between these groups was the text in the button on the home page. The statistical tests that were conducted above, with all the experiment data included (thus the tests for hypotheses 1, 3 and 4) were also done for differences between Suggest and Know. There was never a significant difference.

### Segmenting for Listing Types

The users who clicked the Selling Assistant button on the home page of Suggest or Know were allocated to Version A or Version B equally. Of the 765 users who clicked the button and visited the product page, 350 were allocated to Version A, 349 were allocated to Version B and of the remaining 66 users it is unknown which product page variant they saw. These 66 users did not post a listing. Figure 20 below displays the number of listings per users for all the users who visited the product page. The users are shown in different groups, based on the home and product page variants: all users (who visited the product page), Suggest, Know, Version A and Version B. Please note that these groups overlap (a user can for example be in all users, Suggest and Version B). Also, listings are segmented into groups per type: all listings, normal listings and product page listings. To be able to evenly compare the number of listings per user, the 66 unknown product page users were split between Version A and Version B. So the calculations in the graph in Figure 20 were made based on 765 users in 'all users', 354 users in Suggest, 411 users in Know, 383 users in Version A and 382 users in Version B.



**Figure 20: Listings per user, only for users who visited the product page, split for user groups, segmented in different listing types**

Figure 20 indicates a few things. The first is that users who saw the product page in the Suggest variant posted more listings than users in the Know variant. In this graph, it can clearly be seen that this difference is especially applicable for normal listings (0.068 versus 0.024 listings per user). The difference between Suggest and Know also holds for product page listings but is less apparent (0.076 versus 0.063 listings per use). Seemingly, the users in the Suggest variant were stimulated to post more listings, but the real reason cannot be extracted, because there is no difference in the OLX website version these users saw apart from the text in the Selling Assistant button. Also, there are no expected or logically defensible interaction effects between the home page and product page variants. Therefore, the difference between Suggest and Know is likely caused by chance, or more specifically: the product page versions that users were allocated to.

#### Version A Versus Version B

When looking at the differences between product page variants, Figure 20 indicates that Version A outperforms Version B, for all types of listings. This difference can probably be allocated to the different categories that were featured in A and B. Some categories were more appealing to users than others. Version A probably featured the categories that actually activated users to post something. However, this is just an assumption, because the difference could well be due to coincidence. After all, the number of new listings per user from Figure 20 above is relative. The absolute difference in the number of new listings between Version A and Version B is only 7. Several simple statistical tests indicated there was indeed no statistically significant difference.

#### 4.4.4 Exploring New & Returning Visits

Given the definitions for new and returning users as adapted in the experiment, they only mean something for users who see the Suggest and Know home page for the first time. For each user in the experiment, the home page is the first page they see. After this, if they close their browser, they are already classified as returning. Also, this variable does not have a user as experimental unit, but a page visit (or page load).

#### Differences in Home Page Button Clicks

A valid question that can be explored with this variable is: do users click on the Selling Assistant button on the home page during the first time they see it, or during a later visit? And is there a difference for Suggest and Know? See Table 6 below for an overview of this. It seems that there is about an equal number of users who click the button the first time they see it as number of users who click it in a later visit, for both the Suggest and Know variants. There are no significant differences, so no suggestions can be made regarding as to whether the text in the button appeals to users instantly or only after they have seen it at least once before.

**Table 6: Differences between new and returning visits of users who clicked the home page button, split for Suggest and Know**

Unique users who clicked the home page button			
	New	Returning	Total
<b>Suggest</b>	187	176	354
<b>Know</b>	204	209	411
<b>Total</b>	<b>391</b>	<b>385</b>	<b>765</b>

### Differences in Posted Listings

A relevant question could be: do users post something through the product page the first time they visit it, or later on? This could indicate the user-friendliness and completeness of the sequence of screens users go through. The general differences in posting behaviour between new and returning visits of users, per experiment variant, are visualised in Figure 21 below. The outlier listings are not included. It shows the means and confidence intervals based on the standard errors. The differences do not fall outside of the confidence intervals, so are not suspected to be significant, let alone when using a regression model to control for other variables. However, a notable thing is that the overall difference between the Suggest and Know variant and the Original variant largely originates from returning visits. They seem to have a higher average posting rate than returning visits in the Original variant, whereas the new visits in the different variants show no variance in amount of listings at all. But, no context can be given to this observation, because of the meagre time span for which new and returning visitors have been defined (only during the experiment, not as an OLX user in general). The only notion that can be made is that if the extra listings in Suggest and Know are indeed listings posted through the product page (as suggested in the previous chapter based on Figure 19), users do not post through the product page the first time they visit it. They might be inspired by a highly demanded category shown, close their browser, look around in their house for such an item and return to the product page later on to actually post the listing. This phenomenon has been witnessed by OLX amongst their users before and is also seen in Figure 19: new visits of users have a lower number posted listings than returning visits, through all three variants. Thus it seems to be a general mechanism that occurs, but is more present amongst users in the Suggest and Know variants than in the Original variant.

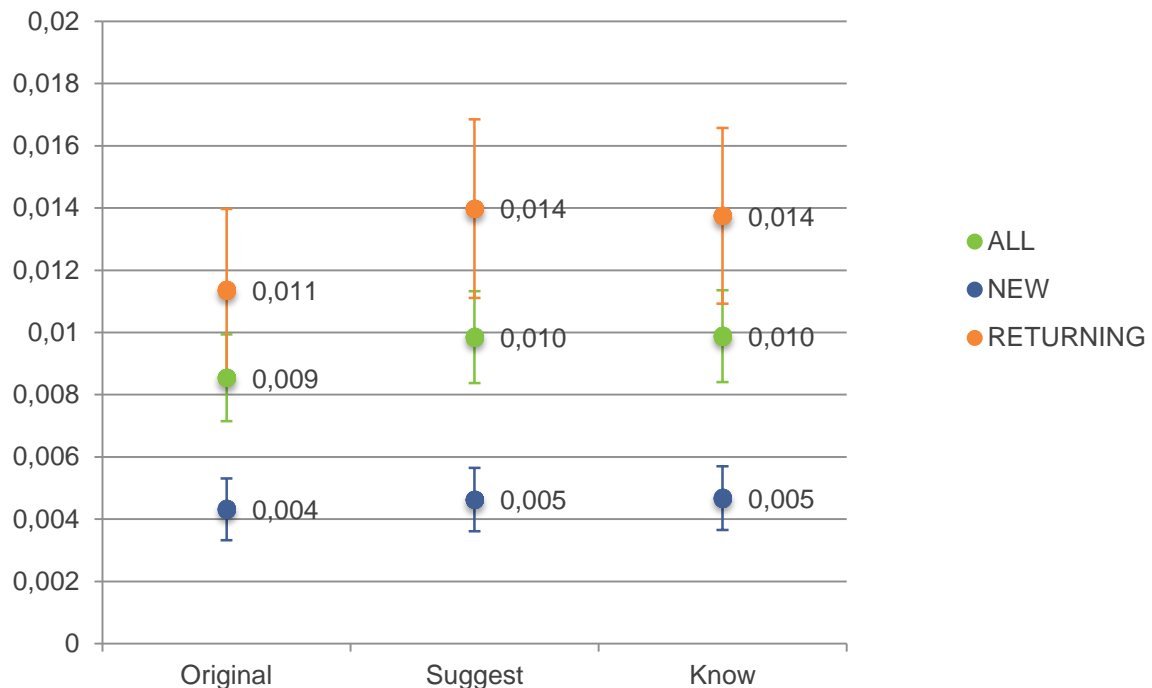
Average Amount  
of New Listings  
per User

Figure 21: 95% confidence intervals for the average number of new listings per user, split per home page variant, segmented for new and returning visits

#### 4.4.5 Concluding on the Experiment Results

With the hypotheses tested and the interesting segments explored, sub question 5 of this research can be answered: what is the effect of the Selling Assistant on the amount of content that is generated by OLX users? Table 7 below features an overview of the hypotheses that were tested and whether they were supported or not.

Table 7: Overview of tested and (un)supported hypotheses

#	Hypothesis	Supported?
1	Users who have seen the Selling Assistant will post more new listings on average than users who have seen the original mobile OLX website.	No
2	Users who have clicked the Selling Assistant button (and seen the product page) will post more new listings on average than users who have not clicked the Selling Assistant button.	Yes
3	The proportion of users that is a lister will be higher for users who have seen the Selling Assistant than for users who have not seen the Selling Assistant.	No
4	The number of listings per lister will be higher for users who have seen the Selling Assistant than for users who have not seen the Selling Assistant.	No

The Selling Assistant caused an increase in the amount of user-generated content. Suggest and Know yielded respectively 18% and 19% more new listings than the control group in the Original variant. These percentages correspond with the amount of listings that were posted through the product page of the Selling Assistant, suggesting that the treatment actually helped users in selecting an item to sell. However, this overall amount of extra listings is not statistically significant, thus it cannot be definitively concluded that the Selling Assistant had a positive effect on the overall amount of user-generated content on the OLX mobile website. The relatively big differences between the

Selling Assistant groups (Know and Suggest variants) and the control group (the Original variant) are not significant because only 2.1% and 2.4% of the users clicked the Selling Assistant button on the home page, in the Suggest and Know variants respectively. The extra productivity of users who engaged with the Selling Assistant still contributes only in a small degree to the overall number of listings, especially considering the fact that the vast majority of users did not post a listing at all.

Within both the Suggest and Know variant, users who actively engaged with the Selling Assistant by clicking on the button on the home page (thus users who experience the full design, including the product page) posted more than 6 times as many listings on average during the experiment than users who did not engage with the Selling Assistant, controlling for other predictor variables such as location, browser, number of page views and source. Thus the Selling Assistant resulted in a significant increase in new listings, for the users who actively engaged with the Selling Assistant themselves.

The question as to whether these interacting users within Suggest and Know were stimulated by the Selling Assistant to post more listings or start posting a listing in general (become a lister) could not be further filled in. There are no significant differences in the number of listings per lister or the proportion of users that is a lister, between the Suggest, Know and Original variants. Exploring the different experiment variants learned that there are no differences between Suggest and Know (which was expected) and Version A slightly outperformed Version B in terms of new listings per user, but not significantly. Looking at new and returning visits of users during the experiment learns that users generally post a listing in a returning visit, not in their first visit. This phenomenon is present in the Original variant and even more recognisable in the Suggest and Know variants, which means that users probably were engaged by one of the suggested categories on the product page, searched their home for such an item and returned to the website later to post a listing.



## 5 DISCUSSION, LIMITATIONS & FUTURE RESEARCH

---

*The results and limitations of the research are discussed, based on the two ways in which this research set out to evaluate gamification: design method application and an evaluation of a gamification treatment in a case online marketplace. This was explained in the research approach in chapter 1.2.3. The discussion and limitations are divided into the more qualitative evaluation of gamification by means of application of the design method and the more quantitative evaluation by means of the controlled experiment with the treatment. Gamification as a field and its applicability to online marketplaces is evaluated by summarising literature and by generalisation of the design method and case results to gamification theory where possible. Also, apart from the theoretical implications, managerial implications will be laid out. The chapter concludes in recommendations for future research; for OLX, online marketplaces in general and gamification researchers. Chapter five gives a provisional answer to the last sub research question:*

*6. What is the suitability of the design method for the way of working of online marketplaces?*

---

### 5.1 Experiment with the Selling Assistant

The gamification design method is evaluated by analysis of the effect of its product: the Selling Assistant. This gamification treatment was applied to OLX as a case, so the quantitative evaluation of the effect of the treatment is limited to the OLX users.

#### 5.1.1 Discussion of Experiment Results

##### Discussion of Outcomes Hypotheses 1 & 2

There are some remarks to be made, regarding the verification of hypotheses 1 and 2. First of all, users who have seen the product page have clicked the Selling Assistant button on the home page. This makes them a special type of users. If clicking the button means that they are the users who actually struggle with the challenge of not knowing what to sell, the Selling Assistant can be called very successful, for those users posted more listings than users who did not experience the full Selling Assistant. This could indicate that the Selling Assistant actually solves their problem. However, if the users who clicked the button are users who are active in general, normally post a lot of listings and are merely curious for a new functionality, the Selling Assistant is not necessarily effective. After all: the Suggest and Know variants did not make a large difference in new listings per user overall (hypothesis 1 was not supported) and can never make an overall difference, if it only affects users who are already active. Then it cannot assist OLX in raising the number of new listings on the long term and help them solve their chicken-egg problem. In order to do this, more users need to be activated to become a lister. This cannot be investigated directly, since no ex ante measurements or ex post surveys of the specific users in the experiment were made.

Second, it seems there is a correlation between interacting with the Selling Assistant and posting listings. Hypothesis 2 proved that visiting the product page is a significant predictor for the number of listings posted. But is the product page also the cause of the product page listings? Did the gamification treatment actually solve the challenge for users who did not know what to sell and would not have posted something if their challenge was not resolved? More questions arise when looking at the normal listings that were posted by users who visited the product page (24 for Suggest and 10 for

Know, in the light green area in Figure 19). Would the users who visited the product page and posted a normal listing have done this, even if the product page did not exist? In other words: did the Selling Assistant product page also have an indirect effect on users? Did users visit the product page, become inspired and post a listing through the normal posting page later on? Or did they visit the product page, find it of no use and continue to post the listing they already wanted to post when visiting OLX in the first place? Unfortunately, these questions cannot be answered here. It is not possible to discern from the experiment data whether or not the product page users in Suggest and Know would have posted a normal listing anyway, if they were in the Original variant and did not experience the Selling Assistant. But, the fact that the extra number product page listings in Suggest and Know is almost the same does support the assumption that the Selling Assistant had an effect on the number of listings that users posted.

Third, one could say that percentage of users interacting with the Selling Assistant button on the home page is quite low. Especially if - according to the starting point of the design phase and the information from the OLX internal reports, interviews and user workshop - identifying a suitable item to sell an OLX is one of the main challenges for users. Such a starting point suggests an engagement of more users than just over 2%. The low click percentage is a feasible explanation for the fact that hypothesis 1 was not supported, because very few users within the Suggest and Know variants actually experience the full Selling Assistant including the product page. The low interaction percentage could have several different causes:

- Appealing directly to the challenge with a button does not engage users who are struggling with this challenge;
- The wording, colour, position or other aspects of the button are not engaging;
- There are not many users who struggle with identifying an item to sell, even though it is the main challenge for OLX users;
- There are other challenges that apply to more users.

#### Discussion of Outcomes Hypotheses 3 & 4

The question as to what type of users were stimulated in their posting behaviour by the Selling Assistant remains unanswered. Hypothesis 3 and 4 could have steered the expectations in two directions. One direction being the assumption that they are users who are generally active and click on new features, now stimulated by the Selling Assistant product page to post more listings. Another direction being the assumption that they are users who normally do not post a listing, but are stimulated by the Selling Assistant product page to become a lister. However, the hypotheses could both not be supported. Still, assuming that hypothesis 3 or 4 could have been supported, no definitive conclusions could have been made. After all, users from independent groups are still being compared based on only their behavioural data. In order to truly evaluate what type of users clicked on the Selling Assistant button and why, qualitative information on users' motivations and historical behaviour on OLX would be needed.

#### Relating the Results to Existing Literature

Using gamification on online marketplaces to stimulate content creation is effective, as long as people actively engage with it themselves. In other words: the first step still needs to be taken by the users. An overall significant improvement in the amount of user-generated content could not be demonstrated. This was concluded in chapter 4.4.5 based on the experiment results. An empirical study by Hamari, who implemented badges as game elements on an online marketplace, yielded very similar results. He stated in his final conclusions: "This study was able to confirm that users who had actively exposed themselves to badges in Sharetribe were also significantly more likely to actively use the service [...]. However, support for the claims that implementing gamified features would alone lead to significant overall increases in usage frequency, quality or social interaction in a utilitarian trading service could not be found" (Hamari, 2013, p. 243). Hamari reported his positive findings based on analysis of a sub-set of active users. His dataset had the same issue as the one in this research: the number of users (experimental units) is vast, but the share of users who actually carried

out activities of interest is quite low. He recommends future studies to use this same subset approach and states that gamification treatments on services with a large user base might be especially effective, since affecting a small proportion of the user base will have a larger effect (Hamari, 2013).

### 5.1.2 Limitations of Experiment

The evaluation method can be related to all empirical gamification studies and literature on online randomised controlled experiments. As stated in chapter 2, Hamari, Koivisto and Sarsa (2014) identified the most important methodological shortcomings in existing empirical gamification studies, which were to be avoided in future research – as stated in the research goal in chapter 1.2.1. Here, the accomplishment of that goal is discussed. First, the pitfalls will be treated one by one and then the most important ones are discussed more elaborately, along with other recognised limitations of the experiment as conducted.

#### Pitfalls from Previous Studies

1. *Too small sample sizes, with N=20;*  
This pitfall was definitely avoided in this research.
2. *If user experiences and attitudes were surveyed ex post, no validated psychometric measurements were used;*  
Not applicable, because no surveys were conducted. However, in hindsight, it might have been better to do so.
3. *Lack of control groups;*  
There was a control group, but it would have been even more valid if a control group with a dummy intervention/treatment had been included. This point is discussed in more detail below.
4. *Multiple game elements were implemented as a whole, so that the individual effect of each element could not be measured;*  
Not really applicable. No game elements were implemented per se, because the design method by Deterding was used. One could say multiple game design lenses were used, so the individual effect of each design lens cannot be measured. However, measuring individual design lens effects was never the goal of this research.
5. *Only descriptive statistics were presented, without mentioning relationships between constructs;*  
This was largely overcome. Multiple regression models were created and several interaction effects between predictor variables were modelled. The relationships between the predictor variables were mentioned, but the hypothesised relations have not been put together in a diagram, nor have all possible interaction effects been tested for.
6. *Very short timeframes for experiments, causing the novelty of an implemented game mechanic itself to be a potential factor, which was not taken into account;*  
Because of time constraints, this pitfall could not be overcome. However, for the Selling Assistant, a much longer timeframe would probably not have been beneficial or yielded in other results. 2 to 3 weeks would have arguable been the best timeframe (R Kohavi et al., 2014). This pitfall also relates to number 3 and is discussed below.
7. *Unclear reporting of results;*  
Cannot be judged by the author, but this pitfall has hopefully been overcome.
8. *No use of multi-level measurement models, including the game mechanics, game dynamics / psychological outcomes and behavioural outcomes;*  
The research as conducted has not avoided this pitfall.

#### Data Collection Methods

Not all results could be logically clarified and follow-up questions from the main hypotheses remained unanswered, only behavioural data during the experiment was gathered. Extra types of data could have added more possibilities to answer these follow-up questions and the hypotheses. The first could be qualitative data in terms of observations and interviews, which should be done during the

prototyping workshops with users and also in later workshops, with a more finished design, in a controlled environment. A second data type could concern psychometrics, which could be collected before, during and after the experiment. The bottom line is that not the only behaviourist view but also cognitivist view on gamification is important: find out what the design did with people, did they like it? Was it fun? How did their intentions and attitudes change? This way, the effect of the treatment can actually be connected to changes in motivational elements of competence, relatedness and autonomy of users (Deterding, 2014a; Groh, 2012; Ryan & Deci, 2000; Werbach, 2014b). A third useful data type would be historical user behaviour, thus also outside of the experiment. This way, users in the experiment can be given a certain profile during the data analysis and their behaviour changes can be assessed more completely. Especially this user profile data would have made it possible to answer a great deal of the unanswered questions in this research. The other two types would have been useful to generalise the results of this research.

Humans are not black boxes. Measuring only the behavioural effect of a treatment that was made with a user-centred design method seems, in hindsight, like an incomplete evaluation. However, a mixed methods approach was not possible in this research. First of all, the time constraints and practical constraints, including the possibilities at OLX, did not allow for ex ante and ex post psychological measurements. Moreover, a trade-off was made and always has to be made: working with a small user base that is aware of being in an experiment because of psychological measurements (not double-blind, so more susceptible for bias) versus working with a large user base in a double-blind experiment and measuring only behavioural data.

### Timeline and Sample Size

Ideally, the number of users needed in each variant for significant results is calculated beforehand. The experiment timeline and user percentage allocation can then be adjusted accordingly. In this case, this was not possible, because only 7 days were available for testing. The ideal period for online experiments like this one is 2-3 weeks. The minimum is 1 week, because at least each day of the week should be included, to control for potential day of the week differences (R Kohavi et al., 2014). The disadvantage of 1 week of testing is that only short-term effects of the Selling Assistant can be seen and effects on a longer term (e.g. because of users having to think longer than 1 week before making the decision to post a listing) can only be speculated upon.

### Experiment Validity

Not only a control group should be used, but also a control group with a dummy intervention. This way, the general bias of an intervention can also be controlled for. An intervention or treatment causes a positive effect in the majority of experiments (Hawthorne effect). A similar effect, mentioned in literature on online controlled experiments and gamification literature, is called the 'novelty' effect (Hamari et al., 2014; Ron Kohavi et al., 2008; Lieberoth, 2014). However, in this case, the novelty effect was not completely applicable. Namely: one would expect this effect to result in many users clicking the Selling Assistant button on the home page, as this is a new feature on the website with a prominent placement and visibility. It turned out that a very low proportion of users actually did this during the experiment. For the users who did click, the novelty effect could still be influential. It is possible that users who saw the product page posted more listings because they were intrigued with this new feature, rather than because it actually solved their (motivational) challenge.

Best practice in online controlled experiments (A/B tests) is to do an A/A test beforehand. In this A/A test, two variants are created which both contain the original website and are thus in fact exactly the same. In such an A/A test the normal variance in user behaviour can be assessed. So, an A/A test is an extra measure to increase the validity of experiment results (Crook et al., 2009; Ron Kohavi et al., 2012; Dan Siroker & Koomen, 2013). Due to time constraints and testing possibilities at OLX, an A/A test was not possible. To solve this for this research, the results of the Original variant were compared with a benchmark of normal mobile website data, in order to check if they align (see H.3). However,

this test was not very successful, as the historical was not fully reliable and also not fully comparable to the data of the Original variant.

Hypothetically, users who were included in the experiment could have deleted their cookies and thereby delete their allocation to an experiment variant. The next time they visited the mobile website, they could be assigned to another experiment variant and thus disturb the experiment results. A small percentage of result skew was expected because of this, but is hopefully equally distributed over all users in the experiment. Likewise, Optimizely uses client-side assignment, which means that users could identify (by viewing the page source) that they participated in an experiment. The client-side assignment also results in slightly longer page load times than normal, which could have had an effect on user behaviour (Crook et al., 2009; Optimizely, 2015; Dan Siroker & Koomen, 2013).

Only users who entered the OLX website via the home page were included in the experiment. This might be a special kind of users with special behaviour, complicating the generalisability of the experiment results to all of the OLX users and online marketplace users on a higher level.

All in all, most limitations consist of minor issues, which do not compromise the possibilities of making conclusions with regard to the research questions, based on the experiment results. They are mostly used as input for the future research recommendations later on in this chapter.

## 5.2 Application of the Design Method

A design method was evaluated by application to a case, which also results in recommendations to improve the design method. This is a more qualitative and practical evaluation of gamification. It considers the way of working of online marketplaces, based on experiences throughout the case study application and some literature. Also the known criticism on gamification design (as described in chapter 2.2.1) is reflected upon.

### 5.2.1 Discussion

#### Way of Working

Gamification provides handholds outside of the thought processes that are normally involved when creating new features on an online marketplace. OLX is quite a big online marketplace, with many skilled and diverse employees all over the world, but none of the people involved in this research had worked with gamification before, or heard of colleagues using game elements. Even though many of the elements could fit into OLX and are closely aligned with user interface design practices and lean working methods in general. The ideas that came out of different phases of the design process (see appendix C.1) were all generally well received by the OLX team. OLX will also continue to test some of the ideas and prototypes that came up during the design process of this research.

The gamification design method structurally combines activities of users with their motivations, needs and practical hurdles, as well as solutions to overcome these. It forces the designer to categorise and connect both the users' thought processes and their actions, which is very helpful. All elements relevant for user challenges are connected in the skill atom, which clearly indicates the interrelations and possible effects of changing one element on another element. This approach is very convenient when looking to pinpoint bottlenecks and specific changes that are needed in an online marketplace, to stimulate certain user motivations and to overcome these hurdles. The results of this research show that for those users who recognised themselves in the challenge which was chosen as bottleneck, the final outcome of the design process also very likely helped them to overcome this challenge. In other words, some changes were made to specific parts of the skill atom, which actually worked out to have an effect on the challenge that was to be resolved.

Also, the design method is very suitable to involve all types of people from the organisation, including designers, programmers, content managers and management executives, during all phases of the method. As Deterding describes in some of the case examples where he applied the design method, most of the design processes are conducted in a single day, with prototypes that are developed and ready to be implemented at the end (2014c). Moreover, users can be incorporated into the design process. This is closely aligned to the 'lean start-up' methods for business development (Ries, 2011), continuous improvement and frequent user contact. Such concepts are widely used in online businesses worldwide, including online marketplaces (Klaassen, 2014). Therefore, implementation of the design method into the current way of working should not be a big hurdle for online marketplaces.

### Influence of Problem Delineation

The problem delineation in the first phase of the design is very important. In the OLX case in this research, it was concluded that the number of new listings needed to increase. The main user challenge that was chosen to be resolved was the fact that users did not know what type of items they should or could sell on OLX. Based on this thought, the Selling Assistant was developed and chosen as most promising treatment to increase the number of new listings overall. However, in the end it turned out that the specific challenge is probably resolved, but the general target outcome of a significant increase in new listings is not met. This is not necessarily due to the concept of gamification, but more due to the decisions made in the design process, by all parties involved. In the case examples published together with Deterding's design, many specific solutions are mentioned too, improving a small part of a system to resolve a specific challenge (2014c). So it could also be more inherent to the design method and not specifically applicable only to the way the method was executed in the OLX case.

### Coping with Gamification Criticism

As can be recalled from chapter 2.2.1, the gamification criticism on designing and applying gamification is mainly based on four accounts. The criticism will be addressed here, in the light of the design process at OLX, in order to determine the way the design method as used in this research was able to cope with it.

1. Not systemic: the design method forces the designer to explicitly address general system qualities and user experiences as result of those system qualities. Step 1 of the design process allows the designer to translate system qualities and experiences into requirements and/or constraints for the treatment design, to make sure these qualities stay in place.
2. Reward-oriented: reward-oriented design is not at all applicable, since the core concept of the design method (the skill atom) focuses purely on motivation as a product of competence, autonomy and relatedness. This puts the focus on invoking intrinsic rather extrinsic motivation (Ryan & Deci, 2000).
3. Not user-centric: the business goals of the system to be gamified are definitely taken into account, but are not put above the user goals. The user goals, motivations and context are the leading principles of design, as they are to be made explicit during the design process and users are involved throughout the ideation and iterative prototyping phases in order to make sure the design complies with their needs.
4. Pattern-bound: using a list of game elements as template for design is not at all applicable in the design method used in this research. One could say it is still pattern-bound, but this concerns the finite list of design lenses that is used. The design lenses do not consist of elements that are directly inserted into a treatment design, but provide handholds to structure ideation thought processes, which does not make it pattern-bound in a confining sense.



### 5.2.2 Limitations of Design Method and Application

For the gamification design, not much can be stated here about its limitations, for it is a new method and there no real benchmark studies or theories to relate to. But, based on the description of the method and general gamification characteristics, some limitations can be identified.

Important to note here is that evaluation of gamification design is an iterative process and not a single event. A first evaluation does not only generate data on the effect of the design, but also serves as input for small design changes and further evaluation of these small changes (Deterding, 2014c; Ferrara, 2012; Hamari et al., 2014). In this research it is only possibly to perform a first evaluation step, due to time constraints. So, in fact, the results of the experiment should be input for a new series of design steps, based on the Selling Assistant. This will be discussed in the recommendations for future research, later on in this chapter.

The current description of the gameful design method steers towards defining specific targets and challenges, which automatically forces the designer to converge his or her thinking into specific solutions. While, when looking at gamification as a holistic socio-technical systems design practice (Deterding, 2014a), which is the view as adopted for this research, an overall system improvement could be realised more. This also corresponds with classical gamification views of adding points, badges and leaderboards to a website, as this proposes to stimulate overall activity, regardless of specific user contexts, activities, needs or challenges.

Another noteworthy point is that the Selling Assistant is not the typical gamification treatment one would expect, given the current gamification field. The Selling Assistant is not necessarily 'fun' and the invocation of motivational elements such as competence, autonomy and relatedness cannot be directly identified. This could limit the potential adoption and generalisability of this study. This point is further explored in chapter 5.3.3.

## 5.3 Theoretical Implications

The possible impact of this study on the gamification literature and related theory - used in this research and in general - is described in this sub chapter.

### 5.3.1 Relating the Results to Online Marketplace Imbalances

Once a problem is identified, gamification allows for very specific motivational affordances to be implemented, stimulating the users who cope with exactly this problem. When the problem or challenge that is focused on is very delineated, the result will likely also be very delineated. This is mainly useful for online marketplaces coping with a chicken-egg problem that is based on a challenge or hurdle that applies for a great share of potential users. Taking this hurdle away with use of gamification can increase the generation of UGC overall, because many users are assisted with the treatment. However, if there are all kinds of small issues, which are highly variable for each individual user, this does not apply. Targeting specific user challenges with gamification does not seem like the best approach to solve the chicken-egg problem, since that is not the real issue in the early stage of an online marketplace. In this stage, it is important to have as many active users as possible, who post as much and as diverse content as possible. This will create a significant user base to start with, which one must cherish by making sure they use the online marketplace as much as possible, see the value of it by experiencing some successful transactions that they arranged themselves (autonomy), grow into the user community (relatedness) and get the feeling of becoming better and better at its use (competency) (Armstrong, 2006; Bakos & Katsamakos, 2008; Eisenmann et al., 2006; Hunicke et al., 2004; Jordan & Hariharan, 2015; J. C. Rochet & Tirole, 2006). Thus gamification is not necessarily a very suitable method to solve the overall chicken-egg problem for online marketplaces and increase the general amount of user-generated content. Then again, if it is applied over and over again -

constantly iterating and improving, tackling different types of challenges for users one at a time - the amount of user-generated content will also increase step by step. Either way: the problem scoping is very important.

When looking to solve imbalances on an online marketplace in a later phase of growth, roughly the same applies. However, in this stage, increasing the overall amount of user-generated content is generally not the problem, but tackling specific issues for specific groups of users that do not create content, due to a shared challenge they face (Albuquerque et al., 2012; Hagiu, 2014; J. C. Rochet & Tirole, 2006). Using gamification to tackle specific problems seems more valuable in that case. More specifically, a solution such as the Selling Assistant can both nudge users to create more content (or nudge more users to create content) and nudge them into the right type of content. The Selling Assistant can do this by suggesting certain categories on the product page, in which the need for new listings is the highest, which solves specific imbalances on the marketplace.

The OLX case touches a special form of UGC, applicable exactly to online marketplaces. Namely, users do not only have to decide to place content, but to be able to do this they are dependent of the actual physical objects they have for sale. In other words, to produce content, a visitor is not only limited by motivation, but also by practical possibilities. This also applies to marketplaces mentioned earlier in the introduction of this report, such as Airbnb (do you have a home to rent out?), Kickstarter (do you have a project you need funding for?) and SnappCar (do you have a car to rent out?). Other platforms, which are of a more social network nature, don't have these practical limitations. Not having something to sell can be a very practical hurdle to post something on an online marketplace, regardless of your motivation and the trigger that stimulates you to sell (with or without gamification). This could make it harder for gamification to be successful on an online marketplace than on a general online platform. The Selling Assistant was partly meant to solve the practical hurdle, but is more connected to user knowledge than to practical user possibilities. After all, if a user is theoretically convinced by the Selling Assistant that it is beneficial to post a listing in a certain category, he/she might still be practically limited by not having an item in such a category that he/she can sell.

### 5.3.2 Gamification Design

Deterding provided a design method that has proven to be effective in this research, since the users who interacted with the gamification treatment produced significantly more content than users who did not. However, the design method can still be refined. By extensively describing the design process, reflecting on its outcomes and giving feedback as to how the method can be improved, this research hopefully helps to further shape the gamification design method. In chapter 2.4.3, several amendments to the gamification method were proposed. Some were of practical nature for this study specifically, but others were more general modifications. Here, those amendments will be reflected upon, in the light of the OLX case and online marketplaces in general.

#### Step 1: Strategy

The first amendment was to incorporate user contexts into the strategy step (the first design step). Perhaps this is something most gamification designers do by nature, but when the user base is large, diverse and multicultural, like in most online marketplaces, this is especially useful. More specifically: online marketplace owners do not meet their users on a day-to-day basis. Their behaviour is analysed every day, but the context in which the users visit the marketplace is not. Therefore, it is good to make this user context exploration explicit in the design method, in order to remind marketplace owners to go out there and get feedback from their users. This context exploration has proved useful in the OLX case in various ways. It provided the basis for the decision to host a workshop with OLX users (which was of great assistance). During the conversations with OLX employees, it caused for repetitive inquiries of users' situations, resulting in qualitative knowledge such as internet speeds and

smartphone usage behaviour, which in turn partly formed the requirements and constraints for the design.

A second amendment in the first design step was closely related to the above, as it also requires the designer to place him or herself in the mind of the marketplace users. For the OLX case, the general qualities that users assigned to the platform were actually very unspecific (free, fast, simple and good quality content) and would have otherwise also been looked after when designing solutions, because of logical reasoning. But, it can theoretically still be a useful addition to the design method. An example can be that users find an online marketplace better than competing online marketplaces because it has a certain characteristic feature. If this is the case, the gamification treatment should not jeopardise this feature, because this will probably negatively influence users' activities on the platform, regardless of the quality of the gamification treatment.

### Step 3: Synthesis

An interesting point, which relates to one of the main drawbacks mentioned in the previous sub chapter, is the practical amendment that was made in the third step of the design method. This step involves the actual system evaluation with the help of the skill atom, which can be regarded as the central element of the gamification design method used (Deterding, 2014c). Because of time constraints and delineation by OLX, not the whole OLX mobile website was evaluated in multiple skill atoms but only the core most-used elements to post listings were put into a single skill atom. So the fact that a very specific gamification treatment was created, which had an effect on only a small proportion of users, could be due to this amendment. If more skill atoms were created, perhaps more (elements of) different user challenges would have been recognised and taken into account in the design.

### Step 4: Ideation & Step 5: Iterative Prototyping

The last two steps of the design method were combined and users were involved not only in ideation but also in iterative prototyping. The idea was that prototypes or mock-ups could provide a more concrete way for users and other workshop participants to think about improvements than starting from scratch with only a skill atom and design lenses. This way, the participants can propose amendments for the prototypes but also generate new ideas. The amendment worked quite well and it was clear that the participants in the user workshop could actually reflect on the prototypes. As can be seen in appendix C.3, they gave much useful feedback which was incorporated into the final design. A downside of this approach is that in the user workshop, it slightly diminished the creativity for the following ideation phase. Many of the suggestions failed to break the realms of what had already been discussed during the prototype evaluations. For the expert workshop, this amendment did not really have a notable effect. The expert had many ideas and many comments during both the prototype evaluation and ideation phases. But, because this was a single participant, no general conclusions can be made.

### Inconsistencies

The amendments from above were made before the application of the method. During the application, some inconsistencies were encountered. The first is the somewhat vague distinction between the concepts of a 'need', a 'motivation' and a 'hurdle'. Something like 'I don't know what to sell' is more of a discouragement. It is an inverse motivation, rather than a relatively small practical gap that prevents a user to perform a certain activity. So this category could be added to step 2 of the design method (the research step). Another issue that was encountered was that the skill atom only features the concept of 'motivation' and not 'need'. However, there is a new concept in the skill atom, called 'goal'. Based on the descriptions by Deterding (2014b), it was assumed that motivation = goal and need = motivation, but this should be specified more consistently.

### 5.3.3 Rethinking Gamification

Strictly speaking, the conclusions and discussions here are limited to the perspective that was chosen within the gamification field in chapter 2. This is the process view (Werbach, 2014a), or gamification as holistic socio-technical systems design practice (Deterding, 2014a), with the corresponding gameful design method by Deterding (2014c). This is also what the word ‘gamification’ refers to in the previous sub chapters, ever since this perspective was adopted at the start of chapter 2. However, where possible, it is very relevant to relate the findings from this study to gamification in general, including the elemental definition and the traditional top-down application of standard game elements. This is relevant because a part of the objective of this research is to reflect on the usefulness of gamification as a whole, which is impossible when the discussion is limited to only a small and new proportion of the field.

#### What is Gamification?

The design method has added value by providing a new way to look at an online marketplace, by looking at user motivation and how this is translated into the system with goals, challenges, etc. Gamification as social-technical design practice can take out the game elements, because game elements are not always identifiable as such. Processes and choice models are used which stem also from other fields, such as user interface design. This makes gamification less identifiable and definable, but does increase its validity as a thorough design practice. So it seems that by evolving out of a top-down game element blueprint marketing tool into a user-centred socio-technical holistic design practice, the characteristic elements of gamification as it is (was) known are not necessarily applicable. This can be related to the Selling Assistant: it is not the typical gamification one would expect given the current field. It is not necessarily ‘fun’ and the invocation of motivational elements such as competence, autonomy and relatedness are not readily discerned. Although not necessarily a problem, it could result the statement that no true gamification treatment was applied and tested in this research, based on the current gamification definitions and views, thereby possibly limiting the adoption and generalisability of this study. Is the Selling Assistant gamification? What is the difference with other user-centred design practices that involve psychology, such as persuasive design? “Gamification attempts to affect motivations rather than attitude and/or behaviour directly, as is the case in persuasive technologies” (Hamari & Koivisto, 2013, p. 2). Based on the experiment data, the distinction between gamification and persuasive technologies cannot be made for the Selling Assistant, because no motivational changes were measured.

So what is gamification? Is gamification defined by the process by which it is created (designed), or by the intentions of the designer, or by the game-like elements that can be identified as addition to a core service or platform, or by the experience of the user of the online marketplace? Are all products as result of the gamification design method (Deterding, 2014c) automatically gamification? Answering this last question with ‘yes’ could mean that gamification treatments are less and less ‘game-like’ and more and more similar to ‘normal’ user interface design. On the other hand, it might also be the case that game elements are growing into the general toolbox of online platform designers, making it harder to pinpoint specific gamified platforms because ‘normal’ is more ‘gamified’. Another answer could be that all treatments as result of the gameful design method still have to fit the traditional elemental (Deterding et al., 2011) or experiential (Huotari & Hamari, 2012) definitions, in order to be defined as gamification. But this would not make sense, as the assumptions for these perspectives and definitions do not align.

#### When is Gamification Successful?

It’s relatively easy to agree with the new gamification view of Deterding (2014a, 2014c), since he summarises the relevant gamification developments of the past years, reflects on the definitions posed (including his own) and incorporates and acknowledges much of the criticism gamification has received. Also, the design method has proven effective in this research. However, adding simple game elements can yield the same results as applying a thorough design method, as the results of this research are similar to the results of Hamari (2014), who ‘simply’ implemented badges on an

online marketplace. Moreover: a recent empirical study by Andreas Lieberoth (2014) showed that participants who interacted with a gamified system (game look and feel and standard game elements) showed the same increase in overall motivation (measured in terms of 'value', 'enjoyment', 'autonomy', 'relatedness' and 'competence') as participants who interacted with a system that had all the interface elements of a gamified system, but no actual game mechanics involved. In other words: applying a thorough user-centered design method, which sees gamification as a holistic socio-technical design practices, was not necessarily more effective than earlier empirical gamification studies.

Important to note here is that there are different ways of evaluating gamification success. The first is more behavioural, based on target activities that are closely related to business goals. The success of gamification is based on the difference in behaviour and is measured through quantitative research methods (like in this study). The second is more cognitive, where motivation increase, enjoyment and 'gamefulness' are important performance indicators to evaluate gamification. This required qualitative research methods, such as interviews, observations and surveys. One could say that the different means of gamifying a system can only be compared when the success criteria (and thus the evaluation methods) are the same.

### Putting it into Perspective

It seems that there are now roughly three types of gamification practice. One is applying gamification as a socio-technical systems design practice, which is user-centred and defines gamification largely based on the intent of the designer and the process through which it is designed (Deterding, 2014a, 2014c; Werbach, 2014a). The second practice defines and evaluates gamification based on the appearance of the gamification treatment, its game-like feel and the emotional added value for users (Huotari & Hamari, 2011, 2012; Nicholson, 2012). The third practice is based on the game elements that are incorporated in the gamification treatment, with the belief that every element has its own effect on users and that researching these effects should be the focus of future gamification studies (Lieberoth, 2014; Thiebes et al., 2014).

The fact that there are multiple views and methods within gamification could be a positive development, in the sense that gamification is evolving from a single method into a research field. However, it also creates confusion, because the term gamification refers to multiple things, making it harder to identify relevant scientific literature. So the boundaries are again to be set. Gamification is splitting into many different directions and gamification as a term might disappear (Werbach, 2014b). It will likely become a set of design practices, which fits and connects to research fields in marketing, psychology and games.

### 5.3.4 Recommendations for Future Scientific Research

Given the results, discussion and limitations of the experiment and the design method application, several recommendations can be made to gamification researchers. Especially the limitations of this study can serve as suitable starting points for future scientific research into gamification.

#### Overall UGC Increase

The results of the experiment in this study suggest that the gamification treatment was effective only for a small proportion of users, who either coped with exactly the challenge that the treatment tried to solve or are willing to actively interact with such a treatment (or both). In other words: gamification can solve mainly specific problems and there has to be a certain willingness to interact for gamification to work. The question remains whether the design method by Deterding (2014b) and gamification as a whole can be used to engage an entire online marketplace, in order to get a significantly higher amount of user-generated content. This could not be confirmed nor renounced in this research and provides an interesting starting point for a follow-up empirical study.

### Differentiation of Gamification Approaches and Cases

An important point to investigate further is the difference of applying gamification from an traditional elemental (Deterding et al., 2011) or experiential (Huotari & Hamari, 2012) perspective and from the new process (or socio-technical design) perspective (Deterding, 2014a; Werbach, 2014a), preferably within the same case. Perhaps an incorporation of classical game elements such as points, badges and leaderboards is the most effective way of increasing the amount of UGC. Although the process perspective and design method by Deterding seem very valuable, their effectiveness could not be confirmed in this research, as it is the first empirical study to adapt the new gamification perspective.

Related to the previous point, the current research could be validated more by applying the same design method again in a different case, with the same general subject and context. If the results are similar, it can confirm the hypothesis that the design method is mainly suitable for specific problems and for users who are willing to actively engage. Thorough documentation of the design method and process in this research allows easy reproduction within another case and increases comparability of future empirical gamification studies, which is currently a problem (Hamari et al., 2014). Gamification literature demands for more empirical research and case studies; especially research that is executed with a valid evaluation methodology (Deterding, 2014a; Farzan & Brusilovsky, 2011; Groh, 2012; Hamari et al., 2014; Nicholson, 2012; Thiebes et al., 2014). Such research would be of great value, as it helps to shape the gameful design method and thus the gamification field.

### Data Collection

The motivation of OLX users to post a listing might have been greatly increased by the Selling Assistant, but this not visible in the experiment results. Incorporating qualitative information on users' motivations (psychometrics) and historical behaviour would create a more complete picture of gamification effects, allowing for more observed correlations to become confirmed causal relations. It is important to connect the behavioural effects of the treatment to changes in motivational elements of competence, relatedness and autonomy of users (Deterding, 2014a; Groh, 2012; Ryan & Deci, 2000; Werbach, 2014b). This way, both the behaviourist and cognitivist views of gamification can be taken into account. Moreover, the analysis and corresponding conclusions can be constructed on better premises.

Also, it is recommended to perform more user segmentations, in order to specify certain user characteristics that might be of influence to gamification treatment effects. In this research, the experiment showed that users from very large and somewhat smaller cities differed greatly in terms of average amount of listings posted. However, the amount of users from smaller cities was too low to acknowledge this difference as statistically significant. Such effects (what is the influence of the level of urbanisation on the susceptibility to gamification?) are very interesting to investigate in further research. Gamification literature acknowledges this need for more granular effect studies into user characteristics and demographics (Deterding, 2014b; Koivisto & Hamari, 2014; Lieberoth, 2014).



## 5.4 Managerial Implications

This sub chapter concerns the practical implications of the conclusions of this research on the decision making for managers and owners of online marketplaces. First, some specific recommendations for OLX are given, following recommendations for online marketplaces in general (which are also applicable to OLX).

### 5.4.1 Recommendations for OLX

#### Future Use of the Selling Assistant

The Selling Assistant has very likely generated extra listings, so in that sense it cannot hurt to implement it into the entire website. However, it takes up quite a prominent spot on the home page, where still only 2% of the users click it. The effect of the current Selling Assistant version might increase when the experiment time is longer. After all, the results showed that the (minimal) differences that were notable between Original, Suggest and Know could be pinpointed in returning visits, rather than new visits. In other words: users need some time after seeing the category recommendations, before they can place a listing. They have to discuss with family members or roommates, might forget about it for a few days, have to explore similar listings to come up with a competitive price, might try to sell it to their friends first, etc.

Another suggestion would be to gather more user challenges related to posting a listing; including choosing a suitable price, writing an attractive title and description, taking good pictures, etc. For all these challenges, a skill atom can be made and game design lenses can be used to generate solutions. Combining all the solutions into a more elaborate Selling Assistant, which helps OLX users with all kinds of posting challenges and has a more general call to action in the home page button, might create a good overall result in terms of increased number of listings. However, it is wise to iteratively test each aspect in a prototype and controlled experiment, to check the possible impact and the number of interested users. A first addition could be to include an average price for each featured category on the product page and observe the difference in posting behaviour. Another change could be to test a version of the Selling Assistant with a different call to action on the home page button, like 'Help Me Sell', or 'Selling Assistant', to see if this engages more users to click it (thus to evaluate if there is enough demand for assisting features, before starting to build and test them).

Applying the Selling Assistant in another OLX country would be very interesting. This could indicate whether the user context research in this research is dependent on countries or not. Also, it could indicate the generalisability of the treatment and design method to OLX as online marketplace in general, rather than just one country.

Lastly, experiment with inserting game-like aesthetics into new version of the Selling Assistant, or at least to test them. The current version is very basic and things like 'juicy feedback' (one of the game design lenses) could improve the gameful experience of the Selling Assistant, which could in turn alter user behaviour. However, this might have consequences for the speed of the iteration and prototyping, as things like this take longer to build.

#### Extra Analysis of Selling Assistant Effects

Apart from the analyses that were done with the experiment data in this research, there are extra interesting Selling Assistant effects that can be explored. For most of them, another experiment is needed, because the current data does not suffice.

- Look at different user paths throughout the experiment, mainly regarding the pages they visit after seeing the Selling Assistant product page. This can give more insight into the user-friendliness of the Selling Assistant, e.g. whether the consecutive steps are self-explanatory. Related to this, it would be interesting to see if the Selling Assistant forms a better preparation for the posting page than not having a step in-between at all. Exit rates from the posting page

between users who see the Selling Assistant and users who see the original website could be compared in order to check this.

- Investigate what conditions and characteristics make users susceptible to the Selling Assistant. This can be based on those who clicked the button: who are they and why did they click? For instance: Do new or experienced users find the Selling Assistant interesting?
- Measuring all the locations of users when possible (rather than only 8 predefined cities) could confirm the suggested differences between city types. Because the amount of data points will be larger, significant results are more easily obtained.
- Use zero-inflated and hurdle models to compare the results with the negative binomial regression models used in this research, see appendix I.1.4

### Additional Metrics & Concepts

In this research the number of new listings posted per user was the main metric on which the Selling Assistant was evaluated during the experiment. However, there are also other relevant metrics that can be used to when applying either the Selling Assistant or a new treatment to increase the amount of UGC on OLX.

- The bounce rate is the percentage of visitors on a certain page that does not continue to another page on the same website, but exits to another website or closes the browser. If more users visit the OLX mobile website, the number of listings posted through the mobile website increases. By decreasing the bounce rate, the number of so-called entered visitors can increase. The bounce rate of a page can give an indication of how user-friendly the page is, if users find what they are looking for, if they are engaged with the page content and thus if they are likely to post a listing in the future (OLX, 2014c)<sup>7</sup>.
- The quality of new listings could be used as a secondary metrics. This way, one can check if a gamification treatment contains extrinsic incentives to create more listings that overrule the intrinsic motivation to put up a listing. Such an incentive could cause a rampant growth of useless listings. For instance: when given money or points for every listing that is added, users might create a large number of fake or copied listings. These listings are of no added value to OLX or its users regarding the core value of the platform (being able to sell items from consumer to consumer) but do satisfy the extrinsic motivation of the user (to gather as many points or as much money as possible).

Also, there are more concepts that can be developed and/or tested, apart from building upon the Selling Assistant. Both treatments in the shortlist at the end of the design process (see chapter 3.2.4) were essentially suitable to be implemented and tested in an experiment, but due to time constraints only the Selling Assistant was chosen. In general, many ideas as result of the gamification design process have not been tested but were well received by OLX. Furthermore, a more complex gamified system could be developed to keep users engaged over time, during different lifecycle stages. So more possibilities for testing, as a follow-up of this research, are definitely present.

### 5.4.2 Recommendations for Online Marketplaces in General

Recommending the Selling Assistant to online marketplaces in general would be unwise, because it goes against the view of gamification as a socio-technical design process, which is highly influenced by specific contextual factors of a marketplace and its users. However, in general, making the offer and demand on an online marketplace more explicit to its users does seem a valuable way of solving supply and demand imbalances. It corresponds with the findings of what makes online marketplaces successful: transparency and full information for users, no middleman, fast and easy transactions (Hagiu, 2014; Seamans & Zhu, 2014). The Selling Assistant does provide a way to suggest very specific categories with a high demand to sellers, in order to even out the supply and demand imbalances on an online marketplace.

---

<sup>7</sup> A presentation on extensive internal data analysis of user behaviour done by OLX

## Designing Gamification

For online marketplaces, using gamification to increase user-generated content can mean many things. It can mean to use a trick toolbox to upgrade your platform, it can mean to structurally think about a specific user problem from the perspective of a game designer, your users and your experts, it can mean to add a few buttons and change a colour here and there, or it can mean anything in-between. An important thing to do is to ask yourself some gamification design fit questions very early on in the process (see chapter 3.2.2), to see which type of solution is likely and what meaning of gamification is probably the most suitable and capable to solve the current problem(s).

A big advantage of the gamification design method by Deterding is that it provides step-by-step guidance. Furthermore, this research provides a detailed reference case and the materials for the design methods are readily available. Instead of requiring a gamification service company that uses its own methods, possibly based on anecdotal experience, the gameful design method provides synthesised knowledge from academic theory and industry practice. More practically, gamification provides handholds outside of the thought processes that are normally involved when creating new features on an online marketplace. The gamification design method (Deterding, 2014c) structurally combines activities of users with their motivations, needs and practical hurdles, as well as solutions to overcome these. It forces the designer to categorise and connect both the users' thought processes and their actions, which is very helpful. Moreover, implementation of the design method into the current way of working of most online marketplaces goes quite seamlessly, as it aligns with widely adopted methods of online companies, including the 'lean start-up' methodology (Ries, 2011) with continuous improvement cycles and frequent user contact (Klaassen, 2014).

## Continuous Improvement

When using the gamification design method, it is recommended to have multiple iterative prototyping and testing phases, with a larger user share involved in each step. Before putting designs live to many users and testing the behavioural impact, a bridge should be made from the qualitative testing of prototypes and mock-ups. This could be in the form of a mixed methods study with 10-100 users, with the actual form of the final design, measuring both qualitative and quantitative results. Ex ante and ex post interviews should be conducted in order to filter out the thoughts and reasoning and improve the design in another iterative loop, before introducing it to a large share of the users and measuring only the quantitative behavioural impact (and losing the possibility to do user interviews). Additionally a balance needs to be found between two sides. One is fully working out all the designs and prototypes, getting exactly the right feedback but not doing many tests and iteration. The other is doing many iterations quick and dirty to see the results as fast as possible, with minimum viable products, doing many iterations.

Lastly, when a successful treatment is found and implemented, online marketplaces should not stop improving their platform. Continuous improvement is essential to keep the flywheel of both the selling and buying side spinning (Hagiu, 2014; Jordan & Hariharan, 2015; Ries, 2011) and Deterding's gamification design method is a very helpful and pragmatic means to assist with this.

## Measuring Success

Using gamification to increase UGC was translated to the amount of new listings for the OLX case in this research, but UGC could also be translated into other user activities on marketplaces. The direct amount of new content posted by users might not be the best and is certainly not the only metric to evaluate the effect of gamification on an online marketplace, when looking to solve challenges related to network effects. Other suggested metrics are engagement (bounce rate), content quality and amount views and reactions for listings. Moreover, users can be distinguished into valuable and less valuable segments. A new user who posts content is more valuable than an experienced user who posts content. After all, getting more buyers to also become sellers is more valuable (and harder) than making sellers more active, in the initial stage of online marketplace growth (Hagiu, 2014; Jordan, Hariharan, & Copeland, 2014).

## 6 CONCLUSIONS & REFLECTIONS

---

*Chapter six gathers all information from the experiment results, discussions, research limitations and the definitive answers to the sub research questions, in order to answer the main research question: what is the suitability and capability of gamification to increase the amount of user-generated content on online marketplaces? When the research questions have been answered, the reflections give a more zoomed out and slightly personal reflection of the research project. The added value for BLOOM, the project complications and the findings during the workshops are discussed.*

---

### 6.1 Conclusions

The sub research questions have been provisionally answered throughout the different chapters of this report, where relevant. Here, in the light of all knowledge gathered, the definitive answers are summarised, in order to come to the conclusions that logically follow out of the answer to the main research question.

#### 6.1.1 Answering the Sub Research Questions

1. *What is the origin and definition of gamification and how can it be delineated from similar fields?*

A relatively new perspective on gamification was adopted, which proposes to combine the other two 'traditional' perspectives into a more socio-technical view, recognising the gaming elements, but also the context and motivational factors. "Gamification is a holistic socio-technical systems design practice" (Deterding, 2014a, pp. 312–313). The goal of gamification is then to afford systemic, emergent motivational experiences in social-technical systems. There are two major differences to identify, when delineating gamification from its most related fields. The first is the difference between gaming and playing. Playing is the primary form of spontaneity, joy and improvisation and occurs without pre-defined rules; whereas gaming is bounded by rules and arbitrary obstacles. The second is the difference between whole games and partial games. Gamification is a partial game, serious games are whole games and playful design or interaction are partial but concern playing rather than gaming (Deterding et al., 2011; Groh, 2012).

2. *What lessons can be learned from gamification criticism and previous gamification studies?*

Gamification criticism mainly stems from the application of gamification as an oversold marketing tool, where game elements are simply added to a platform or service with the expectation that they will be engaging anyway. The main points of criticism are that this type of application is not systemic, reward-oriented, not user-centric and pattern-bound (Burke, 2014a; Chorney, 2012; Deterding, 2014a; Farrington, 2011; Hamari, 2013). Lessons extracted from relevant case studies of empirical gamification research are mainly that the context and users of the system to be gamified must not be ignored, users are not puppets (Deterding, 2014a; Nicholson, 2012; Richards et al., 2014). Also, it was found that a randomised controlled experiment is very useful to quantitatively evaluate the effect of a gamification treatment, especially when dealing with an online platform (Farzan et al., 2008a; Ron

Kohavi et al., 2008). Specific pitfalls of previous gamification studies were also extracted (Hamari et al., 2014)

3. *Which design method can be used to structurally gamify online marketplaces?*

Deterding (2014c) wrote an article in which he reviews relevant game and gamification design methods from scientific and management literature, lists gamification criticism and lessons to be learned and synthesises all into a new and comprehensive 'gameful design' method. This design method was used as reference point for gamification design in this research for three reasons. First, it is the only gamification design method available. Second, it incorporates most of the theory that was found in literature. Third, the general theoretical perspective of Deterding (2014a) was adopted throughout the research, so using the corresponding design method seems fitting.

4. *Through application of the chosen design method: what is the most promising treatment to gamify an online marketplace in order to increase the amount of user-generated content?*

The Selling Assistant is a gamification treatment that was developed during the initial ideation step of the design process. Through two workshops with users and an OLX expert and through discussions with the OLX India team, the treatment prototypes were refined over and over again. It was selected to be implemented and tested in an experiment. One of the biggest challenges for users, which prevents them from starting the posting process, is that they do not know what to sell. The idea of the Selling Assistant is that appealing explicitly to this challenge with a call to action will engage the users who cope with the described challenge. Breaking the challenge into smaller steps by providing guidance and a limited number of choices for items to sell will help those users. The next best action is to select which product to sell from a small list, rather than 'sell something'.

5. *What is the effect of the gamification treatment on the amount of content generated by users?*

The Selling Assistant caused an increase in the amount of user-generated content. The number of listings per user was at least 18% higher for the treatment groups than the control group, but this was not statistically significant. Within the treatment groups, users who actively engaged with the Selling Assistant by clicking on its button on the home page posted more than 6 times as many listings on average during the experiment than users who did not engage with the Selling Assistant, controlling for other predictor variables such as location, browser, number of page views and user source. In the treatment groups, only 2.1% and 2.4% actively interacted with the Selling Assistant, which is the reason that the overall differences between treatment and control are not significant. Looking at new and returning visits of users during the experiment learns that users generally post a listing in a returning visit, not in their first visit. Thus, the results of the experiment in this study suggest that the gamification treatment was effective only for a small proportion of users, who either coped with exactly the challenge that the treatment tried to solve or are willing to actively interact with such a treatment (or both).

6. *What is the suitability of the design method for the way of working of online marketplaces?*

The gamification design method provides handholds outside of the thought processes that are normally involved when creating new features on an online marketplace, by specifically taking a game designer perspective. The ideas that came out of different phases of the design process were all generally well received by the OLX team. It structurally forces the designer to categorise and connect both the users' motivations and their actions, which is very helpful. The design method closely aligns with the 'lean start-up' methods for business development (Ries, 2011), including continuous improvement, many product tests and frequent user contact. Such concepts are widely used in online businesses worldwide, including online marketplaces (Klaassen, 2014). Therefore, implementation of the design method into the current way of working should not be a big hurdle for online marketplaces.

### 6.1.2 Using Gamification to Increase UGC on Online Marketplaces

With the answers of all the sub research questions, including the qualitative evaluation of the design method application and the quantitative evaluation of the gamification treatment in an experiment, the main research question can be answered.

*What is the suitability and capability of gamification to increase the amount of user-generated content on online marketplaces?*

This research set out to investigate how gamification could stimulate the amount of user-generated content on online marketplaces. OLX was used as a case online marketplace, which demanded for more listings, to be created by its users. A novel but theoretically sound gamification perspective was adopted and the corresponding design method was used to create a gamification treatment, through several iterative prototyping and ideation workshops with OLX users and experts. The treatment was quantitatively evaluated in a randomised controlled experiment. Also, the design method itself was qualitatively evaluated by reflecting on its theoretical setup and by assessing its fit to the general way of working of online marketplaces. These evaluations respectively determine the capability and suitability as referred to in the research question.

First of all: the capability, which is the most significant part of this research. A randomised controlled experiment was used to test the effect of the gamification treatment on the OLX user behaviour, for which the main metrics was the number of listings per user. This effect was definitely notable, as the treatment groups showed a distinct increase of posting behaviour as compared to the control group. However, overall significant differences in the amount of generated listings between control and treatment groups could not be confirmed. The gamification treatment was effective for a small proportion of users, who either coped with exactly the challenge that the treatment tried to solve or were willing to actively interact with such a treatment (or both). Those users posted a remarkably higher number of listings than users who saw the treatment but did not interact with it. In other words: in order for gamification to be successful, the first step still needs to be taken by the users. These conclusions closely align with a another empirical gamification study conducted at an online marketplace (Hamari, 2013), which used a different gamification perspective and design approach. This increases the generalisability of the findings of this study to gamification of online marketplaces in general. Gamification allows for very specific motivational affordances to be implemented, stimulating the users who cope with exactly this problem. When the problem or challenge that is focused on is very delineated, the result will likely also be very delineated. This makes gamification less useful for the chicken-egg problem or other early stage network effect problems, as in such cases the overall user base needs to be increased. When looking to solve imbalances on an online marketplace in a later stage of growth, increasing the overall amount of user-generated content is generally not the problem, but rather tackling specific issues for specific groups users that do not create content, due to a shared challenge they face. Using gamification to tackle specific problems seems more valuable in such cases. Furthermore, practical hurdles for users to post content, rather than motivational hurdles, could make it harder for gamification to be effective on an online marketplace than on a general online platform (such as a social network), where these practical hurdles do not apply.

Regarding the suitability of gamification to online marketplaces, Deterding provided a structured and useful design method, which fits quite seamlessly into the way of working of OLX and many other online marketplaces. Its elements of continuous improvement and iterative prototyping as well as frequent users contact with user-centered design are well known concepts for such organisations. An added value of the design method is its way to connect users motivations to activities, to force the designer to connect with the marketplace user and to structurally assess the marketplace from a game designer perspective (with novel insights and ideas that lie outside the normal thought realms of most marketplace employees). The design method synthesises an existing variety of methods and design practices into one, with solutions to most of the known gamification criticism incorporated. This



method creates a potential single starting point for structured gamification research, which can be compared and evaluated on the same grounds. Studies using the design method are more comparable than studies using the game elements approach, because game elements are infinite, not mutually exclusive and highly variable across available gamification theory sources.

So, it can be concluded that gamification is capable, when certain preconditions regarding the marketplace growth stage and the willingness of users to interact are in place, and suitable to increase the amount of user-generated content on online marketplaces. Nevertheless, the exploratory nature of this study should not be forgotten. The adopted gamification perspective and corresponding design method are novel, within the already novel and manifold gamification field. Conclusions made are therefore not necessarily definitive, but provide starting points for further research. Many questions remain regarding the type of users that was mostly affected and the effectiveness of the adopted gamification perspective compared to other perspectives, when applying them to the same online marketplace. This confirms the need for more empirical research; especially with a mixed methods experimental approach that incorporates behavioural experiment data, historical user data and cognitive psychometrics.

## **6.2 Personal Reflections**

This sub chapter features some personal reflections on the practical issues in the design process, experiment and research project in general.

### **6.2.1 Design Process**

In general, both of the workshops conducted were very useful, because the requirements and constraints were mostly shaped during the two workshops. The Indian students were very good at evaluating the prototypes and mock-ups, but not so much at ideation. The difference in terms of ideation between the user workshop with Indian students and the expert workshop was vast. Also, the Indian students had a lot input from user side, but found it difficult to put on the game designer glasses. It might be better to have a more mixed group of people involved, so that all aspects are covered by participant expertise. Meaning: a gamification expert, a product expert, a design expert, a sales expert, users, etc. Ideally, multiple sessions with small groups of around 10 people are held. This way, the sessions can be compared and groupthink can be overcome.

A gamification expert was not used, because the known experts were not available at the time. However, it would be helpful to involve them. This way, experts on all three sides of the important participant sides can be used: the users, the marketplace and the gamification method. A gamification expert could help with the ideation and translation of design lenses to the concerned platform.

### **6.2.2 Experiment**

OLX focused more on user-interface design and business goals and I focused more on the general idea of the treatment to be tested and the design lenses that needed to be expressed to the users through the treatment. Exemplary for this is that a third of the users in the experiment was allocated to Know, as a 'different' version of Suggest, with only different wording of main treatment page button. My preference would have been to use those 33% of available users to test an entirely different treatment. In general, most A/B tests fail to produce significantly different results between treatment and control, so it is wise to test as much as possible and as many diverse ideas as possible (R Kohavi et al., 2014)

Optimizely is not an entirely reliable service to use in experiments like the one in this research. It does not recognise statistical power, recommends shutting down experiments when significance is achieved before sample size is reached, uses one-tailed tests, etc. The objective of Optimizely is to get their clients to think they conduct as many successful experiments as possible, rather than to conduct valid experiments and do valid statistical analysis (Borden, 2014). The statistical tests that are done on the Optimizely web interface are all Z-tests for proportions, so the effects can never be controlled for other variables (Dan Siroker & Koomen, 2013). Also, the web interface does not nearly provide all information needed (for instance, only unique actions by visitors are shown, not a total amount of actions, which was needed to calculate the main metric of the experiment results). Therefore, make sure to conduct all analyses from the raw data yourself.

### 6.2.3 Research Project

#### Added Value for BLOOM

A large share of BLOOM's clients consists of online marketplaces that cope with the problems related to the network effect from day to day. Often, BLOOM works hard to help their clients overcome these problems. For BLOOM, this research has made the following contributions:

- An overview of gamification possibilities and limitations, including underlying theory, to be added to the common knowledge base of the BLOOM consultants;
- A new expertise subject for workshops (gamification) to be given to clients;
- A practically tested design method to apply in client projects;
- A case to refer to when starting new projects at online marketplaces.

#### Project Challenges

Several aspects of this research project were not beneficial for the process speed. The first is a common problem when dealing with multiple stakeholders who have a business interest as main priority, where the priority of the researcher is to conduct scientifically valid research. For instance, small categories were not featured in the Selling Assistant, simply because they would have a small impact on the business goals. Even though some of those categories would probably be very appealing to all types of users, since they consisted of products that are present in almost every home. A second aspect was the dependency on the testing capacity of OLX. The experiment was delayed many times and communication on the planning and expectations was hard, since it was always done via Skype and there were many employees involved at different levels, with somewhat diverging interests. Perhaps, in hindsight, a temporary presence in India would have been helpful in the communication. Thirdly, the novelty of the gamification field makes it hard to define and shape the research. Not many existing case or literature studies were able to take as an example, making the delineation of research methods and relevant literature quite hard.

## BIBLIOGRAPHY

---

- Albuquerque, P., Pavlidis, P., Chatow, U., Chen, K.-Y., & Jamal, Z. (2012). Evaluating Promotional Activities in an Online Two-Sided Market of User-Generated Content. *Marketing Science*, 31(3), 406–432. doi:10.1287/mksc.1110.0685
- Armstrong, M. (2006). Competition in Two-Sided Markets. *The RAND Journal of Economics*, 37(3), 668–691. doi:10.1111/j.1756-2171.2006.tb00037.x
- Bakos, Y., & Katsamakas, E. (2008). Design and Ownership of Two-Sided Networks: Implications for Internet Platforms. *Journal of Management Information Systems*, 25(2), 171–202. doi:10.2753/MIS0742-1222250208
- Batra, A. (2014). *Naspers: India eCommerce Day OLX India [Presentation slides]*. Bangalore, India. Retrieved from <http://www.naspers.com/pdf/presentations/other/olx.pdf>
- Bogost, I. (2011). Gamification is Bullshit. Retrieved April 30, 2015, from [http://bogost.com/writing/blog/gamification\\_is\\_bullshit/](http://bogost.com/writing/blog/gamification_is_bullshit/)
- Borden, P. (2014). How Optimizely (Almost) Got Me Fired. Retrieved March 16, 2015, from <http://blog.sumall.com/journal/optimizely-got-me-fired.html>
- Borys, M., & Laskowski, M. (2013). Implementing Game Elements into Didactic Process: A Case Study. In *Proceedings of Active Citizenship by Management, Knowledge Management & Innovation Knowledge and Learning* (pp. 819–824). Zadar, Croatia. Retrieved from <http://www.toknowpress.net/ISBN/978-961-6914-02-4/papers/ML13-326.pdf>
- Burke, B. (2014a). *Gamify: How Gamification Motivates People to do Extraordinary Things*. Brookline, USA: Bibliomotion.
- Burke, B. (2014b). Gartner Redefines Gamification. Retrieved from [http://blogs.gartner.com/brian\\_burke/2014/04/04/gartner-redefines-gamification/](http://blogs.gartner.com/brian_burke/2014/04/04/gartner-redefines-gamification/)
- Caillaud, B., & Jullien, B. (2003). Chicken & Egg: Competition among Intermediation Service Providers. *The RAND Journal of Economics*, 34(2), 309–328. Retrieved from <http://www.jstor.org/stable/1593720>
- Caillois, R. (1961). *Man, Play, and Games*. Urbana, USA: University of Illinois Press.
- Cambini, C., Meccheri, N., Silvestri, V., Torino, P., Duca, C., & Pisa, U. (2011). Competition, efficiency and market structur in online digital markets. An overview and policy implications. *ERIEP*, (2).
- Cameron, C., & Trivedi, P. (1998). *Regression Analysis of Count Data (Econometric Society Monographs)* (0 ed.). Cambridge, United Kingdom: Cambridge University Press.
- Carron, T., Kordon, F., Labat, J. M., Mounier, I., & Yessad, A. (2013). Toward improvement of serious game reliability. In *7th European Conference on Games Based Learning, ECGBL 2013* (Vol. 1, pp. 80–87). Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84893701039&partnerID=tZOtx3y1>
- Chorney, A. I. (2012). Taking The Game Out Of Gamification. *Dalhousie Journal of Interdisciplinary Management*, 8(May), 1–14. doi:10.5931/djim.v8i1.242

- Crook, T., Frasca, B., Kohavi, R., & Longbotham, R. (2009). Seven Pitfalls to Avoid when Running Controlled Experiments on the Web. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1105–1114). New York, NY, USA: ACM. doi:10.1145/1557019.1557139
- DasGupta, A., Cai, T. T., & Brown, L. D. (2001). Interval Estimation for a Binomial Proportion. *Statistical Science*, 16(2), 101–133. doi:10.1214/ss/1009213286
- Deng, A., Li, T., & Guo, Y. (2014). Statistical Inference in Two-stage Online Controlled Experiments with Treatment Selection and Validation. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 609–618). New York, NY, USA: ACM. doi:10.1145/2566486.2568028
- Denny, P. (2013). The effect of virtual achievements on student engagement. In *Proceedings of CHI 2013: Changing Perspectives* (pp. 763–772). Paris, France: ACM Press. doi:10.1145/2470654.2470763
- Deterding, S. (2013). Skill Atoms as Design Lenses for User-Centered Gameful Design. In *Proceedings of CHI 2013 (Workshop "Designing Gamification")*. Paris, France: ACM.
- Deterding, S. (2014a). Eudaimonic Design, or: Six Invitations to Rethink Gamification. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 305–331). Lüneburg, Germany: meson press.
- Deterding, S. (2014b). Gamification Absolved? Retrieved May 2, 2015, from <http://gamification-research.org/2014/08/gamification-absolved-2/>
- Deterding, S. (2014c). The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. *Human-Computer Interaction*, (Special Issue "HCI and Digital Games," Forthcoming). Retrieved from <http://ssrn.com/abstract=2466871>
- Deterding, S., Björk, S., Nacke, L., Dixon, D., & Lawley, E. (2013). Designing Gamification: Creating Gameful and Playful Experiences. In *CHI 2013 Extended Abstracts* (pp. 3263–3266). New York, USA: ACM.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification." In *Proceedings of the 15th International Academic MindTrek Conference* (pp. 9–15). ACM.
- Dym, C. L., & Little, P. (2009). *Engineering Design: A Project Based Introduction*. Hoboken, USA: John Wiley and Sons Ltd.
- eBay International AG. (2015). Over Marktplaats. Retrieved January 21, 2015, from <http://www.marktplaats.nl/i/help/over-marktplaats/>
- Eisenmann, T. R., Parker, G., & Alstyne, M. W. Van. (2006). Strategies for Two-Sided Markets. *Harvard Business Review*, 84(10).
- Farrington, J. (2011). From the Research: Myths Worth Dispelling - Seriously, the Game Is Up. *Performance Improvement Quarterly*, 24(2), 105–110. doi:10.1002/piq
- Farzan, R., & Brusilovsky, P. (2011). Encouraging user participation in a course recommender system: An impact on user behavior. *Computers in Human Behavior*, 27(1), 276–284. doi:10.1016/j.chb.2010.08.005
- Farzan, R., Dimicco, J. M., Millen, D. R., Brownholtz, B., Geyer, W., & Dugan, C. (2008a). Results from Deploying a Participation Incentive Mechanism within the Enterprise. In *Proceedings of the*

- SIGCHI Conference on Human Factors in Computing Systems* (pp. 563–572). Florence, Italy: ACM.
- Farzan, R., Dimicco, J. M., Millen, D. R., Brownholtz, B., Geyer, W., & Dugan, C. (2008b). When the experiment is over: Deploying an incentive system to all the users. In *Symposium on Persuasive Technology*. Aberdeen, Scotland.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. doi:10.3758/BRM.41.4.1149
- Ferrara, J. (2012). *Playful Design: Creating Game Experiences in Everyday Interfaces*. New York, USA: Rosenfeld Media, LLC.
- Forrester Consulting. (2014). *Unlock The Promise Of Customer Data: How Comprehensive Profiles Hold The Key To Effective Personalization*. Cambridge, USA.
- Gamification Corp. (2014). Vendors. Retrieved from <http://www.gamification.co/vendors/>
- Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression Analyses of Counts and Rates: Poisson, Overdispersed Poisson, and Negative Binomial Models. *Psychological Bulletin*, 118, 392–404.
- Goldsmith, R. E., Pagani, M., & Lu, X. (2013). Social network activity and contributing to an online review site. *Journal of Research in Interactive Marketing*, 7(2), 100–118. doi:10.1108/JRIM-Jul-2012-0029
- Google. (2014). Trends. Retrieved from <http://www.google.com/trends/explore#q=gamification&date=1/2010 58m&cmpt=q>
- Griffith, E. (2014). Meet OLX, the biggest Web company you've never heard of.
- Groh, F. (2012). Gamification: State of the Art Definition and Utilization. In *4th Seminar on Research Trends in Media Informatics* (pp. 39–46). Ulm, Germany: Institute of Media Informatics Ulm University.
- Hagiu, A. (2014). Strategic decisions for multisided platforms. *MIT Sloan Management Review*, 55(2), 71–80. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84893116312&partnerID=tZOtx3y1>
- Hamari, J. (2013). Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications*, 12(4), 236–245. doi:10.1016/j.elerap.2013.01.004
- Hamari, J., & Koivisto, J. (2013). Social Motivations To Use Gamification: An Empirical Study Of Gamifying Exercise. In *Proceedings of the European Conference on Information Systems*. Utrecht, The Netherlands.
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? A Literature Review of Empirical Studies on Gamification. *Proceedings of the 47th Hawaii International Conference on System Sciences*, 3025–3034. doi:10.1109/HICSS.2014.377
- Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, USA: Guilford Press.
- Hense, J., Klevers, M., Sailer, M., Horenburg, T., Mandl, H., & Günthner, W. (2014). Using gamification to enhance staff motivation in logistics. In *Lecture Notes in Computer Science*

(including subseries *Lecture Notes in Artificial Intelligence* and *Lecture Notes in Bioinformatics*)  
(Vol. 8264 LNCS, pp. 206–213). doi:10.1007/978-3-319-04954-0-24

Hoaglin, D. C. (1980). A Poissonness Plot. *The American Statistician*, 34(3), 146–149.

Hunicke, R., Leblanc, M., & Zubek, R. (2004). MDA : A Formal Approach to Game Design and Game Research. In *Proceedings of the Challenges in Games AI Workshop, Nineteenth National Conference of Artificial Intelligence*.

Huotari, K., & Hamari, J. (2011). “ Gamification ” from the perspective of service marketing. In *Proceedings of CHI 2011 (Gamification workshop)* (pp. 11–15). Vancouver, Canada: ACM.

Huotari, K., & Hamari, J. (2012). Defining gamification - A service marketing perspective. In *Proceedings of the 16th International Academic Mindtrek Conference* (pp. 17–22). Tampere, Finland. Retrieved from <http://dl.acm.org/citation.cfm?id=2393137>

Hynninen, P., & Kauppinen, M. (2014). A/B testing: A promising tool for customer value evaluation. In *2014 IEEE 1st International Workshop on Requirements Engineering and Testing, RET 2014 - Proceedings* (pp. 16–17). Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84908614242&partnerID=40&md5=fc68938c8e972094f9c505cd1f3b82dd>

Jordan, J., & Hariharan, A. (2015). a16z Podcast: The Marketplace Rules [audio podcast]. Retrieved March 27, 2015, from <https://soundcloud.com/a16z/a16z-podcast-the-marketplace-rules>

Jordan, J., Hariharan, A., & Copeland, M. (2014). a16z Podcast: The Marketplace Rules. a16z. Retrieved from <https://soundcloud.com/a16z/a16z-podcast-the-marketplace-rules>

Jung, J. H., Schneider, C., & Valacich, J. (2010). Enhancing the Motivational Affordance of Information Systems: The Effects of Real-Time Performance Feedback and Goal Setting in Group Collaboration Environments. *Management Science*, 56(4), 724–742. doi:10.1287/mnsc.1090.1129

Katz, M. L., & Shapiro, C. (1985). Network Externalities, Competition, and Compatibility. *The American Economic Review*, 75(3), 424–440. Retrieved from [http://links.jstor.org/sici?sici=0002-8282\(198506\)75:3<424:NECAC>2.0.CO;2-M](http://links.jstor.org/sici?sici=0002-8282(198506)75:3<424:NECAC>2.0.CO;2-M)

Keynote Systems. (2012). Mobile User Survey: mobile user preferences, habits and expectations. Retrieved November 12, 2014, from [http://www.slideshare.net/keynote\\_systems/keynote-mobile-user-survey-1-h2012](http://www.slideshare.net/keynote_systems/keynote-mobile-user-survey-1-h2012)

Khatibloo, F. (2011). *Personal Identity Management*. Cambridge, USA.

Klaassen, E. (2014). Wat grote bedrijven kunnen leren van succesvolle start-ups. *Holland Management Review*, 158(November-December), 15–21.

Kohavi, R., Deng, A., Frasca, B., Longbotham, R., Walker, T., & Xu, Y. (2012). Trustworthy Online Controlled Experiments: Five Puzzling Outcomes Explained. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 786–794). New York, NY, USA: ACM. doi:10.1145/2339530.2339653

Kohavi, R., Deng, A., Longbotham, R., & Xu, Y. (2014). Seven rules of thumb for web site experimenters. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1857–1866). Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84907031426&partnerID=40&md5=34d17f36b4f0a3cd7be70fc5d1af7f66>

- Kohavi, R., Longbotham, R., Sommerfield, D., & Henne, R. M. (2008). Controlled experiments on the web: survey and practical guide. *Data Mining and Knowledge Discovery*, 18(1), 140–181. doi:10.1007/s10618-008-0114-1
- Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior*, 35, 179–188. doi:10.1016/j.chb.2014.03.007
- Lepper, M., Greene, D., & Nisbett, R. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129–137. doi:http://dx.doi.org/10.1037/h0035519
- Lieberoth, A. (2014). Shallow Gamification: Testing Psychological Effects of Framing an Activity as a Game. *Games and Culture*, in review, 1–20. doi:10.1177/1555412014559978
- Lowe, J. (2014). OLX Announces Expansion in Eleven International Markets. Retrieved April 21, 2015, from <http://www.businesswire.com/news/home/20140916005335/en/OLX-Announces-Expansion-Eleven-International-Markets>
- Naspers. (2014). OLX. Retrieved April 21, 2015, from <http://www.naspers.com/page.html?pageID=20&parentID=70&contentIDChosen=4&operation=olx>
- Nicholson, S. (2012). A User-Centered Theoretical Framework for Meaningful Gamification. In *Games+Learning+Society 8.0* (pp. 1–7). Madison, U.S.: Syracuse University.
- Nussbaum, E., Elsadat, S., & Khago, A. (2008). Best Practices in Analysing Count Data - Poisson Regression. In J. Osborne (Ed.), *Best practices in quantitative methods* (pp. 306–324). Thousand Oaks, California, US: SAGE Publications, Inc. doi:http://dx.doi.org/10.4135/9781412995627.d26
- OLX. (2014a). *Qualitative User Behaviour and Motivation Research [report, not publicly accessible]*.
- OLX. (2014b). *Qualitative User Motivation Research [presentation, not publicly accessible]*.
- OLX. (2014c). *Quantitative User Behaviour Research [presentation, not publicly accessible]*.
- OLX. (2015a). Free Classifieds In India. Retrieved January 22, 2015, from <http://olx.in/i2/>
- OLX. (2015b). India's Largest Marketplace. Retrieved January 22, 2015, from <http://olx.in/>
- Oprescu, F., Jones, C., & Katsikitis, M. (2014). I PLAY AT WORK-ten principles for transforming work processes through gamification. *Frontiers in Psychology*, 5(1), 1–5. doi:10.3389/fpsyg.2014.00014
- Optimizely. (2015). Stats Engine: How Optimizely calculates results to enable business decisions. Retrieved March 16, 2015, from [https://help.optimizely.com/hc/en-us/articles/200039895-Stats-Engine-How-Optimizely-calculates-results-to-enable-business-decisions?flash\\_digest=6cda06633e8c575aedf1b1ce4cad8cd3dc97bc08](https://help.optimizely.com/hc/en-us/articles/200039895-Stats-Engine-How-Optimizely-calculates-results-to-enable-business-decisions?flash_digest=6cda06633e8c575aedf1b1ce4cad8cd3dc97bc08)
- Owyang, J., Tran, C., & Silva, C. (2013). *The Collaborative Economy*. . . . slideshare.net/Altimeter/the-collaborative-economy. San Mateo, USA. Retrieved from <http://www.collaboriamo.org/media/2014/04/collabecon-draft16-130531132802-phpapp02-2.pdf>
- Paharia, R. (2013). *Loyalty 3.0: How to Revolutionize Customer and Employee Engagement with Big Data and Gamification*. New York, USA: McGraw-Hill.



- Peña-López, I. (2010). *The Economic and Social Role of Internet Intermediaries*. Retrieved from <http://ictlogy.net/bibliography/reports/projects.php?idp=1608>
- Pettey, C., & van der Meulen, R. (2014). Gartner Predicts Over 70 Percent of Global 2000 Organisations Will Have at Least One Gamified Application by 2014 [Press Release]. Retrieved from <http://www.gartner.com/newsroom/id/1844115>
- PTI. (2014). India has maximum listings, 80% traffic coming from phones: OLX. Retrieved January 21, 2015, from [http://articles.economictimes.indiatimes.com/2014-11-03/news/55720334\\_1\\_olx-in-mobile-users-traffic](http://articles.economictimes.indiatimes.com/2014-11-03/news/55720334_1_olx-in-mobile-users-traffic)
- Richards, C., Thompson, C. W., & Graham, T. C. N. (2014). Beyond Designing for Motivation : The Importance of Context in Gamification. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play (CHI PLAY '14)* (pp. 217–226). New York, USA: ACM. doi:10.1145/2658537.2658683
- Ries, E. (2011). *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses*. New York, USA: Crown Publishing Group.
- Robertson, M. (2010). Can't play, won't play. Retrieved April 30, 2015, from <http://hideandseek.net/2010/10/06/cant-play-wont-play/>
- Rochet, J. C., & Tirole, J. (2006). Two-sided markets: A progress report. In *RAND Journal of Economics* (Vol. 37, pp. 645–667). Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-34248324442&partnerID=tZOtx3y1>
- Rochet, J.-C., & Tirole, J. (2003). Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, 1(4), 990–1029. doi:10.1162/154247603322493212
- Roe, D. (2012). Forrester: Why Customer Experience Management Needs To Be Personalized. Retrieved from <http://www.cmswire.com/>
- Rohlfs, J. H. (2003). *Bandwagon Effects in High Technology Industries*. Cambridge, USA: MIT Press.
- Ryan, R., & Deci, E. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67. doi:10.1006/ceps.1999.1020
- Sathe, G. (2015). Goods Worth Rs. 56,200 Crores Lying Unused in Homes Across India: OLX.
- Schell, J. (2010). Visions of the Gamepocalypse [Online Presentation]. San Francisco, USA: Long Now Foundation, July 27.
- Seamans, R., & Zhu, F. (2014). Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers. *Management Science*, 60(2), 476–493. doi:10.1287/mnsc.2013.1785
- SimilarWeb. (2015). OLX.in Traffic Overview. Retrieved April 21, 2015, from <http://www.similarweb.com/website/olx.in>
- Siroker, D., & Koomen, P. (2013). *A/B Testing: The Most Powerful Way to Turn Clicks Into Customers* (1st ed.). Wiley Publishing.
- Siroker, D., Koomen, P., Kim, E., & Siroker, E. (2014). Systems and methods for website optimization. USA: USPTO. Retrieved from <http://patft.uspto.gov/netacgi/nph-Parser?Sect2=PTO1&Sect2=HITOFF&p=1&u=/netahtml/PTO/search-bool.html&r=1&f=G&l=50&d=PALL&RefSrch=yes&Query=PN/8839093>

- Tanz, J. (2014). How Airbnb and Lyft Finally Got Americans to Trust Each Other. Retrieved from <http://www.wired.com/2014/04/trust-in-the-share-economy/>
- Thiebes, S., Lins, S., & Basten, D. (2014). Gamifying Information Systems - A Synthesis of Gamification Mechanics and Dynamics. In *Proceedings of the Twenty Second European Conference on Information Systems* (pp. 1–17). Tel Aviv, Israel.
- Volkova, I. (2013). Four Pillars of Gamification. *Middle-East Journal of Scientific Research (Socio-Economic Sciences and Humanities)*, 13, 149–152. doi:10.5829/idosi.mejsr.2013.13.sesh.1427
- Warmelink, H. (2011). Towards a Playful Organization Ideal-type : Values of a Playful Organizational Culture.
- Webb, E. N. (2013). Gamification: When it works, when it doesn't. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8013 LNCS(PART 2), 608–614. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84880738442&partnerID=40&md5=4c73d0859a382e3e9997e3b8eaabe222>
- Wenzler, I. (2008). The Ten Commandments for Translating Simulation Results Into Real-life Performance. *Simulation & Gaming*, 40(1), 98–109. doi:10.1177/1046878107308077
- Werbach, K. (2014a). (Re)defining gamification - A process approach. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8462 LNCS, 266–272. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84900527633&partnerID=40&md5=b39e7df7309890a9f69dcc42ca275a06>
- Werbach, K. (2014b). Gamification [Coursera online course]. Retrieved from <https://www.coursera.org/course/gamification>
- Werbach, K., & Hunter, D. (2012). *For the Win: How Game Thinking Can Revolutionize Your Business*. Philadelphia, USA: Wharton Digital Press.
- Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, 27(8), -. doi:8
- Zhang, P. (2008). Motivational Affordances: Reasons for ICT Design and Use. *Communications of the ACM*, 51(11), 145–147. doi:10.1145/1400214.1400244
- Zichermann, G., & Cunningham, C. (2011). *Gamification by Design* (First.). Sebastopol, USA: O'Reilly Media.
- Zichermann, G., & Linder, J. (2013). *The Gamification Revolution: How Leaders Leverage Game Mechanics to Crush the Competition*. New York, USA: McGraw Hill.

## APPENDICES

<b>A</b>	<b>Relevant Empirical Gamification Research .....</b>	<b>88</b>
<b>B</b>	<b>Extensive Description of Gamification Design Method as Applied in this Research .....</b>	<b>90</b>
<b>C</b>	<b>Chronological Walkthrough of Ideation &amp; Iterative Prototyping Step in Design Method .....</b>	<b>93</b>
<b>D</b>	<b>Selling Assistant Details .....</b>	<b>113</b>
<b>E</b>	<b>Experiment Details .....</b>	<b>115</b>
<b>F</b>	<b>Excluded Metrics and Hypotheses in Experiment .....</b>	<b>118</b>
<b>G</b>	<b>Extensive Data Preparation .....</b>	<b>121</b>
<b>H</b>	<b>Extensive Data Validation .....</b>	<b>130</b>
<b>I</b>	<b>Justification of Statistical Distributions and Tests Used .....</b>	<b>137</b>
<b>J</b>	<b>Extensive Results of Statistical Tests .....</b>	<b>143</b>

## A RELEVANT EMPIRICAL GAMIFICATION RESEARCH

---

Descriptions of four well-executed empirical gamification studies according to Hamari, Koivisto and Sarsa as mentioned in their literature review (2014). See chapter 2.2.2 for the conclusions.

### A.1 Badges in an Online Marketplace

The first study was done by the same author as the literature review: Juho Hamari (2013). He investigated the effect of gamification on a utilitarian peer-to-peer trading service, quite comparable to OLX, with a field experiment in the existing service. He implemented a badge achievement system into the service, with a runtime of 1,5 years. The experiment setup was 2 x 2 design, with 3234 who were randomly assigned into groups. The two variables in this 2 x 2 design were the possibility to see other users' badges and the possibility to see which achievements unlocked which badges. Within this study, there is no specific gamification design method that is applied. Even though, in the same article, Hamari states: 'the mere addition of game elements does not necessarily guarantee successful gamification. However, in popular discussion the idea prevails that gamification simply refers to adding game mechanisms into a service, which in turn automatically becomes more engaging and attains a better retention of customers' (2013, p. 237). The author considers badges to be the primary and most used form of gamification and uses this as the single argument to implement badges as a game element. He did use a self-developed method to design the badges, named the 'badge game design pattern' (Hamari, 2013). Coincidentally, Hamari is also the main other of the literature review discussed at the start of this paragraph. So the acknowledgement of his own empirical study as 'well executed' should be treated with a grain of salt.

### A.2 Points and Reputation in an Enterprise Social Network

In the second empirical study, the authors tried to incentivise the internal social network of IBM, in order to stimulate user contributions (in terms of content). They identified the following incentive systems from a range of psychology research on participation motivations: rewards, explaining community benefit, goal-setting, reputation and providing self benefit. They chose to test a point-based incentive system (thus a combination of both rewards and reputation), based on the social nature of the service to enhance and the simplicity of the incentive system (Farzan et al., 2008a). The point system was implemented based on the content that was preferably added on the social network. For short-term content generation, it was found to be effective. However, the extra amount of content generated by the experimental group (compared to the control group) already declined after a week (Farzan et al., 2008a). After the first experiment, the authors decided to expand the point-based system with a reputation system, where users got a certain status based on the number of points they had gathered. This time, they also interviewed users from the experimental group in order to qualitatively extract their motivational changes, due to the incentive system that was implemented. Their long term goal was to gradually expand the incentive system with new features, testing the effect of every step, in order to create a sustainable and effective incentive mechanism for all users of the social network (Farzan et al., 2008b).

### A.3 Social Components in Gamification

The third empirical study concerns a statistical (meta-)analysis of the results of an online questionnaire, regarding the social motivations of people to use and continue using gamified services. A psychometric measurement model was created, incorporating the relations between social factors, which might influence one's attitude towards gamification. According to the results, the social factors that were identified are strong predictors for the perception of gamification and thus 'social elements are essential for creating engaging gamification services' (Hamari & Koivisto, 2013, p. 8). In

gamification design, it is important to allow social interaction and exposure between users, as well as having the users aligned and committed towards common goals (Hamari & Koivisto, 2013).

#### **A.4 Performance Feedback and Challenges in a Group Collaboration System**

The fourth study examines the effect of providing feedback and designing for optimal challenge on the contribution of each individual in a digital group collaboration system. The goal is to maximise this contribution, which in this case meant: to maximise both the quality and quantity of ideas generated in a messenger-like group collaboration system. The authors use the theory on motivational affordances and corresponding design principles for ICT systems by Zhang (2008) and recognise that ‘the design of the human-computer interface is an important determinant of a system’s motivational affordance’ (Jung et al., 2010, p. 727). Feedback and designing for optimal challenge were chosen as design principles to implement because of their simplicity and seemingly easy implementation (Jung et al., 2010). The evaluation of both the feedback and challenge added to the system was done with two 2x2 controlled experiments (experiment 1 with axes anonymous/pseudonyms and feedback/no feedback; experiment 2 with axes feedback/no feedback and explicit challenge set/general challenge set), with respectively 260 and 205 students. The students were asked to generate ideas on improving the parking situation at their university (Jung et al., 2010). The results were that performance feedback increased both the quantity and quality of the ideas generated. Setting an explicit challenge in combination with providing performance feedback was even more beneficial, while setting a challenge without feedback decreased the idea quantity and quality. Also, performance feedback turned out to increase performance in general, especially when users did not contribute anonymously. Thus, having a social component where performance can be compared is generally beneficial (Jung et al., 2010). This study is quite valuable, for many parallels can be drawn with the OLX case: a utilitarian peer-to-peer system with individual contributions of users who have a cognitive involvement, where the goal is to increase the quantity and quality of user-generated content.

## B EXTENSIVE DESCRIPTION OF GAMIFICATION DESIGN METHOD AS APPLIED IN THIS RESEARCH

In this appendix, an overview is given of the gamification design method for online marketplaces which was used in this research and which was adapted from the gameful design method by (Deterding, 2014c). Below, the original design steps and the **amendments** as applied in this research are listed. An overview of the steps is shown in Figure 6 below.

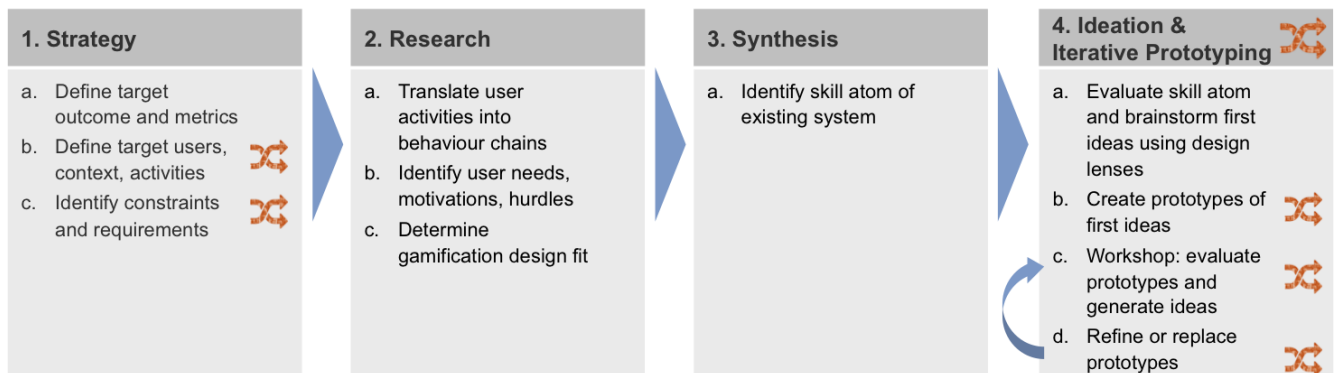


Figure 6 (copy): The gamification design method (Deterding, 2014b) with amendments, as applied in this research

### 1. Strategy

#### a. Define target outcome and metrics

The target outcome is the factor or factors that need to be changed in the online marketplace that will be gamified. They need to be quantified into metrics (or key performance indicators), preferably the smallest number of metrics possible. During evaluation, the success of the gamification treatment will be determined with these metrics.

#### b. Define target users, context, activities

This step is focused on data gathering, in terms of desk research, observation, historical data analysis or interviews. The goal is to distinguish the target audience within the system that can contribute to the target outcome, by influencing the metrics that were defined. **With this, as many contextual factors as possible should be included. Think of age, nationality and corresponding cultural dimensions, economical context, device used to visit the online marketplace, time spent on the online marketplace, number of users, individuals or groups, etc. This is important in order to actively think as the online marketplace users when designing gamification.** Also, for the target users, a list of activities should be made that they can perform and that will influence the metrics. This list should be prioritised by (1) the impact of activity on the metric and (2) the extent to which the addition of motivation and fun can help stimulate this activity.

#### c. Identify constraints and requirements

Constraints for the gamification design should be formed as hard delineations, while requirements can be seen as objectives, to which the design should comply as much as possible. Therefore, the constraints will mainly be based on technical and legal limitations and resources (time, budget, people) and the **requirements will mainly be based on general qualities that the users associate with the online marketplace and to which the gamification treatments should comply as much as possible.** More formally, this can be seen as a simple list of requirements (as in engineering design (Dym & Little, 2009)) which answers three questions: (1) what should the gamification do? (2) what should be the general characteristics of the gamification? (3) what are the technical/practical constraints to implement the gamification?

## 2. Research

### a. Translate user activities into behaviour chains

Based on the metric(s) and ranked activities that users can perform in order to influence the metric(s), behaviour chains can be created. These show the different activities of users, their interrelatedness and their influence on the metric(s), which can be either positive or negative.

### b. Identify user needs, motivations, hurdles

For each relevant activity in the behavioural chain, the designer determines the general need of the user to which this activities relates, the motivation(s) to perform the activity and the hurdle(s) that the user must take in order to perform the activity. This information can again be gained through desk research, interviews, historical data analysis, etc. and expands the existing behaviour chain. The goal of this step is for the designer to understand the broader psychological context in which the online marketplace users operate.

### c. Determine gamification design fit

With the system, user motivation and user background context drawn, the designer is able to determine whether gamification is the truly the right method to apply. In practice, this will have been decided earlier on in the process, but it is wise to perform another check at this point, to estimate the effectiveness of gamification treatments in general. The questions the designer should ask and which should be positively answered are originally based on work by Werbach and Hunter (2012, p. 49):

1. Does the activity connect to an actual user need?
2. Is lacking motivation a central issue or opportunity (and not e.g. poor usability)?
3. Does the target activity involve an inherent challenge with a learnable skill?
4. Is affording experiences of competence an effective and efficient way of improving motivation (and not e.g. defusing fears)?

## 3. Synthesis

### a. Identify skill atom of existing system

The designer will compose skill atom(s) out of the existing online marketplace, with help of the 'Lens of Intrinsic Skill Atom'. The starting point for this is the most opportune activity (as identified in step 1b) that users can perform on the online marketplace to influence the metric(s).

## 4. Ideation & Iterative Prototyping

### a. Evaluate skill atom and brainstorm first ideas using design lenses

For design lens, the designer needs to look at its applicability to fill in the gaps in the skill atom from step 3. If applicable, a first idea can be generated around this game element and possible other ideas can be listed.

### b. Create prototypes of first ideas

**Of the most promising first ideas that were generated in step a**, a prototype needs to be created that will serve as input in the workshop in the next step. Prototypes can have many shapes, both digital and physical, but a rule for these prototypes is that they should enable actual interacting and they should relieve or take away the core challenge in the skill atom (Deterding, 2014c).

### c. Workshop: evaluate prototypes and generate ideas

**In a workshop with the users of the online marketplace or participants that resemble these users, the prototypes will be evaluated and new ideas to either adapt or replace the prototypes can be generated.** This workshop can be set up in many ways, but there are some general guidelines are: introduce the concept of gamification, introduce the online marketplace and its skill atom, use as concrete questions as possible to evaluate the prototype and generate ideas, use small groups of people at the same time to allow discussion, split participants into teams, make sure there are resources for brainstorming (post-its, flip overs, markers, etc.), create an open



atmosphere, record the workshop for future reference (Deterding, 2014c; Ferrara, 2012; Werbach & Hunter, 2012).

d. Refine or replace prototypes

Based on the applicability to fit into the skill atom, the requirements and constraints, the metrics and all other contextual knowledge of the online marketplace and its users, the top ideas and comments should be considered for implementation into new or existing prototypes.

After this, the iterative cycle containing step 4c and 4d can be repeated as often as needed. The following iterative prototyping cycles will contain prototyping ideas, evaluating the prototypes in a workshop and adjusting the prototypes based on the workshop. Also, new ideas can be generated in a workshop and prototyped for the next cycle.

## C CHRONOLOGICAL WALKTHROUGH OF IDEATION & ITERATIVE PROTOTYPING STEP IN DESIGN METHOD

This appendix describes the entire ideation and iterative prototyping step (see step 4 of the gamification design method used, in appendix B) as executed in this research. With this description, it is easier to follow the choices that were made in order to end up at the final design of the treatment that was implemented on OLX for the controlled experiment. Also, it allows for a more enriched context for readers. The steps described in this chapter are laid out in figure X below.

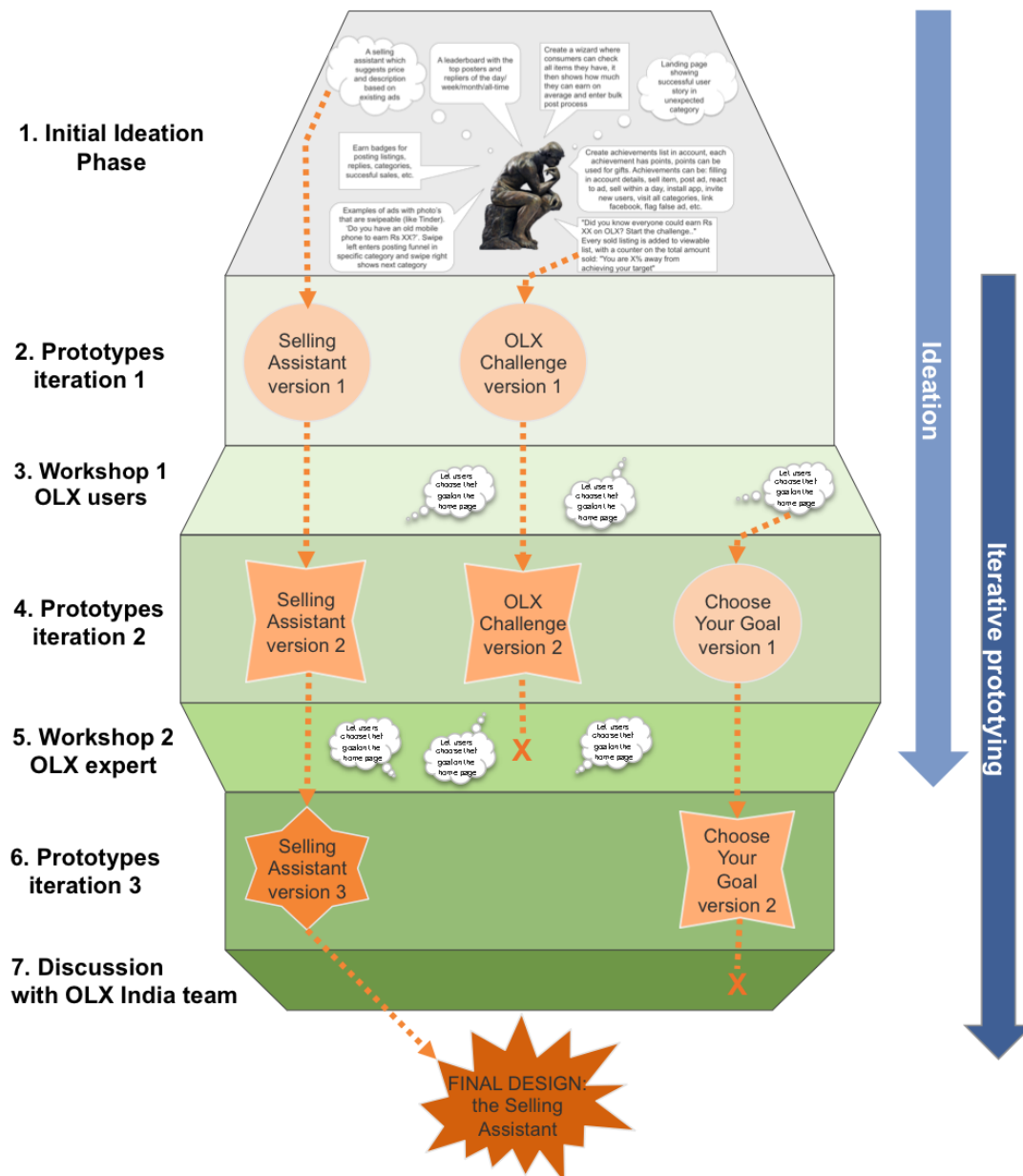


Figure 9 (copy): Ideation and iterative prototyping, step 4 of the gamification design method

## C.1 Initial Ideation

During the initial ideation, the following sources of input were (mostly implicitly) used to create initial ideas. In other words, these sources were of inspiration to the first ideas:

- personal communication with, amongst others, the following people:
  - CTO, Naspers Classifieds;
  - Senior Product Manager Mobile, OLX India;
  - Business Improvement Leader, Naspers Classifieds;
  - Eric Klaassen (Partner, BLOOM);
  - Jim Bijwaard (Consultant, BLOOM);
- gamification literature as mentioned in the bibliography;
- the game design lenses of the gamification design method used (Deterding, 2014c);
- various online gamification sources, mostly blog posts;
- video games, board games and other games that have been played;
- several courses on (serious) gaming at Delft University of Technology.

The resulting ideas were the following:

1. Telling consumers the average worth of unused products that an average person has in their house (based on OLX data): "Did you know everyone could earn \$XXX on OLX? Start the challenge.." Every listing they sell and remove adds up into a list that they can view, with a counter on the total amount of money sold: "You are X% away from achieving your target". This also stimulates removing sold listings.
2. Ask (new) consumers which unused products they have in their house (through a short checklist). Each product will be priced based on the average price of existing similar listings. At the end of the wizard consumers can see the potential amount they can earn on OLX: "You can earn \$XXX, start listing items now". Then they continue into the listing wizard with all the selected products and average prices already pre-filled.
3. Presenting (new) consumers a Tinder-like app, which shows them average listings in all types of categories with an average price: "Do you have an old lamp, to earn \$XXX?" "Do you have an old couch, to earn \$XXX" etc. Each listing can be swiped to 'yes' or 'no', where 'no' displays the next example and 'yes' continues to a pre-filled listing form.
4. Setting a goal for OLX users in a simple form on the landing page: sell X ads within X time and you will earn X amount of money (based on the average selling price of OLX ads in general).
5. Give feedback in every step of the posting funnel: at which step are you, how many steps left, well done, etc.

And more, see Figure 22 below.

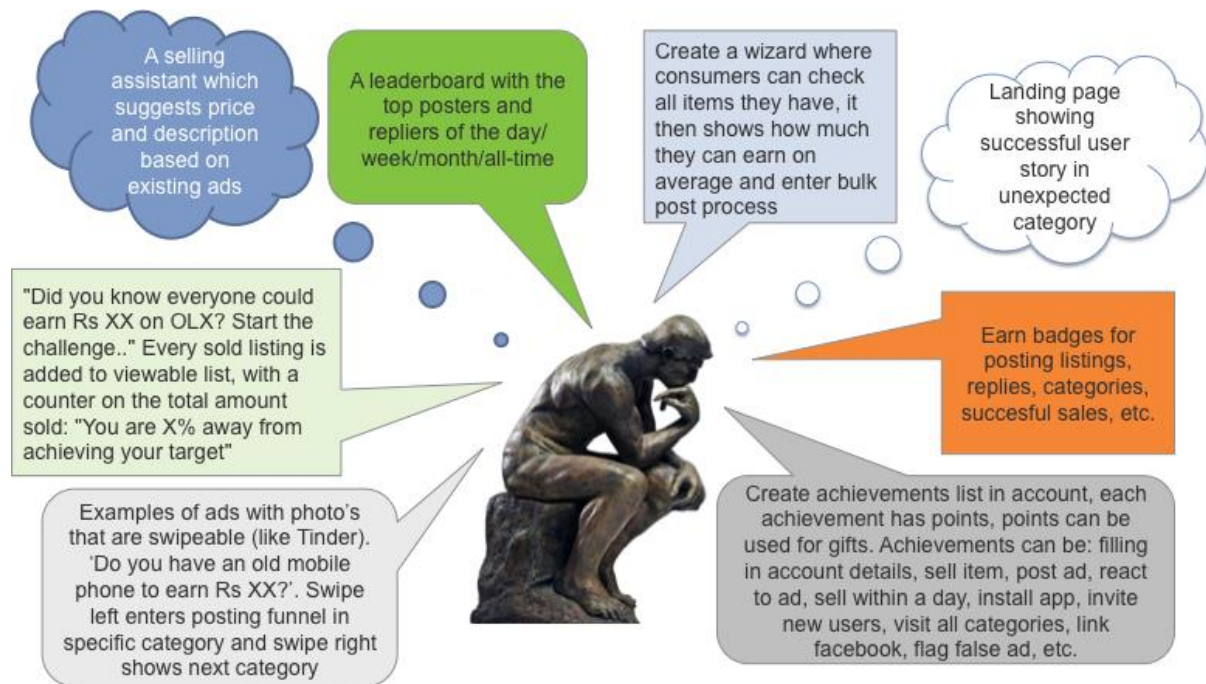


Figure 22: Results of initial ideation

## C.2 Mock-ups/Prototypes - Iteration 1

From the concepts and ideas, the following two were translated into a tangible form, namely a mock-up or even a clickable prototype.

### C.2.1 The Selling Assistant (prototype) – version 1

Ideas 2 and 3 (see the list in the previous subchapter C.1) were combined into one, which was named 'the Selling Assistant'. This is the concept that was the final result of the gamification design, as in: the gamification treatment that was tested on OLX with a controlled experiment. For an extensive description of this concept, see chapter 3.4.

Version 1 of the concept was a clickable prototype, featuring several items that were regularly sold on OLX. The idea was that showing users items that could be sold on OLX, including the average price that can be earned, is a method to invoke user to actually start selling this item. A few screenshots of the prototype can be seen in figure X below. The OLX homepage would feature a new button ('Don't know what to sell'), which redirects to a Selling Assistant, featuring one item and average price at a time. When a user would choose 'Yes', he/she would enter the posting funnel for this item and category. When a user would choose 'No', he/she would be shown the next item. This process could be enhanced by turning it into a swiping movement, making flipping through the items almost a game itself (much like the swiping movement in the popular dating-app 'Tinder').

The following game design lenses (Deterding, 2014c) were used as foundation for this concept:

- *Scaffolded complexity*: the concept of posting a listing is broken down to a less complex task, which is also the first task to perform when posting a listing – selecting an item to sell.
- *Appeal to motivations*: a motivation for users to sell their items on OLX is to earn money. By showing the average amount of money one could earn for each item, this motivation is directly being appealed to.
- *Limited choice*: Instead of selecting from a list of 11 main categories and then from a number of sub categories, users are shown one product at a time. Limiting the choice might also

make the choice less difficult. The categories are created by OLX and may not be self-evident for each user, while a product certainly is recognisable.

- *Templates*: providing a constrained set of products as starting point, including average prices, might partly take away the fear or inability for users to start from scratch with posting a listing.
- *Traces of others*: given the fact that each item has an average price that is shown, the existence of predecessors is suggested to each user, making it easier for new users to follow.
- *Sensual objects*: a swiping movement can make the activity more fun.
- *Onboarding*: the button on the homepage which leads to the 'Selling Assistant' says 'Don't know what to sell?'. This might create a strong want in the user to start, because not knowing which item to sell is one of the biggest challenges for OLX users.

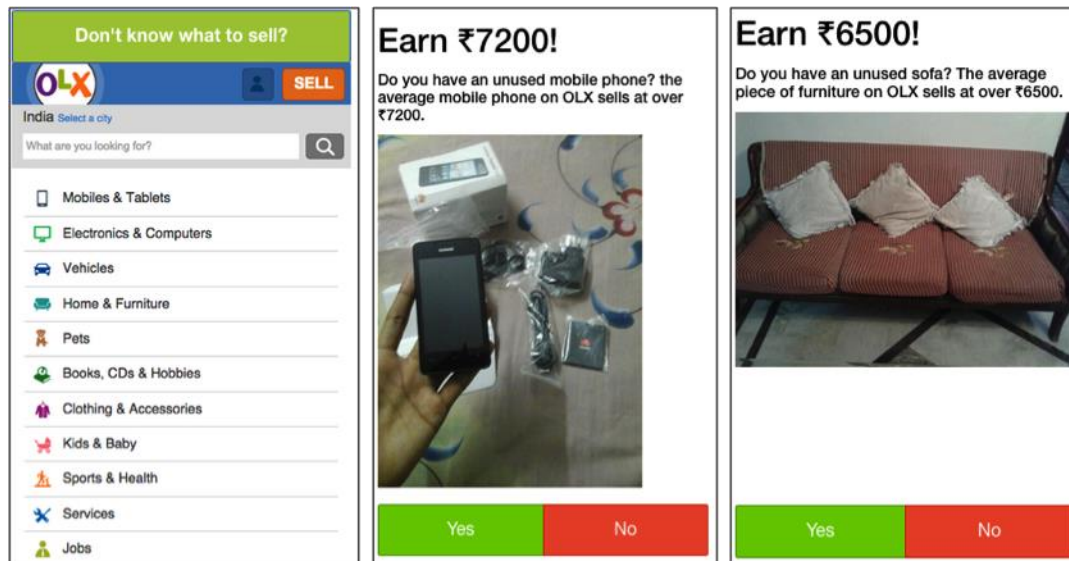


Figure 23: Selling Assistant prototype version 1

### C.2.2 The OLX Challenge (mock-up) - version 1

Idea 4 (see the list in the previous subchapter C.1) was translated to 'the OLX Challenge' and is quite simple. By introducing a challenge – including winning conditions, a time limit and an appealing and tangible reward – users might be willing to perform certain actions (post listings) in order to complete the challenge. The reward of the challenge is no external sum of money, but the amount that users actually earn with the listings they post. The line of reasoning is that users do not realise how much they can earn on average with a few listings. By making this explicit, it may feel as a reward for users, since it is a tangible sum of money that has not been explicitly tangible before.

The following game design lenses (Deterding, 2014c) were used as foundation for this concept:

- *Interim goals*: the challenge is split up into three parts, namely posting three listings.
- *Traces of others*: on average, other users earned an amount of XXX on OLX, so can you.
- *Appeal to motivations*: a motivation for users to sell their items on OLX is to earn money. By showing the average amount of money one could earn for each item, this motivation is directly being appealed to.
- *Graspable progress*: every step (posted listing) of the challenge is visualised after completing a step, so the progress within the challenge is made clear on further steps are implicitly suggested.

See a few screenshots of the mock-up in figure X below. The homepage (first and far left screen in figure X) would feature a pop-up, informing users of the new OLX challenge. By clicking the pop-up, users would go to the next page (second screen), which states the conditions and a more elaborate explanation of the OLX challenge. By clicking a button ('Start the challenge now'), users would enter

the normal posting funnel (third screen). After posting a listing, the normal 'thank you page' (fourth and far right screen) is extended with a message that informs users (1) how many percent of the challenge they have completed and (2) how much they can possibly earn on OLX based on the price of the listing they just posted.

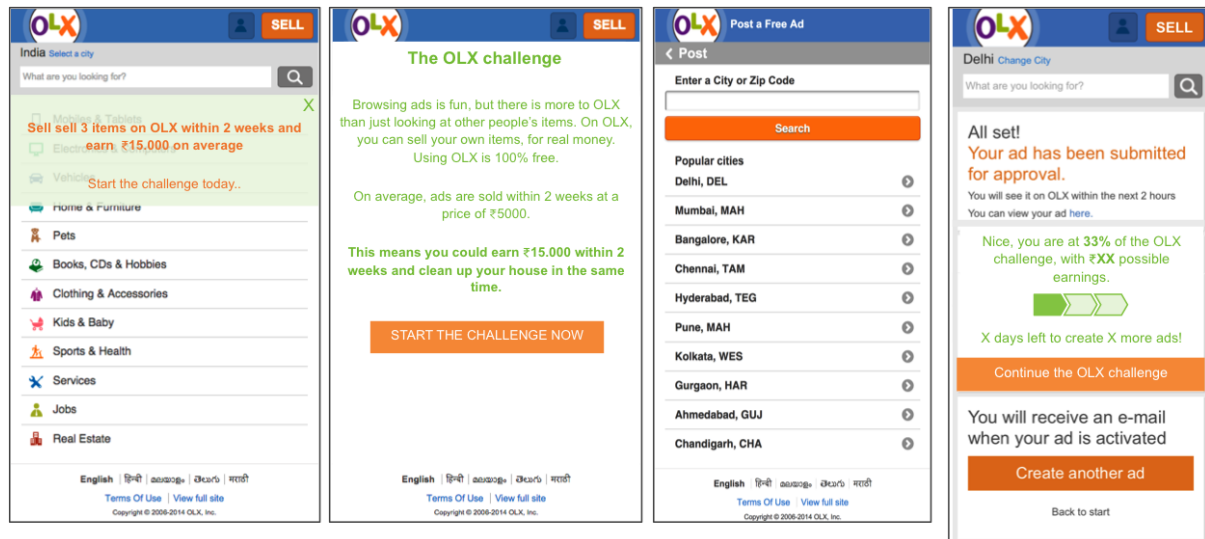


Figure 24: OLX challenge mock-up version 1

## C.3 Workshop 1 – OLX users (Indian students)

The version 1 mock-up of the 'OLX challenge' and version 1 prototype of the 'Selling Assistant' were tested on a panel of Indian students at Delft University of Technology on November 13<sup>th</sup>, 2014. The setup and results of this workshop are described here.

### C.3.1 Goal

There were four main goals in the workshop with the Indian students:

1. To evaluate prototype/mock-up of first ideas;
2. To generate new ideas based on game design lenses (Deterding, 2014c);
3. To identify possible hypotheses about segments for which the ideas and prototypes might or might not be effective;
4. To deepen contextual and interaction knowledge about users.

### C.3.2 Setup

#### Introduction round

At the start of the workshop, a word of welcome was given and the goals of the session were introduced, as well as the background and context of the research. Also, a short round was made, to identify the background of each participant. This way, some remarks can be made on the generalisability of the participants to the OLX India users. Questions included were:

- What's your name?
- What's your age?
- What's your study programme?
- Where in India do you come from?
- Do you know OLX? Have you ever used it? Posted replies or listings? Do you like it?
- Do you know similar platforms (Quikr, Marktplaats, Craigslist etc.)?
- What do you know about gamification?



## Presentation

After the introduction round, a short presentation was given. This covered an introduction of OLX and the layout of its mobile website, concluding in the participants having to post a listing through the OLX mobile website, namely the chair they were sitting on during the workshop. After this (which took around 5 minutes) the participants were asked to write down their first reactions, both positive and negative, to the posting process. They were given 3 minutes to do this and the following handholds to pay attention to, but not limit themselves, to these points:

- number of steps;
- clarity of fields to fill;
- biggest challenges;
- attractiveness to start posting an ad;
- feedback you received from OLX.

All participants were asked to give their feedback and this was summarised on a flip board.

Next, the OLX skill atom as developed with use of the design method by Deterding (2014b) was explained – see chapter 3.3.3. Also, an introduction to gamification was given, including some successful examples. The most important information was given to the participants on a hand-out: the OLX case (target users, metrics, skill atom) and some working materials from the gamification design method (Deterding, 2014b), including the design lenses and a skill atom worksheet.

## Evaluation of Pre-Developed Gamification Concepts

For the prototype of the ‘Selling Assistant’, one of the participants was asked forward to click through the prototype. Afterwards, all participants were given 3 minutes to individually write down their first reactions to the prototype, with the entire context in mind that they had been given in the presentation and hand-out (skill atom, motivation, fun, meaning, requirements, constraints, Indian user context). All participants were asked to give their feedback and this was discussed and summarised on a flip board.

Questions that were asked to deepen the knowledge of the participants’ names resembled:

- First reaction? Pro’s and cons?
- How does this motivate users to look at items in and around house to sell?
- How does this remind users of making money on OLX?
- What can be adapted to prototype to improve usability, connection to goals, needs and motivations of users, connection to OLX qualities, etc.
- In what other ways can we use <game design lens> (e.g. *juicy feedback*) to make users aware of what they can sell?
- In what other ways can we use <game design lens> (e.g. *scaffolded challenge*) to make users aware of the fact that they can earn money?
- Is there a way to use <game design lens> (e.g. *next best action*) on OLX to make the connection between items to sell and the money to make?

For the mock-up of the ‘OLX challenge’, this process was repeated.

## Ideation

For the last part of the workshop the participants were divided in groups of 2-3 people, based on their experience with OLX or similar platforms. This gave room for hypotheses on differences in effect between first time users and experienced users. The participants were asked to identify gaps in the developed skill atom (see chapter 3.3.3), using the design lenses, in order to come up with new ideas for possible gamification treatments. They were given 20 minutes to do this and afterwards they presented their ideas, following in a general discussion.



### C.3.3 Participants

The participants were Indian students, since they were the closest to OLX users that were available within the time and location (Netherlands) constraint of this research. Some of the participants were actual OLX users, as turned out during the workshop. A group of Indian students was invited for the workshop via Facebook. 5 participants were eventually included in the workshop. Groups are preferred over single persons with an interview and/or survey, because discussion and creativity are needed.

In the introduction round, the following data was collected about the backgrounds of the participants:

Participant	Age/ gender	Origin	Know OLX?	Posted on OLX?	Replied on OLX listing?	Know similar websites (Quikr, Ebay, Marktplaats)?	Know gamification?
1	24/m	Mumbai	Yes	Yes	Yes	Yes	Little bit
2	26/m	Mumbai	Yes	No	Yes	Yes	Heard of it
3	23/m	Surat	Yes	No	No	Yes	No
4	26/m	Pune	Yes	Yes	Yes	Yes	No
5	25/m	Chennai	Yes	Yes	Yes	Yes	No

**Table 8: Participant data, workshop 1**

### C.3.4 Documentation

In total, the workshop took around 2,5 hours. Materials used included a beamer, whiteboard, laptop, video camera, flipboard, pens/markers, post-its, game design lenses, skill atom worksheet, drawing paper and food/drinks (as gamified reward for participants).

The entire workshop was videotaped and photographed (see figure X below). Also, digital copies were made of the whiteboard and flipboards (which included the summarised opinions and input of the participants).



Figure 25: Workshop 1 photographs

### C.3.5 Results

#### Reactions to posting process in general:

- upload of pictures is slow;
- approval timeline and criteria of listing are not clear;
- email with approval and listing info is delayed;
- selecting city (first step in posting page) is slow, not responsive to text input;
- picture should be mandatory, because it tells you much more than a description;
- the description should be guided with a suggested text and content, or with more fields to fill, depending on selected category. Is also nice to have more data on this for OLX;
- selling something is not prominently shown on mobile web page - design and layout are more likely to invoke browsing categories rather than selling something;
- quite an easy process, because you are guided through the process very nicely through many small steps - however, having a lot of steps in the process is not necessarily positive.

#### Evaluation prototype 'Selling Assistant' version 1

- No pop-up and button not featured on top of the page. Should look like it's integrated into the OLX site and associated with OLX, otherwise people will think it's third-party advertising.
- Don't show the most popular categories first. Try to find categories of products that everyone has lying around in their house but that are not the first thing you would think of to sell. The popular categories are obvious and showing them to users does not give them any possible sell insights.

- Try to make pictures smaller, because the average phone is a feature phone (not smartphone) and uses 2G or even EDGE connection.
- Swiping would be better, nicer and more fun than pressing buttons, but this is complicated for the old mobile phones. Swiping has to work perfectly, otherwise it will only be a nuisance and negative influence.
- Do not use buttons with 'yes' and 'no', because it seems very definitive. Rather, just say 'next' or 'previous' and have an extra button or clickable picture to go into the posting funnel.
- Use logo's or graphical representations of items, rather than actual photo's. This is better for load times and also more generic. Pictures are specific products from other people and thus you are more inclined to say you do not have that item, because you actually don't. People will associate their own items faster to generic representations.
- If there is not swipe movement, or other way to invoke fun, it is better to put more categories into each step. This way, the number of steps for users to go through is smaller, which is better (because users have a short attention span and page loads are long).
- The average price should be communicated to attract people, money is a good motivator for OLX India users. However, the average price is not the price that should be the guideline for all objects in the category (people might sell very old sofa's for 5000 rupee if the average sofa price is 5000, even though the old sofa is worth below the average, just because this is the only guideline price they have). Using extra mandatory description fields, you might be able to suggest a price range.
- The order of categories should be changed every time you use it.
- Show the selling conversion of each category and have people select an item to sell based on this. Disadvantage: success to the successful.

#### Evaluation mock-up 'OLX challenge' version 1

- The challenge per se is not very attractive, there has to be an actual reward. For instance: money, free premium ad, extra percentage from OLX on top of the listing price (is susceptible for cheating however).
- The reward is actually selling your old items and getting money for it. However, if a user has not experienced this, it does not count as a reward yet.
- Watch out for 'pop-up feel'.
- 'Do the challenge' is not necessarily motivating. It has to be more specific.
- Progress bar or something more juicy is nice.

#### Ideation

Oriented on setting goals for OLX users towards selling rather than buying:

- Post X ads within X days and get free food coupons.
- Post X ads within X days and participate in a lottery to be featured in the new OLX TV commercial.
- Split users into 2 groups when they visit the site: 'do you want to buy or sell'? Choosing 'buy' will lead you directly to the buy page, while choosing 'sell' will lead you to a new page, which is focused on selling.
- Change text of 'sell' button to 'make money' and have don't know what to sell thing come in at second step of posting page (selecting a category).

#### General remarks

##### About OLX India

- Testimonials do not work, because in India they are not mostly not real and bought. In general, Indians do not really trust them.

- It is very difficult for online classifieds to establish in Indian market, because the unorganised sector is very big in selling and buying second hand items. In every Indian city or village, there are even people (kabadiwalas) who come by the houses – generally every week – to pick up items that people do not want anymore. They pay the material price and recycle the materials in the items (thus take them apart). They name the price and come by your house (thus much easier and personalised than using OLX). And if not, there is always someone around who knows someone who has a shop who and who will buy your unused goods. Many shops buy and sell second hand and much is also donated to family and friends.
- For buying, it is logical to use OLX, because you will find a large array of items that are not to be found in offline shops. When selling, you will only use OLX when you have a collector's item of which only certain people will acknowledge the true value, but this is a small user case.
- Smaller cities are good target, because they do not have the degree of access to second hand shops that big cities do provide.
- Everyone knows the OLX slogan (OLX pe bech de), it is an internet meme and many jokes are made with it.
- Try to make clear that old stuff could have value for other people and more value than selling it on the streets.
- A few years ago, when OLX was released, the participants in the workshop already tried OLX. However, back then they did not get any response to their listing or they could not find any interesting listings to buy. The success and number of listings should be confirmed on the website, to convince these type of users.

#### About gamifying OLX mobile web

- Easy differentiation between users who have successfully sold something on OLX on the one hand and users who have only posted without success, have only replied or are new to OLX on the other hand. For the first category, the reward (and the working and successful feeling of OLX) is already clear: selling your old items for a relatively high price. The latter category still needs to be convinced of some reward, which cannot just be the working of OLX, because they have not experienced it.
- Hard to implement reward-based gamification, because there is no transaction taking place on OLX.
- Probably the users who use the mobile website are either first time users who have not downloaded the app yet, or users who don't have a smartphone but a feature phone.

## **C.4 Mock-ups – Iteration 2**

### **C.4.1 The Selling Assistant – version 2**

All feedback of the participants in workshop 1 was taken into account, because technical malfunctioning and thereby failing effectiveness of the treatment would very negatively affect the results of the research. The feedback can be recognised in the modifications done, as can be seen a screenshot of the mock-up version 2 in figure X below. By clicking a green button on the home page (far left screen in figure X), users enter the Selling Assistant page (middle screen). In the Selling Assistant, users can select a product, which redirects them to the posting page for a new listing (right screen), with the category field prefilled based on the product they have chosen in the Selling Assistant. The posting page (right screen) is the original page and has not been modified in this treatment.

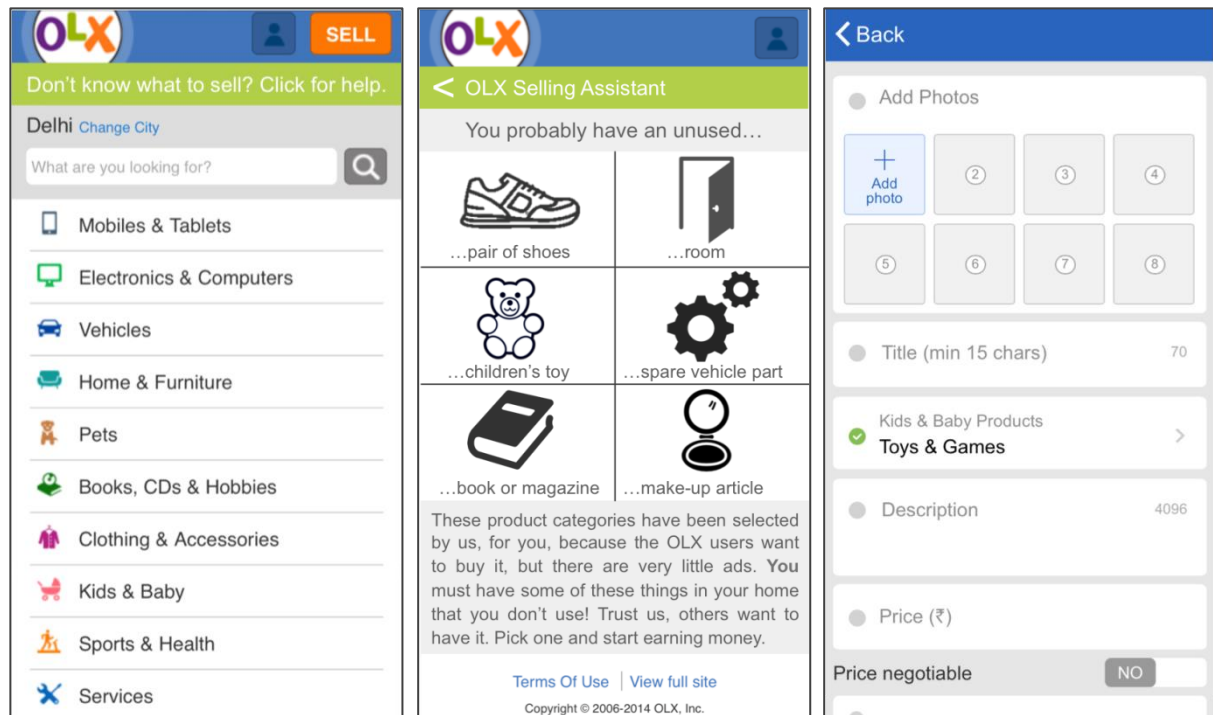


Figure 26: Selling Assistant mock-up version 2

### C.4.2 The OLX Challenge – version 2

Being featured in an OLX TV commercial was incorporated as the possible prize when completing the OLX challenge, based on the suggestions in the user workshop. Except for a few small layout changes, no more big changes were made to the mock-up of the OLX challenge. Below, in figure X, the changes from OLX Challenge version 1 to version 2 can be seen.

The main comments of the Indian students in workshop 1 were focused on the fact that a simple challenge with the reward of selling something on OLX would be no reward. Users who had not experienced the feeling of successfully selling something on OLX would not be engaged in the challenge without an extra (external) reward, according to them. The goal for workshop 2 was to find out if an OLX expert felt the same about this general principle.

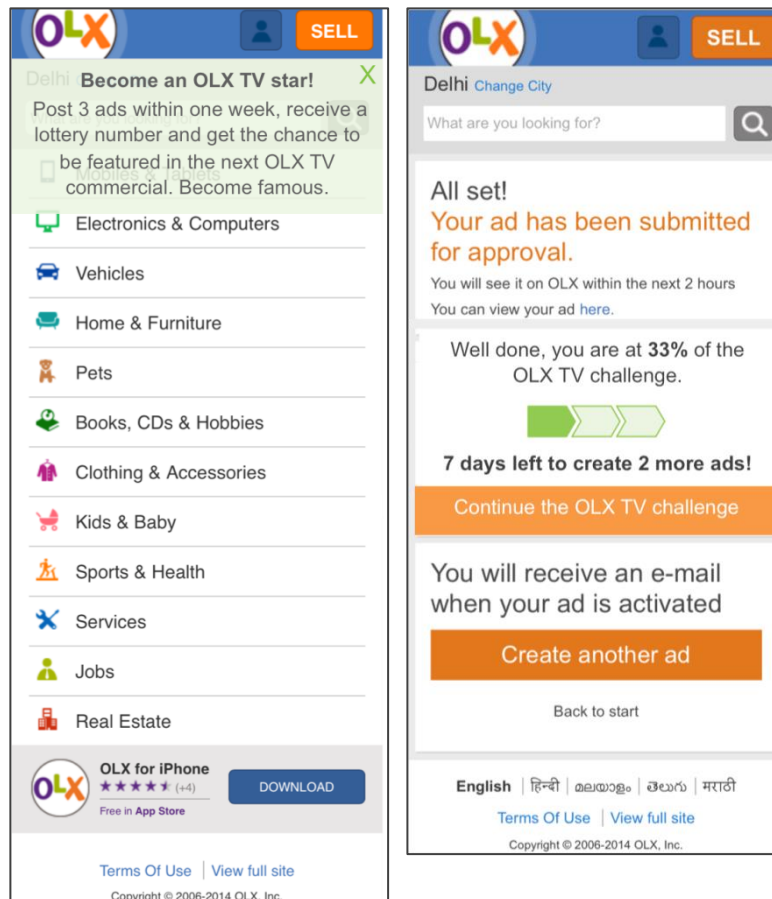


Figure 27: The OLX Challenge mock-up version 2

### C.4.3 Choose Your Goal - version 1

Based on the ideas from workshop 1, a new concept was introduced: 'Choose Your Goal'. Because much of the old OLX homepage was focused on getting OLX users to buy something (the majority of the links and content led to exploring existing ads), while the goal of the gamification treatments is to increase the number of new listings. The current landing page of the mobile website naturally redirects users into browsing categories or searching for a certain item, thus setting their implicit goal to buying rather than selling. The idea is to make clear on the homepage that users can both sell and buy on OLX. By translating these goals into two buttons on the home page, both buying and selling are equally featured as goals that OLX users can pursue. This concept tries to activate sellers by letting them select their 'path of destiny' or goal of their visit on OLX. Also, the landing page by itself should explain more about what OLX is, what you can do (both buying and selling) and have attractive on boarding. This will decrease the bounce rate and lead to more new listings in the long run.

The following game design lenses (Deterding, 2014c) were used as foundation for this concept:

- *Interim Goals*: The two buttons give the users a sense of freedom of which goal to pursue. The buttons function as explicit interim goals towards the core challenge and goal of a user.
- *Limited Choice*: only two choices for interaction are offered on the homepage, rather than the list of categories, the sell button, login button, search bar, location selector, etc.
- *Onboarding*: by introducing users to the basic concepts of interaction and functioning of OLX right from the beginning, they can have a better understanding of what OLX is and does and what they can do with it.
- *Next best action*: suggest a next best action to users, without taking away an interesting choice.



See figure X below for a screenshot of version 1 of the mock-up of the 'Choose Your Goal' concept. The featured screen is the OLX homepage, which would have two buttons. The orange button leads to the page to post a new listing (enter the posting funnel), while the gray button leads to the normal homepage, where the different categories with listings can be explored (buy-orientated).

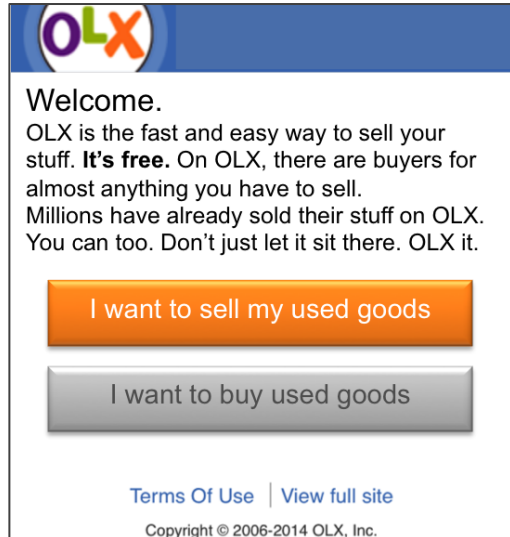


Figure 28: Choose Your Goal mockup version 1

## C.5 Workshop 2 – OLX expert

The version 2 mock-up of the OLX Challenge, version 2 mock-up of the Selling Assistant and version 1 mock-up of Choose your Goal were evaluated in a workshop with an OLX expert: the Business Improvement Leader at Naspers Classifieds. He has experience at different marketplace all over the world. The setup and results of this workshop are described here.

### C.5.1 Goal

There were four main goals in the workshop with the OLX expert:

1. To evaluate iterated mockups of gamification ideas;
2. To generate new ideas based on game design lenses (Deterding, 2014c);
3. To identify possible hypotheses about segments for which the ideas and prototypes might or might not be effective;
4. To deepen knowledge on the OLX platform and its users.

### C.5.2 Setup

The setup of this workshop is similar to the workshop with students. Only now, at the beginning, only gamification was explained and no clarification was needed on OLX.

### C.5.3 Participants

There was one participant: the Business Improvement Leader at Naspers Classifieds. He has extensive knowledge on OLX and online marketplaces in general, through his experience at eBay, Markplaats and OLX. Also, his creativity helped with the generation of new ideas. No more experts were invited to the workshop, because the goal was to have an open and critical mind towards the ideas and mock-ups, without possible involvement of other stakes or group think.



### C.5.4 Documentation

In total, the workshop took around 1,5 hours. Materials used included a laptop, whiteboard, pens/markers, post-its, game design lenses, skill atom worksheet and drawing paper. Below, in figure X, a photograph of the workshop is shown.



Figure 29: Workshop 2 photograph

### C.5.5 Results

#### Evaluation mock-up Selling Assistant version 2

- Great idea, very relevant, fits mobile experience quite well and is appealing to the many users who struggle with the hurdle 'don't know what to sell'.
- New users will find this very helpful, if they do not know what they should sell.
- Show the number of interested buyers for each product, in order to underline the demand for a new listing in this product category. This can be distilled from the average number of page views in the category, or the number of replies to a listing in the category.
- Adding an average price to each product will make it even easier for users to decide to post, especially if the price is higher than they expected. Also, this takes away another barrier for users, namely choosing a price for their ad.
- Shorten the amount of explanatory text. Mobile phones are not suitable for long reads and you want to convince users to post a listing as fast as possible, otherwise they will be gone. Especially the average phone in India is not suitable for long reads. Explanation or a convincing message can also be put into an audio file that starts when the page is opened.
- The icons are very nice, but could be made juicier (*lens of juicy feedback*) by turning them into animated gif files that move. Gifs are small files, thus would not have a big impact on the speed and technical feasibility (given the internet connection and the phones in India). This would allow for instance to alternatively display the icon and an average price plus number of interested buyers.
- On the long term, this treatment can also be used to balance the supply and demand of the different OLX categories, by making the categories that are featured dynamic. The categories in which the ratio sellers/buyers is the lowest can be featured in the treatment.
- Instead of only featuring just a fixed list of categories that are appealing to almost all users (to get the highest chance of reaching a user with such unused item), the categories can also be dynamic and based on for instance:

- 2 categories that have recently been viewed by the user (bought a new couch, now get rid of the old one);
- 2 categories similar to a category in which the user has posted a listing (interested in CDs, here are more);
- 2 random categories.

#### Evaluation mock-up OLX Challenge version 2

- Being featured in a TV show is a very nice prize for the treatment in India. However, it is not easy to realise this prize within a short time frame for this research.
- The challenge is interesting for users and not too hard.
- The quality will probably suffer badly for listings that are posted in this challenge. The only requirement in the challenge is now a number of listings within a certain time frame, so users might game the challenge by posting fake listings.
- For OLX, this treatment is not very new or exciting. Similar treatments have been tested in other countries (Kenya, South Africa, Brazil) with for instance tablets as a prize. These types of treatments are regularly effective on the short term, but the users involved in the challenge do not keep on posting listings and using OLX on the long term.

#### Evaluation mock-up Choose your Goal version 1

- Something similar has been tested on OLX Thailand and this was chosen as the winner homepage, so was positive for the number of listings that were posted.
- This treatment is aimed at getting users to enter the website (thus decrease the bounce rate), rather than directly nudging them to post a listing. Naturally, by having more users who enter the website, this will statistically mean that more listing are also placed. However, this is more a long-term effect, than something that you can see within an A/B test.
- The introductory text on OLX is not suitable for mobile experiences.
- For new users who are not familiar with the concept of OLX, this might be hard to understand as homepage. If they do not grasp the consequences of selecting either one button or the other, they are not likely to continue. Maybe a list of the most recently posted ads, in a sort of gallery, can explain the OLX concept quickly. In other words: if new users see two buttons (sell/buy) and a list of second-hand items for sale with a price, they know instantly what OLX is.
- You could do a multivariate test with different variations of buttons and button texts.
- Add buttons for users who visit OLX with other goals than buying or selling. This also helps for OLX to find out what it is that users come to do when they visit the website. Extra buttons can be:
  - What is OLX? → show an explanation of the OLX concept
  - I have some time to kill → show the list of latest ads for people to explore
  - Inspire me what to sell → show the list of latest ads for people to explore

#### Ideation

- Something that makes users laugh and feel good always works quite well. There does not even have to be a business or informative component. For instance: placing a funny animated gif from an OLX TV commercial at the 'thank you' page after users have posted an ad.
- An indication of the number of deals made for which amount of money on OLX today. Underneath it a button that says: 'I want that too'.
- Relating to the kabadiwala explicitly (see comments from user workshop): 'The kabadiwala will offer you XXX Rupees for this sofa, while user X sold it on OLX for X% more'.
- Changing the home page into a thank you message. 'Thank you for visiting OLX' with an image of a happy OLX India manager with thumbs up. Beneath it a link 'click here to enter'.
- Ask users to set a personal selling, buying or exploring goal every X times they visit OLX and show this goal to them when they

### General remarks

- Don't worry about the experiment influence of commercial sellers, they are mostly active on the desktop website.
- Keep in mind the practical complications for OLX India mobile web users: sun on the screen, small screen full of scratches, slow internet connection, no smartphone but feature phone, etc.
- The Selling Assistant has the most potential for OLX and fits the research requirements.

## C.6 Mock-ups – Iteration 3

Following workshop 2 with the OLX expert, the 'OLX Challenge' treatment was removed from the shortlist of treatments. The 'Selling Assistant' and the 'Choose Your Goal' treatments were adjusted based on the comments.

Please note that the OLX mobile web page changed to a new platform just before this iteration 3. Therefore, the general layout of the web pages changed. The mock-ups were also adjusted accordingly.

### C.6.1 The Selling Assistant – version 3

The main changes that were made include removing much of the explanatory text and adding information on the average number of views and the average price per featured category/product. Screenshots of Selling Assistant mock-up version 3 can be seen below:

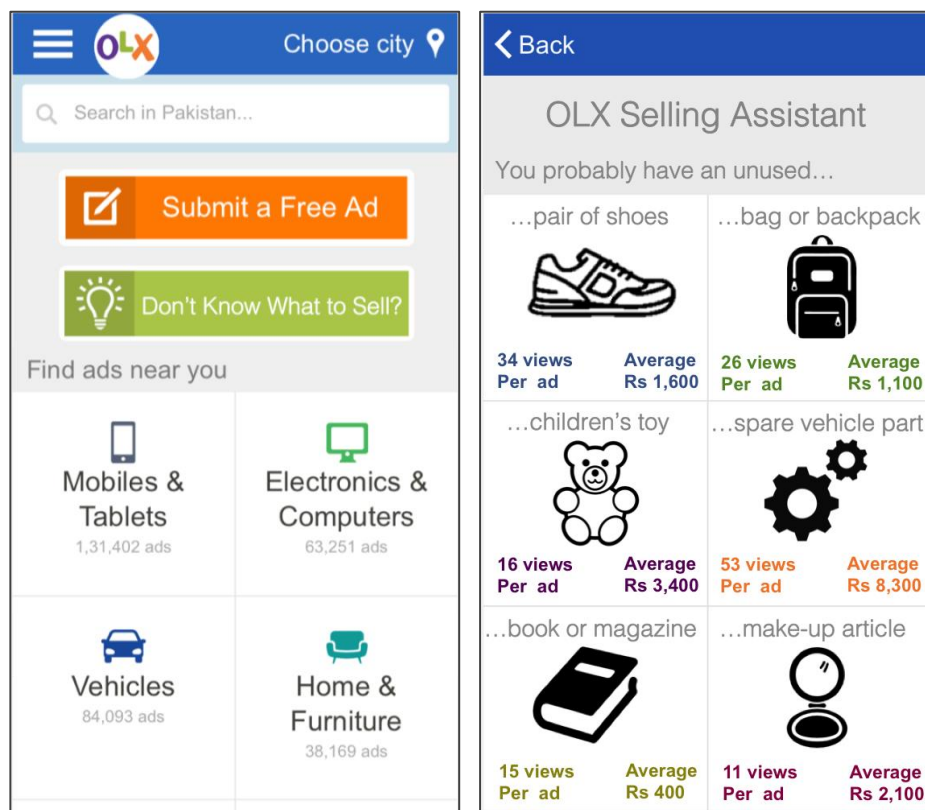


Figure 30: Selling Assistant mock-up version 3

### **C.6.2 Choose Your Goal – version 2**

The changes from version 1 to version2, based on workshop 2 with the OLX expert, can be seen in figure X below. From version 1, the explanatory text was removed and a list of the most recent listings posted was inserted, to replace the function of explaining the OLX concept to new users. The orange button leads to the posting page, the green button leads to the original homepage (where different categories with listings can be explored) and clicking a featured listing redirects users to the specific page of that listing.

Figure X shows is another possible variation, with some new buttons, which could be used to test what users come to do on OLX and to increase the chance of appealing to a users motivation for his/her visit. These were based on workshop 2 with the OLX expert.

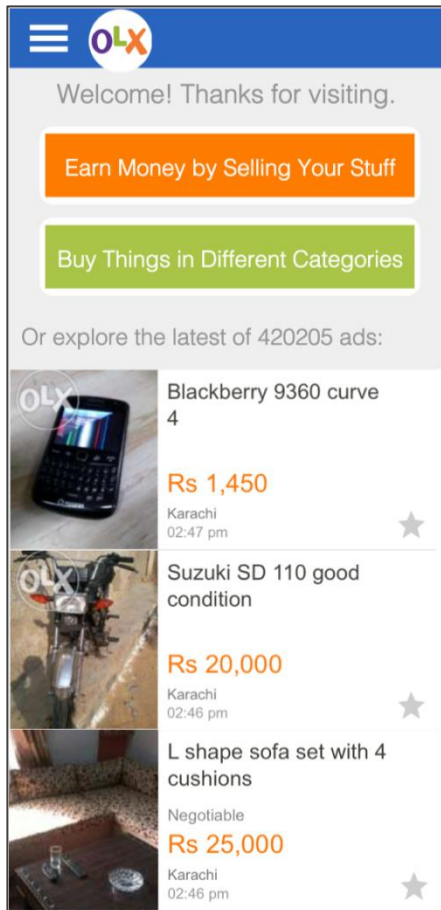


Figure 32: Choose Your Goal mockup version 2, variation A

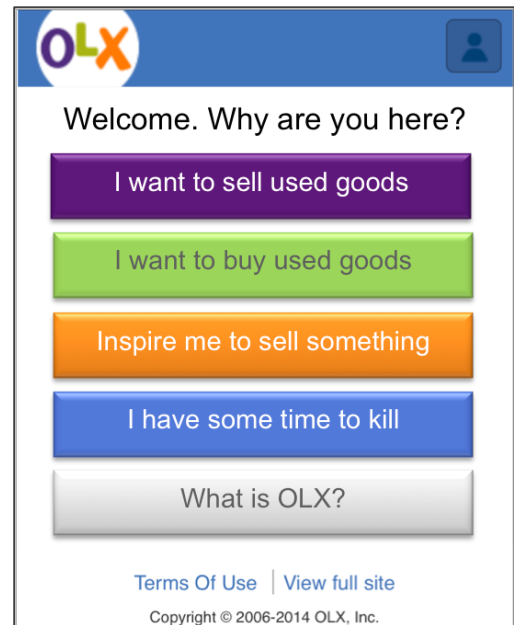


Figure 31: Choose Your Goal mockup version 2, variation B

## C.7 Discussion with OLX India team

The Selling Assistant mock-up version 3 and Choose Your Goal mock-up version 2 were presented to the OLX India team. During several meetings<sup>8</sup> the two treatments were evaluated.

### Evaluation mock-up Selling Assistant version 3

- There is too much info on the Selling Assistant page. The average price is less relevant, because it could result in users copying the average price for their ad, rather than thinking of an appropriate price themselves. So only the average number of views should be shown. This information cannot be dynamic, because that takes too much development time for this research. An average number of page views per category will be taken over the last few months and shown as static number.
- The average number of views per listing does not engage all users instantly, because this has no direct meaning to them. The text should be changed to 'interested buyers per ad'.
- The button leading to the Selling Assistant on the homepage should be less prominent, because the other button directly leads to the posting page, which is the starting point for users to post a new listing. So the main focus of the homepage should still be on that button.
- The colour scheme of the Selling Assistant page is too extreme and distracting.
- The categories featured in the Selling Assistant cannot rotate dynamically, due to restricted development time. However, it is unwanted for users to see the exact same categories with

<sup>8</sup> Personal communication with Senior Product Manager Mobile (OLX India) & UX Design Manager (OLX India). December 2014 & January 2015

same numbers over and over again. Adding a second Selling Assistant page with 6 other categories could solve this.

#### Evaluation mock-up Choose your Goal version 2

- The variation with many different buttons is too distracting for users. It is useful for OLX to know exactly how many users visit the mobile web page with which purpose, but this can be done with other means as well.
- This treatment will not contribute directly to the overall goal for this case study and research, namely using gamification to increase the number of new listings on OLX. It is interesting, but might fit better into another project.

#### Conclusion

Due to limited time and development resources for this project, only one treatment can be implemented into the mobile website and evaluated in an A/B test. The Selling Assistant fits the requirements and constraints (see sub chapter 3.3.1), probably contributes to the metrics (see chapter 3.3.1) and zeroes in on the main hurdle and discouragement for users (see sub chapter 3.3.3). The Choose Your Goal treatment does fit the requirements and constraints, but is less revolutionary, game-like and does not obviously directly contribute to the metrics defined and user motivations/needs/hurdles that were identified. Therefore, only the Selling Assistant was developed further and eventually tested.

## C.8 Mock-ups – Iteration 4

All comments from the discussion with the OLX India team were taken into account when developing mock-up version 3 of the Selling Assistant into mock-up version 4. See figure X below. The middle and right screen are the two variations of the Selling Assistant page, with different categories featured, which will rotate between user visits.

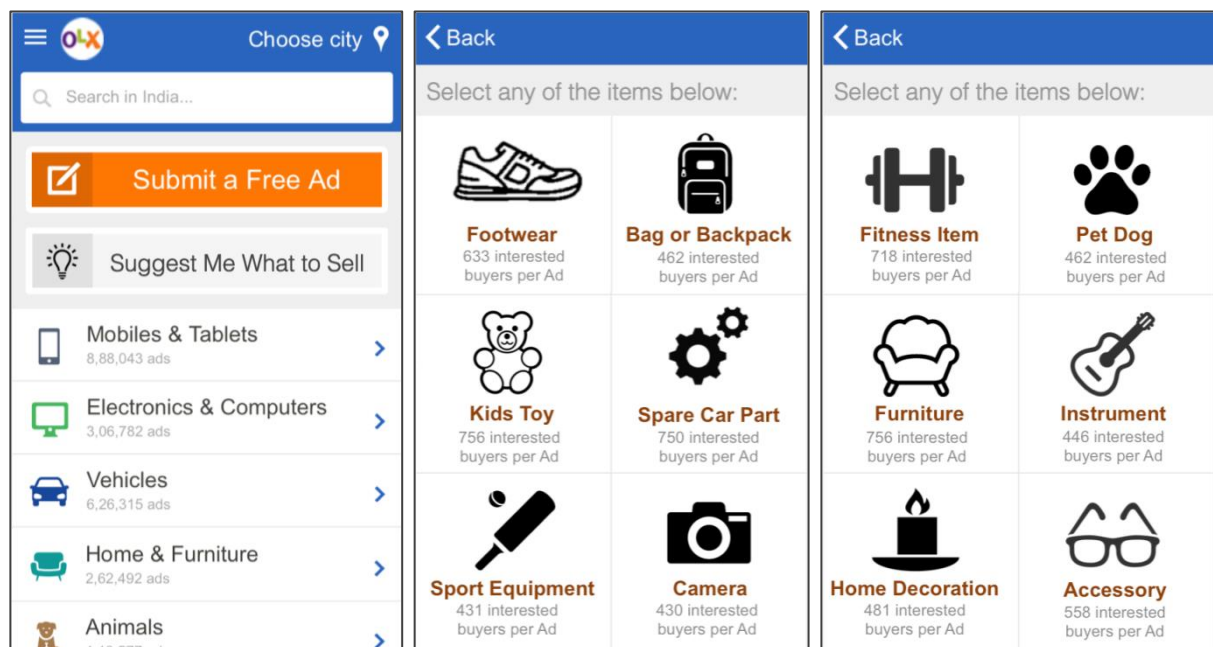


Figure 33: Selling Assistant mock-up version 4

The version 4 mock-up of the Selling Assistant was taken by the OLX designers and transferred into the OLX branding layout. Also, the featured categories were changed. In the mock-up, these are categories with a high ratio of page views per listing and with a high penetration amongst average users (e.g. everyone has an item in that category). This was changed to a mix of 12 categories, which had (1) the highest number of page views per listing and (2) accounted for at least 1% of the total

number of listings on OLX. This way, an increase in the number of listings in a category would actually have an effect on the platform in general and thus the metrics that were identified. If only very small categories are featured, the total potential of the treatment in terms of new listings per user is still relatively low.



## D SELLING ASSISTANT DETAILS

### D.1 Design Lenses Used







 <b>SCAFFOLDED COMPLEXITY</b> <p><i>To neither bore nor frustrate, good challenges grow with the user's skills.</i></p> <ul style="list-style-type: none"> <li>•How might you track the user's skill growth?</li> <li>•How can you increase the difficulty and choice of available goals and actions with the user's skill?</li> <li>•How might you turn »more of the same« into »different«, »more complex«?</li> <li>•How might you automate away what the user has mastered?</li> </ul>	 <b>APPEAL TO MOTIVATIONS</b> <p><i>Good feedback elicits the emotions and motivations that drive the activity.</i></p> <ul style="list-style-type: none"> <li>•What motivations and emotions driver your users to engage in your target activity?</li> <li>•How can you appeal to them in image, sound, text?</li> </ul>	 <b>ONBOARDING</b> <p><i>Good challenges make learning the system intentional part of the experience for newcomers.</i></p> <ul style="list-style-type: none"> <li>•How might you create a strong want in the user to start?</li> <li>•How might the user experience success in the first minute?</li> <li>•What is the core goal-action-feedback loop of your system? How might the user learn it by doing it in the first minute?</li> </ul>
 <b>LIMITED CHOICE</b> <p><i>Well-designed actions do not overwhelm users with too much choice.</i></p> <ul style="list-style-type: none"> <li>•How might we limit the number of actions offered?</li> <li>•How might we highlight important actions?</li> </ul>	 <b>INTERIM GOALS</b> <p><i>To support a sense of progress and direction, structure the user's path with clear interim goals.</i></p> <ul style="list-style-type: none"> <li>•How might you show the path to mastering the core challenge through explicit or implicit goals?</li> <li>•How might you break down over-sized tasks with interim goals?</li> <li>•How might you give the user a sense of freedom of choice what goal to pursue when?</li> </ul>	 <b>TEMPLATES</b> <p><i>Things to adapt are easy invitations that ease over the fear of the blank page.</i></p> <ul style="list-style-type: none"> <li>•Can you provide a construction manual?</li> <li>•Can you provide some randomly generated starting point?</li> <li>•Can you provide a constrained set as starting point?</li> <li>•Can you offer something half-done?</li> <li>•Can you offer something to adapt?</li> </ul>

Figure 34: Game design lenses used for the Selling Assistant (Deterding, 2014c)

## D.2 Full Experiment Workflow

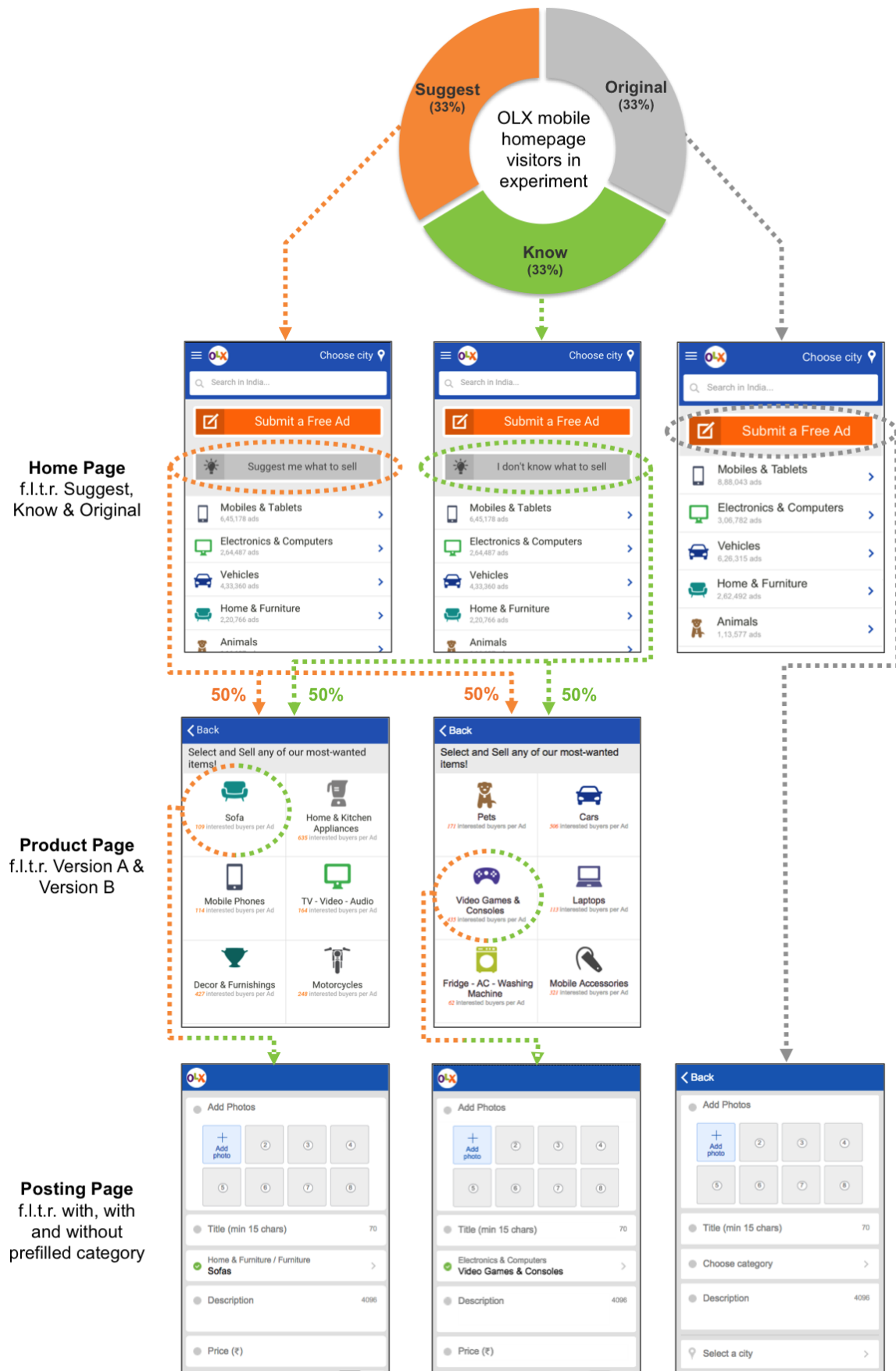
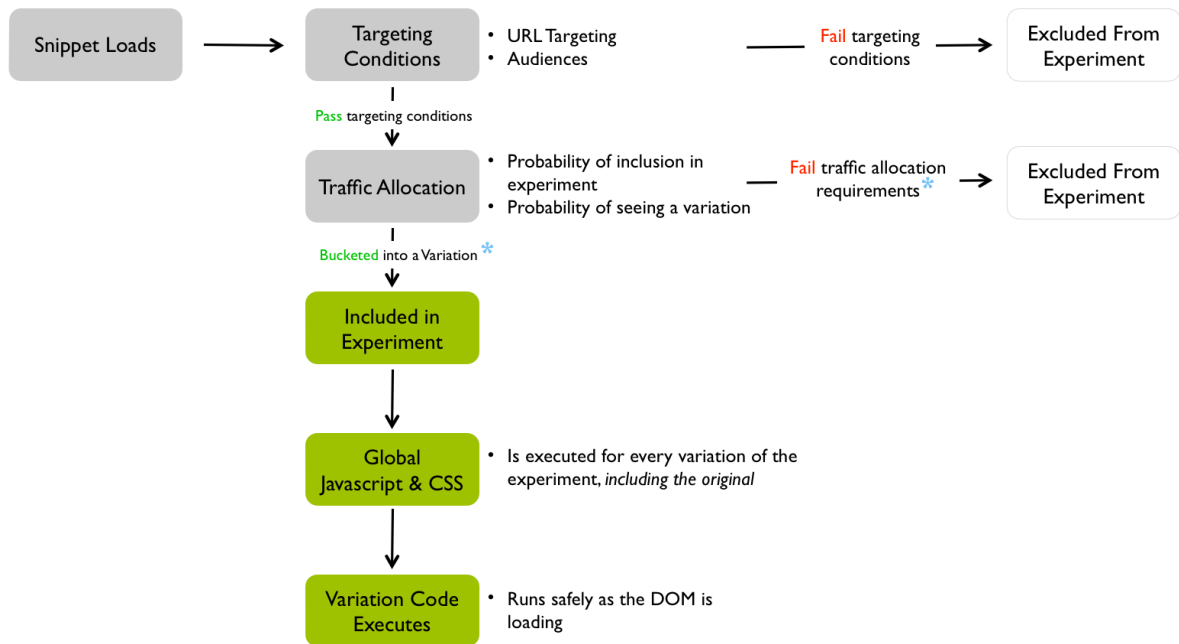


Figure 35: Full experiment workflow of the Selling Assistant

## E EXPERIMENT DETAILS

### E.1 Optimizely Traffic Allocation Scheme



\*User is cookie'd into this state. They will see this same version every time they revisit the site.

Figure 36: Sampling method of Optimizely, used in experiment, from <https://help.optimizely.com/hc/en-us/articles/200040335>

## E.2 Javascript Code For Specific Optimizely Data Collection

```

1 ▼ /* _optimizely_evaluate=force */
2
3 window.optimizely = window.optimizely || [];
4 jQuery( document ).ajaxSuccess(function( event, xhr, settings ) {
5     if (settings.url == "http://olx.in/i2/ajax/posting/") {
6         window.optimizely.push(['trackEvent', "posting_pv_i2_in"]);
7     }
8     if (settings.url.indexOf("posting/confirmmessage") != -1) {
9         window.optimizely.push(['trackEvent', "posting_sucess_i2_in"]);
10    }
11 });
12
13 ▼ (function() {
14     window.optimizely = window.optimizely || [];
15     if(document.cookie.indexOf('gamification_newsreturning') > -1) {
16         optimizely.push(['setDimensionValue', 'NewVsReturning', 'returning_visitor']);
17     }
18     else {
19         optimizely.push(['setDimensionValue', 'NewVsReturning', 'new_visitor']);
20     }
21
22 ▼ function setCookie(c_name,value,exdays,c_domain) {
23     c_domain = (typeof c_domain == "undefined") ? "" : "domain=" + c_domain + ";";
24     var exdate=new Date();
25     exdate.setDate(exdate.getDate() + exdays);
26     var c_value=escape(value) + ((exdays==null) ? "" : "; expires="+exdate.toUTCString());
27     document.cookie=c_name + "=" + c_value + ";" + c_domain + "path=/";
28 }
29
30 setCookie('gamification_newsreturning', 'true', 30);
31 })();
32
33 ▼ $( document ).ready(function() {
34     var city = jQuery.parseJSON(localStorage.optimizely_data).asyncInfo.location.city;
35     if (city == "NEWDELHI" || city == "CHENNAI" || city == "BENGALURU" || city == "PUNE" ||
36         city == "JAIPUR" || city == "CHANDIGARH" || city == "COIMBATORE" || city == "MUMBAI") {
37         optimizely.push(['setDimensionValue', 'location', city]);
38     } else {
39         optimizely.push(['setDimensionValue', 'location', 'other']);
40     }
41 });
42

```

```

42
43  /* _optimizely_evaluate=safe */
44
45  (function() {
46    window.optimizely = window.optimizely || [];
47    function checkPageExit(page, parentPage, cEvent) {
48      this.i2Mgr = i2.FrameManager;
49      this.parentFrame = '';
50      this.bindFlag = 0;
51      var Obj = this;
52      this.init = function() {
53        this.i2Mgr.$().on({
54          'framepushed' : function(e, frame) {
55            frame.$().on({
56              'frameready' : function() {
57                if(frame.params.action === page) {
58                  if(Obj.i2Mgr.parent() && Obj.i2Mgr.parent().params &&
59                     Obj.i2Mgr.parent().params.action){
60                    Obj.parentFrame = Obj.i2Mgr.parent().params.action;
61                  }
62                  if(!Obj.bindFlag) {
63                    Obj.checkExitRate();
64                  }
65                }
66              }
67            });
68          }
69        });
70      };
71    };
72    this.checkExitRate = function() {
73      this.i2Mgr.$().on({
74        'framepopped' : function(e, frame) {
75          if(frame.params.action === page && Obj.parentFrame === parentPage) {
76            window.optimizely.push(['trackEvent', cEvent]);
77          }
78        },
79      });
80      this.bindFlag = 1;
81    };
82    this.init();
83  }
84  var bRateSell = new checkPageExit('sell','index', 'exit_from_suggest');
85  var bRateAdding = new checkPageExit('adding','sell', 'exit_from_sell_posting');
86
87  })();
88

```

## **F EXCLUDED METRICS AND HYPOTHESES IN EXPERIMENT**

---

### **F.1 Metrics**

#### **F.1.1 Listing Quality**

The listing quality could be not measured, while this could be an interesting. For instance, the ratio between posted new listings (Gross New Listings) and accepted listings (Net New Listings) cannot be assessed, because the accepted listings could not be measured in the experiment. Also, the exact listings that were posted in the experiment cannot be individually determined, so there is also no manual sample inspection of listing quality possible, to assess potential differences between the Original and Suggest/Know variants. However, listing quality was especially a relevant issue if the gamification design entailed many extrinsic motivations, such as rewards in terms of money. This could have proposed a risk to the listing quality of listings posted through the gamification design. The Selling Assistant does not contain such rewards, so there is no logical expectation of a lower listing quality.

#### **F.1.2 Posting Page Visits**

Not possible to be used as metric, because data collected in experiment was not valid. This could have given more insight into the user-friendliness of the Selling Assistant, e.g. whether the consecutive steps are self-explanatory. It would be interesting to see if the Selling Assistant forms a better preparation for the posting page than not having a step in-between at all. Exit rates from the posting page between users who see the Selling Assistant and users who see the original website could be compared in order to check this.

### **F.2 User Segments**

#### **F.2.1 New and Returning Users**

Looking at new and returning user behaviour would be very interesting, but users cannot be qualified as this. The classifications of new and returning for users are limited to the experiment time, due to the unavailability of historical user data. In other words: a new user in the Original variant might be someone who has been using OLX for over a year and has sold and bought many items through the platform. He or she is classified as new in the experiment during the first visit and as returning for the second visit and possible visits afterwards. The expected differences between new and returning user behaviour - originated from the workshops and OLX internal research - are based on the overall characteristic of a user being new or returning and are not limited to the experiment time. The expected differences could therefore not be translated directly to testable hypotheses.

The hypotheses below are the ones originally set up after the user workshops and might be able to serve as input for further research.

For the Selling Assistant, one can assume that new users are more eager to press the button on the home page leading to the Selling Assistant, than a returning user. This is because new users often do not fully grasp the concept of the types of items that are sold or sellable on OLX. This was confirmed in the user workshop (see appendix C.5.5). Returning users already have some experience, thus may find the call to action on the button leading to the Selling Assistant of less help. Also, according to novelty effects, new features on a website usually yield higher conversions for the first period that

they are featured (R Kohavi et al., 2014). Since the difference between new and returning users has been confirmed by OLX CLM research before and there is an expected direction of the difference, we can state a one-tailed hypothesis here.

*Hypothesis X1. The percentage of new users that sees the Selling Assistant home page and presses the button leading to the Selling Assistant product page is lower than the percentage of returning users that sees the Selling Assistant home page and presses the button leading to the Selling Assistant product page.*

The difference between new and returning users when it comes to the number of new listings per active user is expected to be exactly the other way around. New users generally are not likely to post a new listing, because they are not familiar with the OLX platform yet and do not decide to post an item on the spot. Once they have seen what OLX is, they look for an item to sell and return to the site to fill in the details and post the listing (OLX, 2014a, 2014c). This is expected for users who interact with the original website and also for users who interact with the Selling Assistant.

*Hypothesis X2. Within the users who see the original mobile OLX website, new users post a lower number of new listings on average than returning users.*

*Hypothesis X3. Within the users who see the Selling Assistant, new users post a lower number of new listings on average than returning users.*

If hypothesis X1 holds and there is indeed a difference in the posting behaviour between new and returning users, hypothesis X1 can also be segmented for new and returning users. However, as hypothesis X1 is two-tailed, the segmented hypotheses also need to be two-tailed.

*Hypothesis X4. New users who see the Selling Assistant will post a different number of new listings on average than new users who see the original mobile OLX website.*

*Hypothesis X5. Returning users who see the Selling Assistant will post a different number of new listings on average than returning users who see the original mobile OLX website.*

## **F.2.2 Location**

During the user workshop, location (more specifically: city size) was identified as a factor that could determine the success of a gamification design and OLX in general (see appendix C.3.5). Especially small cities inhabitants were expected to make more use of OLX and post more listings, in comparison with inhabitants of large cities. Because of technical limitations and because of the small number of users visiting OLX from small cities (the focus of OLX in terms of customer acquisition and marketing lies on large Indian cities), small city users were not specifically registered. So tier 1, tier 2 and tier 3 cities cannot be compared. Only a selection of tier 1 cities (New Delhi, Mumbai, Bangalore, Chennai) can be compared with a small selection of tier 2 cities (Pune, Coimbatore, Jaipur and Chandigarh). Those are the cities that were registered, but they all have over 1 million inhabitants, so can barely be classified as 'small' cities with a low density of offline second hand shops. On top of that, mobile location data turns out to be quite inaccurate, because it often relies on network internet connection (which is area wide) rather than Wi-Fi. Moreover, it turned out that users post listings from multiple locations, so this cannot be genuinely incorporated.

Again, the original hypotheses are featured below.

Indian cities are classified into 3 Tiers, based on their population, infrastructure, education, living costs, political stability, etc. Between cities, there are notably differences in the OLX penetration, brand recognition and user posting behaviour (OLX, 2014a). The workshop with OLX users learned (see appendix C.5.5) that big cities have more offline locations for users to sell goods of relatively low



resale value, so the effect of the Selling Assistant on the number of new listings per active user might be bigger in small cities.

The following three segments are distinguished:

1. Tier 1 cities (Chennai, Delhi, Mumbai were registered during the experiment);
2. Tier 2 cities (Chandigarh, Pune, Coimbatore, Jaipur were registered during the experiment);
3. Other cities (unknown which cities exactly).

Based on this, two hypotheses can be identified for the location segments.

*Hypothesis Y1. Users from Tier 1 cities post a different number of new listings on average than users from Tier 2 cities.*

*Hypothesis Y2. Within the users who see the Selling Assistant, users from Tier 1 cities post a different number of new listings than users from Tier 2 cities.*

If both hypotheses turn out to be true, the following two hypotheses can be tested to dig further.

*Hypothesis Y3. Within the users from Tier 1 cities, users who see the Selling Assistant post a different number of new listings than users who see the original mobile OLX website.*

*Hypothesis Y4. Within the users from Tier 2 cities, users who see the Selling Assistant post a different number of new listings than users who see the original mobile OLX website.*

### **F.2.3 Listing Categories**

Differences in listing categories were expected, based on the different product page version. However, no data was collected to research this. The original reasoning is featured below.

For different OLX listing categories, the Selling Assistant is expected to have a different effect. After all, the Selling Assistant mainly stimulates the posting of new listings in the categories that are actually featured in the Selling Assistant. Therefore, the number of new listings per active users is expected to be higher within the categories that were featured in the Selling Assistant, as opposed to categories that were not featured.

## G EXTENSIVE DATA PREPARATION

---

### G.1 Filtering

#### G.1.1 Timeline and Sample Size Filter

Optimizely normally has a web interface that allows experiment owners to monitor the results. The raw data as extracted from the experiment was not fully in line with the web interface. The raw data was collected longer than the web interface showed. In order to align the two and to limit the experiment to exactly seven days, the final few logs were removed from the experiment. This is also because after this time, some adjustments were made to the experiment setup, which could result in false data. The experiment was cut from Monday February 23d, 16:00h (CET) until Monday March 2nd, 16:00h (CET). After this cut, the Optimizely web interface showed a number of 51052 unique users were included, while the raw data contains 51103 unique users. This is a difference of 0.1%, so negligible, meaning the raw data is aligned with the web interface in terms of experiment timeline.

#### G.1.2 User ID Numbers

The user ID numbers were visually inspected and matched with the logged IP addresses. For some user IDs, multiple IP addresses were logged, so it was assumed that the user IDs are leading in determining the unique users. After all, a user can access the website from multiple internet connections, with the same phone. Then, the IP addresses were deleted from the dataset. Also, the user ID numbers are random positive and negative numbers greater than  $1 \cdot 10^{18}$  or smaller than  $-1 \cdot 10^{18}$ . They are converted to string values by adding "id\_" in front of the values, otherwise they are misinterpreted by MS Excel and R Studio (because the numbers are too large for 32-bit integer vectors).

#### G.1.3 Duplicate Log Filter

Duplicate logs (rows) were removed: when all values for the entire row were the same as for another row. This was not true, because the timestamp went down to milliseconds. Therefore, if a page load was logged double, they still had a different timestamp. So a new column ('time\_short') was created with only the time in hh:mm:ss. Then duplicate logs were removed, meaning that all duplicate logs that happened in the same second are removed. This turned out not fully accurate, as a page load could be logged double, once at 23/02 15:05:01:943 and again at 23/02 15:05:02:012. So the 'time\_short' column was changed to format hh:mm:s, meaning that only every 10 seconds are shown. Then duplicate logs were removed. 10 seconds is a more credible range in which a user performs the exact same action. After this, the column 'time\_short' was removed again.

### G.2 Merging

The experiment resulted in two datasets, one concerning all users behaviour data from the 'home page experiment' and one from the 'product page experiment'. Both of them were filtered according to the procedure mentioned in the previous sub chapter. The fact that there were two datasets was due to implementation in Optimizely, which concerned two linked experiments with two variant allocation algorithms.

The home page data set contains all the posted listing logs that the product page data contains. According to the product dataset, 43 users posted 63 listings. These users were either in variant Know or Suggest, not in original. All of those users also did an event "click\_on\_suggest\_button" in home page data, so these are events posted through the product page and then posting page of one

of the Selling Assistant. These logs are also in the home dataset. 52 product users did 110 `posting_pv_i2_in`, an event indicating that the selling page has been seen. These events and users have all been logged in the home page. All their logged actions that are in the posting data is also included in the home page data. Of all the users who posted a normal listing according to the product data (23), 20 are also logged in the home page data and 3 are not. The 20 users all clicked the button to the product page in the Know or Suggest variant. The 3 users posted 3 listings. These 3 listings are negligible, since they are only  $3/554 = 0.5\%$  of the total number of listings recorded on the home page.

However, not all the users overall match 1-on-1 between the two datasets. Of the users in the home dataset, 765 users clicked on the button leading to the product page. Only these users should be included in the product dataset. However, there were 1110 users in the product dataset, of which 767 were also in the home dataset. Of the 765 users, 699 were in the product dataset and 66 were not. Why are 66 users not allocated but they did click the button? Individual inspections shows they do continue browsing after seeing the product page. They all did not post a listing through the product page.

Since the posting data is included and the most important logs are also included (new or returning, page loads, city, etc.), the most important thing left to do is to incorporate the product page variant to each home page user, since this is the only data point that is logged into the product page data but not into the home page data. This is done by selecting only the users in the product data who have clicked on the Selling Assistant button on the home page (which can be seen in the home data). The correctly assumed users who saw the product page and of whom a product page variant value should be added to the dataset are those who clicked in the suggest button in the home page. Only for those users it is possible to have been included into the product page. For these users, the variant id number for the product page (corresponding with Version A or Version B) is extracted and then only the unique users are selected. This resulted in a table with all user ID numbers of the 699 users who clicked the button on the home page and the product variant they were in. With this table, an extra column in the home data was created. For the 66 users who clicked the product page button according to the home dataset, but were not included in the product dataset, it was logged that they were on the product page, but it is unknown in which variant exactly. None of them posted a listing, so this will not influence the main analyses.

### G.3 Structuring

The whole dataset needed to be transformed in such a way that each row did not represent a page load log, but a user, because the experiment and the analyses are based on the fact that each experimental unit is a user. Because the values for the variables were collected per page load, some users had different values assigned to them for different page loads. For example, some users visited the site from both Mumbai and New Delhi or with both Chrome and Safari as browser during consecutive visits. In these cases the value, which occurred the most for that user, was assigned.

Also, some derived data variables could be added, which were useful for the analyses. Also, some redundant or unnecessary variables, which were automatically logged by Optimizely, were removed.

- *Time* was derived from the timestamp, with only the time included.
- *Day* was derived from the timestamp, with only the day included.
- *Post* was used in the original dataset (where each log was a page load). It was derived from the description of each page load URL. Listings through the product page contained '`posting_sucess_i2_in`' and were logged the same way in the post variable. Normal listings are those where the URL contains "posting/confirmmessage" and "added". This was verified by OLX.
- *Post sum* was derived from the post variable and counted the number of non-empty values for each user
- *Revenue* was empty, not needed and removed

- *Campaign* contained specific campaign names for OLX, not needed and removed.
- *Home variant ID* and *Product variant ID* were recoded from Optimizely ID numbers into their actual names (Suggest, Original, Know, Version A and Version B)
- *Product page* was derived from the URL description. If a user had a log which contained 'click\_on\_suggest\_button', a 'yes' would be logged as product page value, otherwise a 'no' would be logged as product page value.

## G.4 R Scripts Used for Data Preparation

```

1 setwd("~/Dropbox (Bloom)/Bloom team/Voor Bauke/Afstuderen/Data/Experiment 1")
2 require(data.table)
3 require(ggplot2)
4
5 #filtering duplicate logs in all datasets
6 DT = read.csv("./filtering/home_timefiltered.csv")
7
8 DTtime = data.table(DT)
9 DTtime = data.table(DTtime[, short_time := substr(DT$timestamp, 1, 18)])
10 DTtimeunique = DTtime[!duplicated(DTtime[,2:16, with=F]),]
11 write.csv(DTtimeunique, "./home_timefiltered_unique.csv")
12
13 DTp = read.csv("./filtering/product_timefiltered.csv")
14 DTptime = data.table(DTp)
15 DTptime = data.table(DTptime[, short_time := substr(DTp$time, 1, 18)])
16 DTptimeunique = DTptime[!duplicated(DTptime[,2:16, with=F]),]
17 write.csv(DTptimeunique, "./product_timefiltered_unique.csv")
18
19
20 #data from home page
21 DataHome = read.csv(
22   |   "./Data Preparation/4. Manual duplicate posts filtered/data_complete_final_v2.csv")
23 DTH = data.table(DataHome)
24 View(head(DTH))
25
26 #data from product page
27 DataProduct = read.csv(
28   |   "./filtering/product_timefiltered_unique.csv")
29 DTP = data.table(DataProduct)
30 View(head(DTP))
31
32 #identifying who was in wich dataset, if they posted, if they did or did not click
33 # to the product page and if this was recorded into both datsets or not
34 P_users = unique(DTP$user_id)
35 H_users = unique(DTH$user_id)
36 length(unique(DTP$user_id))
37 length(unique(DTP[homepageexperiment == ""]$user_id))
38 table(P_users %in% H_users)
39
40 a = unique(DTP[homepageexperiment == ""]$user_id)
41 table(a %in% H_users)
42

```

```

42
43 P_posting_users = unique(DTP[description == "posting_sucess_i2_in"]$user_id)
44 H_posting_users = unique(DTH[description == "posting_sucess_i2_in"]$user_id)
45 table(P_posting_users %in% H_posting_users)
46 length(unique(DTP[description == "posting_sucess_i2_in"]$user_id))
47 nrow(DTP[DTP$description == "posting_sucess_i2_in"])
48 length(unique(DTH[description == "posting_sucess_i2_in"]$user_id))
49 nrow(DTH[DTH$description == "posting_sucess_i2_in"])
50
51 length(unique(DTP[description == "click_on_suggest_button" & user_id %in%
52   P_posting_users]$user_id))
53 P_click_users = unique(DTP[description == "click_on_suggest_button"]$user_id)
54 H_click_users = unique(DTH[description == "click_on_suggest_button"]$user_id)
55 table (H_posting_users %in% H_click_users)
56 table (H_posting_users %in% H_click_users_wel_in_P)
57 table (H_posting_users %in% a)
58 table (H_click_users %in% P_users)
59
60 H_click_users_niet_in_P = setdiff(H_click_users, P_users)
61 H_click_users_wel_in_P = intersect(H_click_users, P_users)
62 b = intersect(a, H_users)
63 table (b %in% H_click_users_wel_in_P)
64
65 H_URL_users = unique(DTH[post == "URL"]$user_id)
66 P_URL_users = unique(DTP[post == "URL"]$user_id)
67 table (H_URL_users %in% P_users)
68 table (P_URL_users %in% H_URL_users)
69 table (P_URL_users %in% H_click_users)
70 setdiff(P_URL_users, H_URL_users)
71
72 P_posting_pv_users = unique(DTP[description == "posting_pv_i2_in"]$user_id)
73 table (P_posting_pv_users %in% H_users)
74 nrow(DTH[(user_id %in% P_posting_pv_users == "TRUE") & description ==
75   "posting_pv_i2_in"])
76 nrow(DTP[description == "posting_pv_i2_in"])
77
78 nrow(DTP[(user_id %in% P_posting_users == "TRUE")])
79 nrow(DTH[(user_id %in% P_posting_users == "TRUE")])
80 table(DTP[(user_id %in% P_posting_users == "TRUE")]$description %in% DTH[(user_id
81   %in% P_posting_users == "TRUE")]$description)
82 table(DTP[(user_id %in% P_posting_pv_users == "TRUE")]$description %in% DTH[(user_id
83   %in% P_posting_pv_users == "TRUE")]$description)
84
85 nrow(DTH[description == "exit_from_suggest"])
86 nrow(DTP[description == "exit_from_suggest"])

```

```

87
88 nrow(DTH[post %in% c("URL", "posting_sucess_i2_in")])
89
90 nrow(DTH[description == "exit_from_sell_posting"])
91 nrow(DTP[description == "exit_from_sell_posting"])
92
93 #making an overview of those users who were actually in a product page variant and
94 #also clicked to the product page
95 P_correct = DTP[user_id %in% H_click_users == "TRUE"]
96 length(unique(P_correct$user_id))
97 write.csv(P_correct, "./filtering/product_timefiltered_unique_correctusers.csv")
98
99 P_lookup = P_correct[, c("X", "revenue", "time", "description", "user_ip", "browser",
100   "source.type", "mobile.visitors", "campaign", "newsreturning", "location",
101   "location.segmentation", "productpageexperiment", "short_time") := NULL]
102 P_lookup = P_lookup[!duplicated(P_lookup[,1:2, with=F]),]
103 P_lookup = P_lookup[, c("homepageexperiment", "post") := NULL]
104 length(unique(P_lookup$user_id))
105 write.csv(P_lookup,
106   "./filtering/product_timefiltered_unique_correctusers_lookuptable.csv")
107
108 #merging the data from the product page and home page
109 DTall = merge(DTH, P_lookup, by = "user_id", all= TRUE)
110 write.csv(DTall, "./filtering/data_complete.csv")
111 length(unique(DTall$user_id))
112
113 #transforming all the data per page load to all data per user, and also for
114 #specific segments only
115 CompleteDataFinal = read.csv("./data_complete_final_v2.csv")
116 DTuser = data.table(CompleteDataFinal)
117 DTuser = DTuser[newsreturning == "returning_visitor"]
118 DTuser[, "post_count" := 0]
119 DTuser[ post %in% c("URL", "posting_sucess_i2_in"), post_count := 1]
120 sum(DTuser$post_count)
121 DTuser = DTuser[,.(max(variant_id_home), max(variant_id_product), .N,
122   post_sum = sum(post_count)), by = user_id]
123 write.csv(DTuser, "./data_per_user_returning.csv")
124
125 #checking for outliers
126 nrow(DTH[user_id == "id_-1020306231096922545"])
127 sum(DTuser$N)
128 sum(DTuser$post_sum)
129 length(DTuser$user_id)
130

```



```

130
131 DataUserVariant = read.csv(
132   "./Data Preparation/4. Manual duplicate posts filtered/data_per_user_variant.csv")
133 DUV = data.table(DataUserVariant)
134 CountTable = table(DUV)
135 boxplot(post_per_user ~ user_segment, data = DUV, main =
136   "Number of Posts per User Per Segment" )
137 p = ggplot(DUV, aes(user_segment, post_per_user))
138 p + geom_boxplot()
139
140 #adding a variable with the amount of posted listings per user
141 CompleteDataFinal = read.csv("./data_complete_final_v2.csv")
142 DTnopp = data.table(CompleteDataFinal)
143 DTnopp = DTnopp[!(user_id %in% H_click_users)]
144 length(unique(DTnopp$user_id))
145 DTnopp[, "post_count" := 0]
146 DTnopp[post %in% c("URL", "posting_sucess_i2_in"), post_count := 1]
147 sum(DTnopp$post_count)
148 DTnopp = DTnopp[,.(max(variant_id_home), max(variant_id_product), .N,
149   post_sum = sum(post_count)), by = user_id]
150 write.csv(DTnopp, "./H1/h2_post_by_users_who_did_not_visit_product_page.csv")
151
152 # Adding source type, browser, location, city tier, posting pv
153 CompleteDataFinal = read.csv(
154   "./Data Preparation/4. Manual duplicate posts filtered/data_complete_final_v2.csv")
155 DT = data.table(CompleteDataFinal)
156 DTz = DT
157 DTz[, "post_page_view" := 0]
158 DTz[post_pv %in% c("URL", "posting_pv_i2_in"), post_page_view := 1]
159 sum(DTz$post_page_view)
160 ▼ DTz = DTz[,.(max(variant_id_home), max(variant_id_product), max(browser),
161   max(source_type), max(location), max(city_tier), .N,
162   post_pv_sum = sum(post_page_view)), by = user_id]
163 View(DTz)
164 write.csv(DTz, "./listingsperuser/extra_variables_per_user.csv")
165

```

## G.5 Data Samples

In this sub chapter, data samples of the datasets per page load (page view) and per user are featured. These datasets were the results of the data preparation as described throughout this appendix.

### G.5.1 Data per User

Row 10 selected rows of the 51103 rows in total (excluding the variable names) are featured in Table 9 below

**Table 9: Data sample of experiment data, organised per user**

user_id	variant_id_ home	product_ page	variant_id_ product	post_ sum	page_ load_sum	browser	source	location
id_-1000178741835331583	suggest	no	none	0	3	safari	campaign	
id_-10153114168949320	know	yes	version b	0	40	safari	direct	
id_-1017564587089262331	original	no	none	0	7	gc	direct	MUMBAI
id_-1037336742191132417	know	yes	version a	0	8	safari	search	MUMBAI
id_-1048039267013908132	know	no	none	0	5	ucbrowser	referral	NEWDELHI
id_-1050224514046011315	know	no	none	0	32	ie	campaign	other
id_-1070939110976660655	suggest	yes	unknown	0	15	safari	direct	MUMBAI
id_-1070985791772810730	suggest	no	none	0	3	safari	referral	
id_-1071108401551248317	suggest	no	none	0	15	gc	referral	other
id_-1071273562369571808	know	no	none	0	2	gc	search	MUMBAI

## G.5.2 Data per Page Load

In Table 10 below, 10 selected rows of 575355 rows in total (excluding the variable names) are featured.

**Table 10: Data sample of experiment data, organised per page load**

timestamp	day	time	user_id	variant_id_home	variant_id_product	description	post	browser	source	newvsreturning	location
2015-02-23T15:05:05.535Z	23/02/15	15:05:05	id_-329999780502107408	suggest	none	http://olx.in/i2/		safari	referral	new_visitor	
2015-02-23T15:05:48.104Z	23/02/15	15:05:48	id_-6902023482032285940	know	none	engagement		safari	search	new_visitor	
2015-02-23T15:06:02.843Z	23/02/15	15:06:02	id_1214530041258516941	suggest	none	http://olx.in/i2/?category=1411			direct	new_visitor	
2015-02-23T16:37:43.649Z	23/02/15	16:37:43	id_-5610757710983378548	suggest	none	http://olx.in/i2/posting/configrmmessage/819268895/unlogged/added/	URL	gc	campaign	returning_visitor	
2015-02-23T16:37:46.645Z	23/02/15	16:37:46	id_-3571123000393989389	suggest	none	http://olx.in/i2/cars/q-pondicherry/?page=8			campaign	returning_visitor	MUMBAI
2015-02-23T17:44:38.768Z	23/02/15	17:44:38	id_-2825372486073478539	know	none	http://olx.in/i2/posting/configrmmessage/819283013/unlogged/added/	URL	gc	direct	returning_visitor	NEWDELHI
2015-02-23T18:44:21.174Z	23/02/15	18:44:21	id_4817133100839614567	suggest	version a	posting_sucess_i2_in	posting_sucess_i2_in		search	returning_visitor	other
2015-02-23T18:44:23.061Z	23/02/15	18:44:23	id_4205402957416249956	original	none	http://olx.in/i2/hyderabad/q-gorillaglass/?search%5Bregion_id [...]			direct	returning_visitor	
2015-02-23T18:44:24.194Z	23/02/15	18:44:24	id_4325040415327512574	know	none	http://olx.in/i2/item/2-bhk-flat-just-rs-6000-IDTqS1L.html#:b3804e740c			search	new_visitor	
2015-02-24T01:18:36.656Z	24/02/15	01:18:36	id_-7508089033728164863	know	version b	click_on_suggest_button		gc	direct	new_visitor	CHANDIGARH

## H EXTENSIVE DATA VALIDATION

Before testing actual hypotheses, it is wise to explore the data and validate if it corresponds with the experiment as implemented. This is done in three steps: (1) checking the different user segments and their overlaps and missing values, (2) checking for extreme and duplicate values for relevant variables and (3) fitting the control group (Original variant) to the OLX mobile user population.

### H.1 User Segment Overlaps and Missing Values

It is important to see if the various segments for users are mutually exclusive or not. For some, such as the product and home page variants, this should be the case, due to the experiment setup. For others, such as location and new/returning user, this does not have to be the case and needs to be inspected.

#### Home page variant

First and foremost, the distribution of the users and the page loads per variant is checked. This will confirm if the experiment group sizes have been implemented and executed correctly and if there is approximately the same number of logged user activities per variant. The users should be almost perfectly distributed, due to the experiment setup. The number of page loads can differ (because the users saw different websites), but should not be extreme. See Figure 37 below.

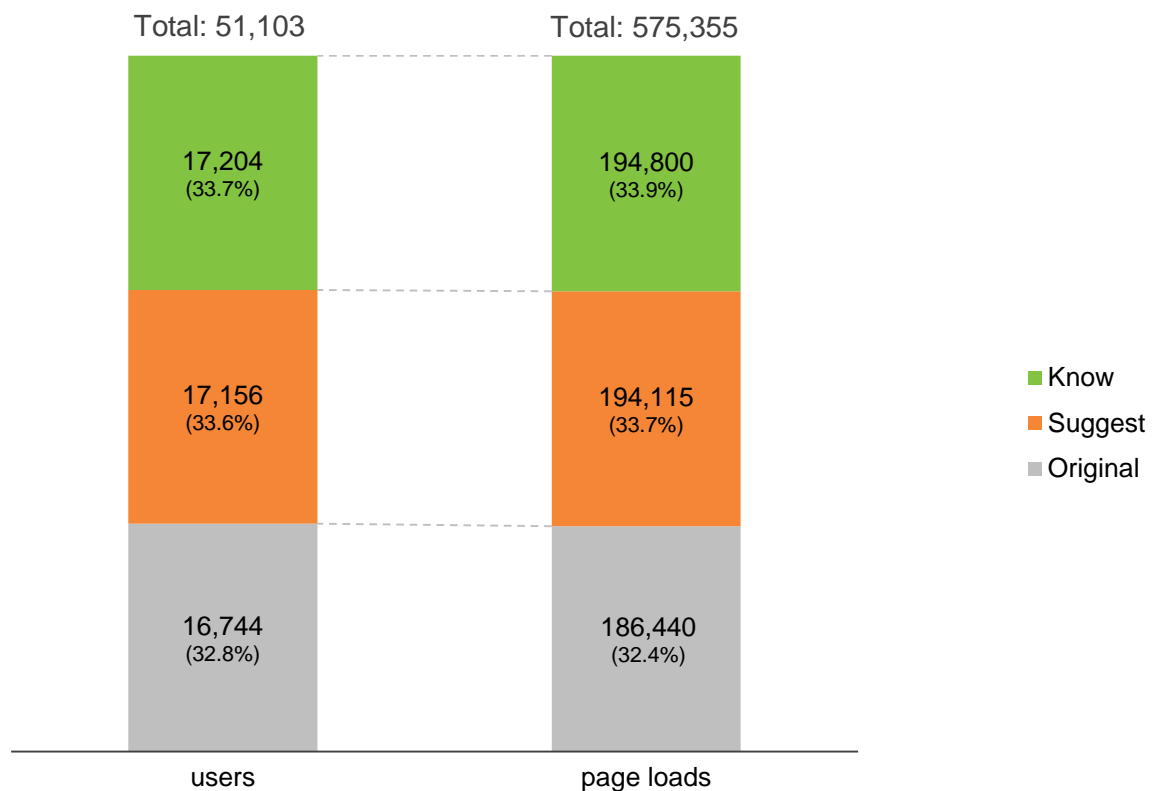


Figure 37: Home page variant user distribution

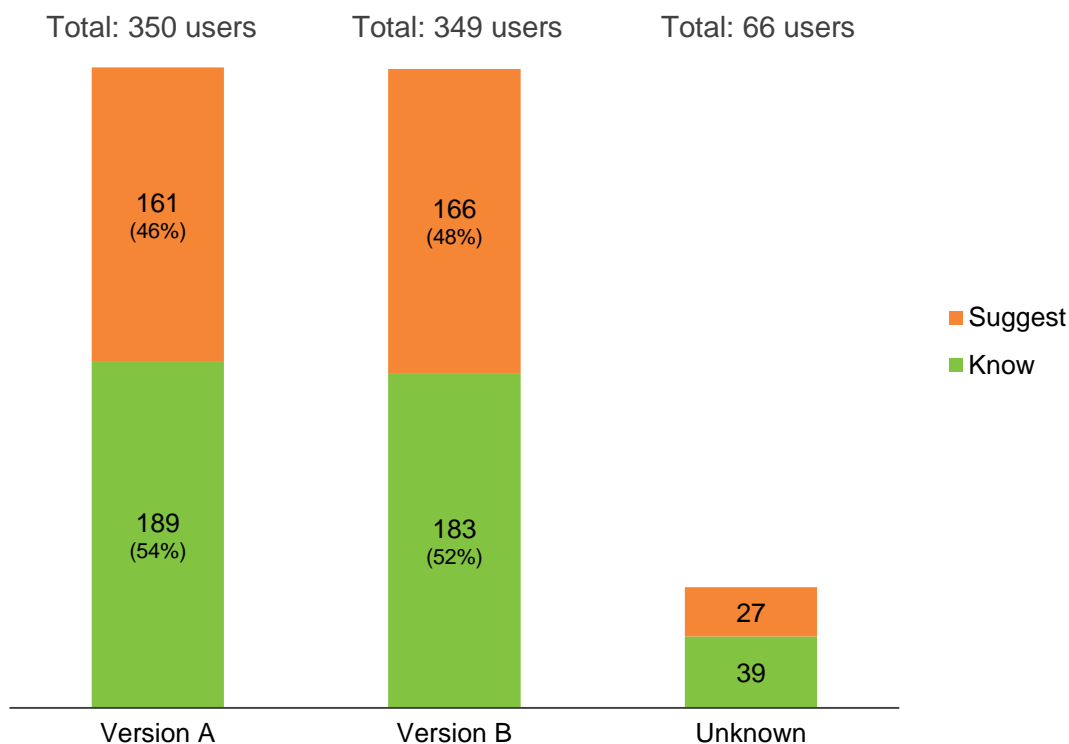
The distribution seems to be fairly even, although the Original variant is just under the Suggest and Know variants, both in terms of users and page loads. The imperfect user distribution can be possible due to Optimizely algorithms. The difference in page loads per variant could be due to the influence of the Selling Assistant in the Know and Suggest variants.

Also, it is important to check if the users in the experiment have only been assigned to one home page variant. This is true for Original: there are no users who were also included in Suggest or Know.

However, there is 1 user who has been included in both Suggest and Know. This is 1 out of >17000 so it is negligible. Probably, this is a user who cleared his/her cookies after being included in one variant and then was included into the experiment again (after another visit to OLX) and assigned to another variant. Inspection of this user's page loads shows that he/she was assigned to a new variant when returning to OLX on a new day. He/she was also depicted as a new visitor again, so it is likely that this cleared his/her cookies. There are no missing values for the home page variant variable.

### Product page variant

As for the product page variants (Version A and Version B), the distribution of users between them should be equal. Also, there is a number of users who clicked the Selling Assistant button on the home page, but was not allocated to a product page variant in the experiment.. These users did see a product page, but it is unknown which version. The distribution of users between the product page variants, from the Selling Assistant home page variants (Suggest and Know) can be seen below in figure 38.



**Figure 38: Product page variant user distribution, by home page variant**

As seen in figure 38, the total numbers of users allocated to Version A and B is almost equal – 350 and 349 respectively – which is correct according to the experiment setup. The number of users who clicked the Selling Assistant button on the home page but was not allocated to a product page variant is 66, which equals 8.6% of the total 765 users who clicked the Selling Assistant button in all variants. This is probably because they did not continue afterwards (non of these users posted a listing) and therefore Optimizely did not have the chance to log their product page allocation in a later page load log. The exact user behaviour and the difference between the Suggest and Know variants in terms of product page views is discussed in the results chapter.

A related check is whether users were assigned to multiple product page variants, while they are supposed to be assigned to only one. This is not the case: users only were in either Version A or Version B.

### New and returning

It is expected that all users were logged as new visitor and only a share was logged as returning visitors. This can be explained on a logical basis, since every user who is included into the experiment will have a first page load, being a new visitor. Then, after they close their browser or page and return the OLX mobile website, they are a new visitor. Not all users are expected to do this.

Of all the 51103 users in the experiment, 50868 have been logged as new user and 19239 have been logged as returning user. This equals respectively 99.5% and 37.6% of the total number of users in the experiment. Of the 0.5% of users who are never logged as new, 70 are only logged as returning user and the remaining 153 users are neither logged as new or returning visitor. This error margin, probably due to cookie deletion or user connection errors, of 0.5% can be negligible. The 37.6% of users who return to the website during the experiment week seems valid. After some individual user inspections it becomes clear that some users are logged as returning right at their second page load, even if this occurs only a few seconds after the first page load. This can happen if they close their browser in between those page loads and such measurement 'errors' are inevitable.

A single value for new or returning could not be allocated to each individual user, thus could not be included in the general analysis as a predicting variable for the number of new listings that were posted. However, when widening the definition of the experimental unit to visits rather than users, they can be assigned as new or returning. Based on this, some analysis is done later on in the results chapter.

### Browser

It is expected that some users use multiple browsers, but that most users stick with one. This is also related to the number of mobile phones they have and use to access the internet and the probability with which Optimizely can detect if this is in fact the same user (which is not accessible for this research). Apart from this fact, there could be several reasons for multiple browsers to be used by the same user in the experiment. The first reason is that users simply have multiple browsers installed on their mobile phone and use both of them alternately. Another reason is that users accessed the mobile website with multiple phones with different browsers. Another reason could be that they also used a PC or tablet to manually access the mobile website.

The browser overlap was tested for Safari and Google Chrome, since these are the most-used browsers in the experiment (90% of the total data logs). There were 25870 users who used Chrome and 20615 users who used Safari. Of the Safari users, 47 also used Google Chrome. So it seems only a small number of users has accessed OLX with multiple browsers during the experiment. All in all, it seems it is only a small percentage of users (~0.2%) who uses multiple browsers, so this should not affect the analysis.

When transforming the data from per page load to per user, the most occurring browser for each user was chosen and allocated to each individual user. There were 675 users (1.3%) for which no browser could be allocated.

### Location

It is likely that some users accessed the mobile website from multiple cities during the experiment, however, the great majority of users is expected to stay in one city. There are several reasons for users to connect from multiple cities. One could be that people move, which is a valid reason. It could also be because locations could not be accurately established with mobile internet connection, but

this seems less likely for such a big distance between Mumbai and New Delhi. This all depends on the exact location allocation algorithm at Optimizely, which is not accessible for this research.

The city overlaps for users were tested for New Delhi and Mumbai, since these are the most allocated locations in the data (63% of all the data logs with a location included New Delhi or Mumbai). There were 9662 users accessing the mobile website from New Delhi and 14280 users accessing the mobile website from Mumbai. Of the users from New Delhi, 377 also accessed the website from Mumbai. The effect on the analysis seems small still, because it includes only 3.9% of the inspected users from New Delhi also connected from Mumbai. However it is negligible, so needs to be thought in coming analyses.

When transforming the data from per page load to per user, the most occurring location for each user was chosen and allocated to each individual user. There were 13,996 users (27.4%) for which no location was detected. 11,028 users (21.6%) were not in one of the eight cities that were actually measured, so were assigned with value 'other'.

#### Source

All users could be assigned a source value. The distribution of 'campaign' (5,595 users), 'direct' (20,518), 'search' (20,001) and referral (4,989) are again based on the most occurring source value for each user.

## H.2 Extreme & Duplicate Values

An important thing to check is if any extreme values or outliers exist, since they can have an irrelatively large impact on the results. Also, samples of data were manually inspected for any remaining duplicate logs.

#### Listings per User

The number of listings per user was examined, for all users that were included in the experiment. In the figure below, the individual users are plotted sided by sided on the x-axis and the number of listings they posted is plotted on the y-axis. As one can see, most of the users posted 0 or 1 listings, as expected. There are few users who posted more (for example, people who are moving and have many things to sell at once). One user posted 29 listings in 1 week, which seems very unlikely. To verify this, the data logs (page loads) of the users with 4 or more posted listings were inspected.

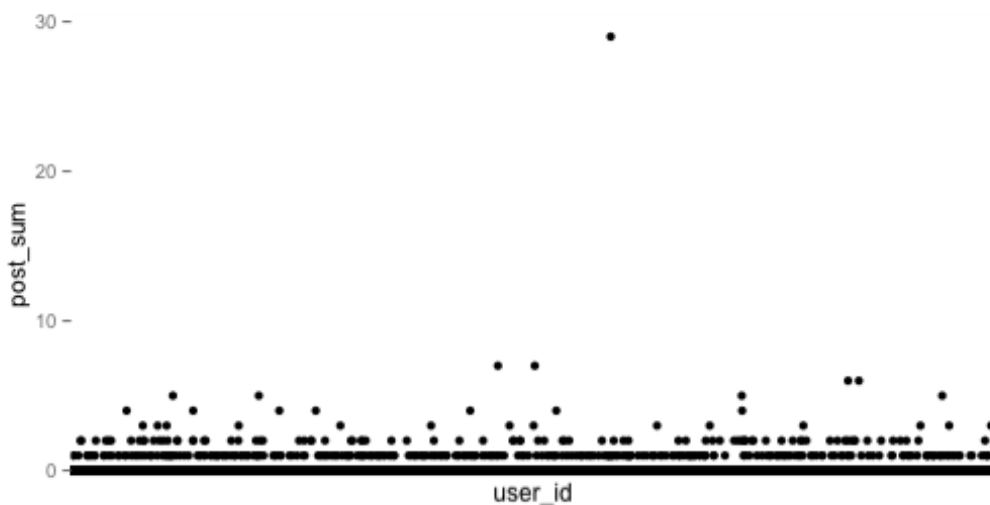


Figure 39: Number of listings posted per user during experiment



Listings posted in the normal way can be easily verified, since the confirmed posting URL contains a unique listing ID number. For postings through the Selling Assistant product pages, the URL was recoded to the name of the posting event ('posting\_success\_i2\_in'), so the listing ID numbers could not be identified. These logs were counted as duplicates when no other page was logged in-between (it is impossible to post second listing without visiting the posting page again) or when the posting logs had only a few seconds between one another (it is impossible to fill out the complete posting page and post a listing in a few seconds). All posting logs were manually examined: many of the users had posting URLs that were logged more than once, resulting in false logs of posted listings. 114 duplicate posting logs were removed; the first recorded log in time was retained.

The reasons for duplicate logs to exist are numerous, but there are two reasons most likely. The first is that the internet connection times out when the posting confirmation page (after the posting page has been completely filled out) is loaded. The obvious thing to do for a user is to reload the page. This can cause the posting confirmation to be logged twice into Optimizely. Another likely reason is that, after the posting process has been completed, the user shuts down his/her mobile browser. When he/she opens up the browser at a later point in time, the last page (the posting confirmation page) is reloaded, which also causes a duplicate log in Optimizely.

After removing duplicate logs manually, the boxplots in figure 40 below were generated, as well as the overview of user posting counts in table 11. Clearly, the median number of listings per users is 0. And still, there is one extreme outlier. This is the user with 29 posted listings, of which 2 were duplicate, so now 27 posted listings remain. Both within all the users and within the users in the Know variant, this user is an extreme outlier, but all the listings posted are valid, unique and posted throughout several days on different times. So, this outlier will not be automatically removed in the coming analyses. However, in some of the hypothesis tests the results with and without the outlier will be displayed, in order to make well-considered conclusions later on. When the outlier is not taken into account, the 27 listings posted will be transformed to 6, since this is the next highest number of listings in the entire dataset. This means that 21 listings can be considered as outlier listings, within the Know variant.

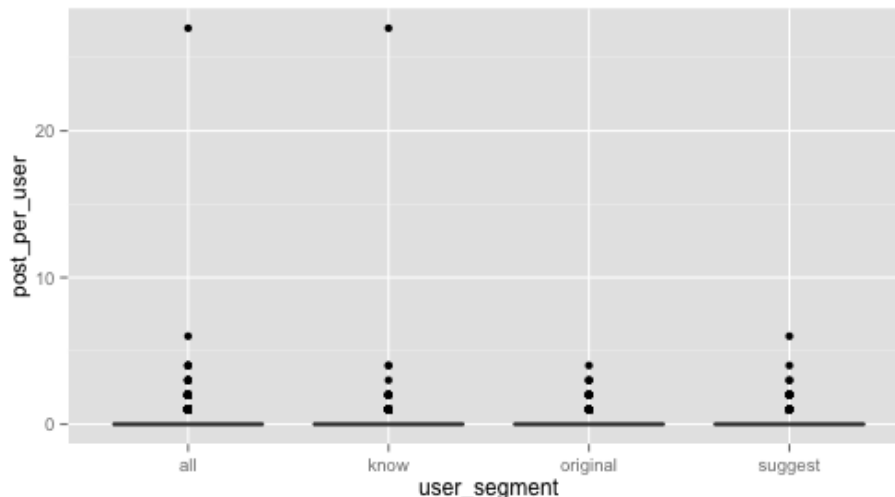


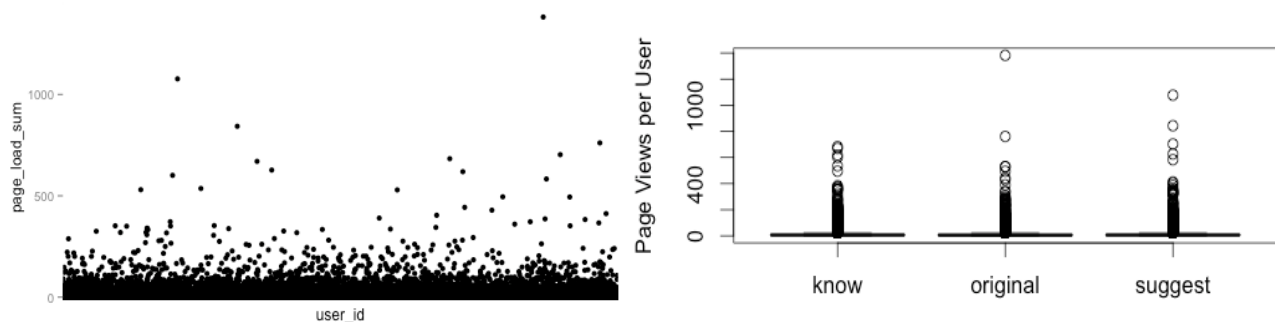
Figure 40: Number of listings per user, within different user segments

**Table 11: Number of users with a certain number of posts, within different user segments**

		Listings per User								
		0	1	2	3	4	6	...(7, 26)...	27	
User Segment	All	50687	372	32	6	4	1	0	1	
	Know	17056	133	10	1	2	0	0	1	
	Original	16620	110	10	3	1	0	0	0	
	Suggest	17011	129	12	2	1	1	0	0	

### Page Views per user

As explained earlier in sub chapter 4.2.4, a log is recorded into Optimizely for each page that a user (who is included in the experiment) loads on the OLX mobile website. Because of several reasons, some of which explained in the paragraph above, duplicate logs can occur. This has been the case with posting logs. Therefore the overall page loads per user were inspected. In the figure below, two graphs are shown. In the left graph, the individual users are plotted side by side on the x-axis and the number of pages they loaded (thus number of page they visited) on the y-axis. The right graph shows boxplots of the page views per user per variant. There are only a few users who visited more than 500 pages in the experiment week. All the logs of these users were manually inspected, to see if there were any abnormalities. Even though these users did have many duplicate logs, they all spent multiple hours during multiple days browsing through the OLX mobile site. Also, removing the duplicate logs automatically is not an option, since it is possible for users to return to the same page twice, so some duplicate logs would be removed unjustly. Every duplicate page load within 10 seconds has already been removed, so these types of double logs have been taken care of. Any duplicate page loads with a longer time in-between than 10 seconds should be manually inspected, along with the page loads before and after, to judge whether this was an actual user action or an accidental duplicate log. This takes several weeks and is not error-proof, so will not be done. Since the same policy has now been applied to all data logs, the relative difference can be assumed to be equal (assuming that the accidental duplicate logs are equally distributed over the users, but this cannot be checked).

**Figure 41: Number of pages loaded by users during experiment. Left: per user. Right: per user per variant.**

## H.3 Fit of Control Group to Population

Because no A/A test was performed (see explanation in 4.2.5), it is wise to benchmark the results of the Original variant with the general mobile website performance, during the same period, to check if they match. This cannot be done perfectly, because the users included in the Original variant are only users who landed on the home page and who were located in India during their visit (see 4.2.2). Also,

the (historical) population data comes from different OLX databases and is not completely aligned<sup>9</sup>. Therefore, the Original variant performance is benchmarked with all the mobile web data (not only the HTML5 i2 mobile website, but also the older versions that are still live) during the same period (23-02-15 until 02-03-15, the timeline of the experiment). Still, this comparison should give a general indication of the validity of the absolute metrics values that come out of the experiment data. Note: the relative comparison between the Original, Suggest and Know variants is not tested here.

The average number of new listings per user for the original variant differs 58% with the average number of new listings per user for the OLX mobile website overall, during the same period. For the i2 website only, this is 89%, but as explained above, the OLX i2 mobile web data is not entirely trustworthy. The conclusion from this data is that somehow the users in the experiment Original variant behaved differently than the normal OLX mobile web users. This could be due to the fact that they are a different user group, because they are only users who landed directly on the home page. Another reason for the difference shown in the table above is that not all page loads were collected correctly in Optimizely. Either way, there is no definitive reason to doubt the controlled randomised experiment and doubt the validity of the relative comparison between the Original, Suggest and Know variant, which is the main subject of this research. A better approach for this would be to have the actual population data of the new listings per user, in order to compare distributions, outliers, etc. Then a Chi-square goodness of fit test could also be conducted. However, this data is not available.

---

<sup>9</sup> Personal communication with and OLX internal data provided by Caspar Schönau (Director of Strategy, OLX). March, 2015.

## I JUSTIFICATION OF STATISTICAL DISTRIBUTIONS AND TESTS USED

### I.1 New-Listings-per-User Data

This type of data was used to test hypotheses 1, 2, ... and is commonly known as 'count data'. The number of events (new listings) per experiment unit (user) is counted. The complete new-listings-per-user dataset as result of the experiment, which was used to verify hypothesis 1 and 2, is analysed in this sub chapter.

#### I.1.1 Basic Descriptive Data

The graph in figure 42 below shows the number of new listings per user on the x-axis and the count of users with that number of new listings (frequency) on the y-axis. Please note the logarithmic scale of the y-axis and the fact that the outlying user in Know with 27 new listings was left out of this graph, for readability purposes.

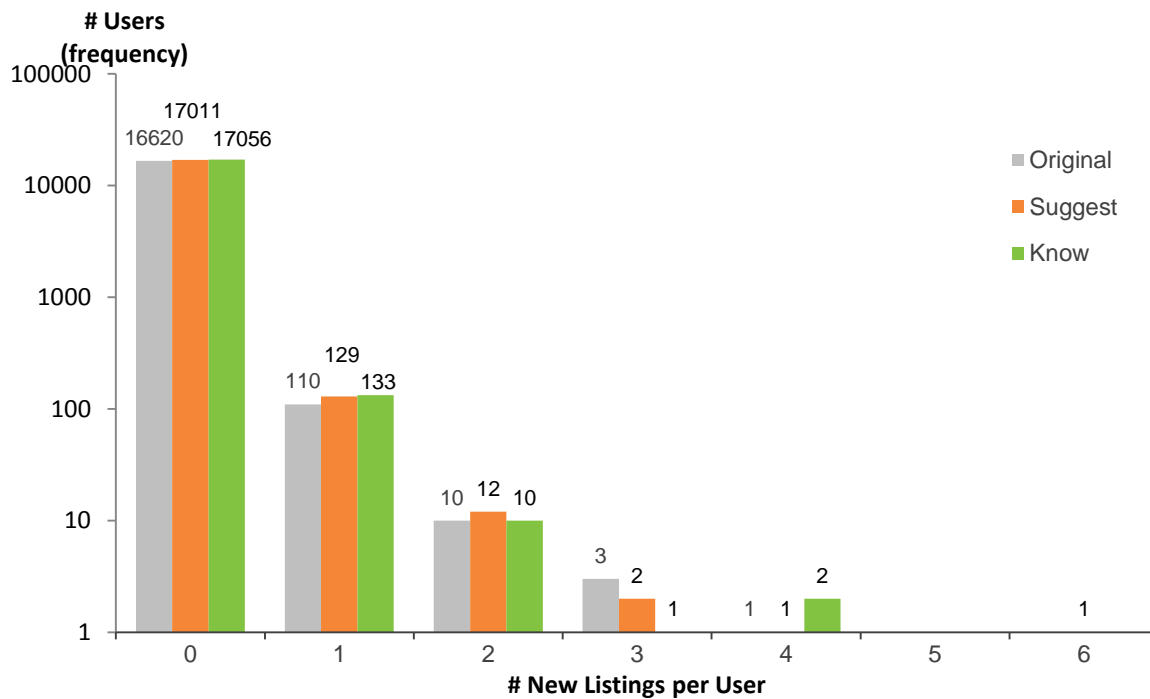


Figure 42: Number of users with a certain number of new listings posted, split per experiment variant

The table below shows the N, sum, mean, variance, standard deviation, minimum and maximum value for the number of new listings per user, split for the Original, Suggest and Know variants. Normally, when an experiment contains a large number of users, the variances are naturally low (Dan Siroker & Koomen, 2013). Here, this is not the case.

Table 12: Descriptive statistics of new listings per user, split per experiment variant

	$\chi^2$ value	Degrees of Freedom	P-value
All	11046.48	5	0
Original	881.2632	3	1.03E-190
Suggest	4228.586	5	0
Know with outlier	5380267	26	0
Know without outlier	4235.401	5	0

### I.1.2 Normality Inspection

Because of the high N of the three variants, a normality assumption according to the central limit theorem could be made. QQ-plots of the quantiles from each variant against the theoretical quantiles of a normal distribution will be used to visually examine if the distribution of new listings per user of Original, Suggest and Know can approach the normal distribution. See Figure 43 below.

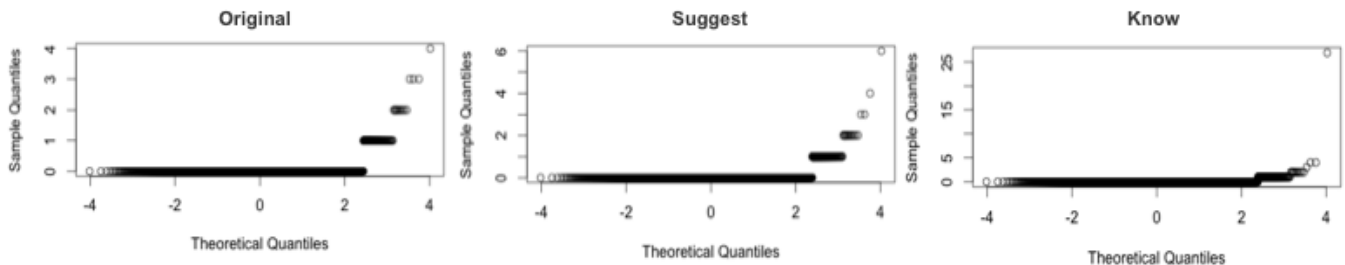


Figure 43: QQ-plots of variant quantiles (y-axis) against theoretical normal quantiles (x-axis)

Clearly, there is no good match, because the plots do not approach a straight 45 degree line. Especially Know deviates from the normal distribution, due to the outlier who posted 27 listings (see page 133 for more on this).

In order to confirm the beliefs on non-normality from above, Shapiro-Wilk tests for normality were conducted, where the null-hypothesis is that the dataset given comes from a Normal distribution. Multiple samples with  $n=500$  and  $n=5000$  (maximum sample size for the Shapiro-Wilk test in R) from the Original, Suggest and Know variant overall new-listings-per-user data were taken and tested. All Shapiro-Wilk tests resulted in p-values smaller than  $2.2 \times 10^{-16}$  ( $< 0.05$ ) so the null-hypothesis could be rejected. All in all, it is clear that normality for new-listings-per-user data from the experiment cannot be assumed and it is very likely that this data does not follow a Normal distribution.

### I.1.3 Two-Sample Comparison with Mann-Whitney U Test

In order to test the individual difference between the values of two samples, a Mann-Whitney U test (or Wilcoxon rank-sum test) is used which is the non-parametric alternative to the independent-samples t-test. It is mainly useful for data that is non-normal with independent treatment and control groups, with a categorical independent variable that influences an ordinal or ratio dependent variable. It is more efficient than the t-test for non-normal distributions and nearly as efficient as the t-test on normal distributions. This is the case for both hypothesis 1 and 2. Also, the Mann-Whitney U test is used in a very similar gamification study on an online marketplace (Hamari, 2013).

There are no conditions for data distributions, as the test is non-parametric. The assumptions of the test are based on the experiment setup and the type of data that is compared:

- There is one dependent variable that is measured at the continuous or ordinal level: the number of listings per user.
- There is one independent variable that consists of two categorical, independent groups: Original and Suggest groups, or Original and Know groups, or users who did and did not visit the product page, etc.
- The samples of the independent variables are independent and the experimental units within the samples are independent. This is met because the users are all in only one variant or in one group (have or have not visited the product page). The users in the experiment are included due to random sampling by Optimizely. Also, it is very unlikely that the users influenced each other in their posting behaviour.

These assumptions are met for the new-listings-per-user-data of the experiment and the experiment setup (as long as two independent samples are compared).

### I.1.4 Possibilities for Regression Models

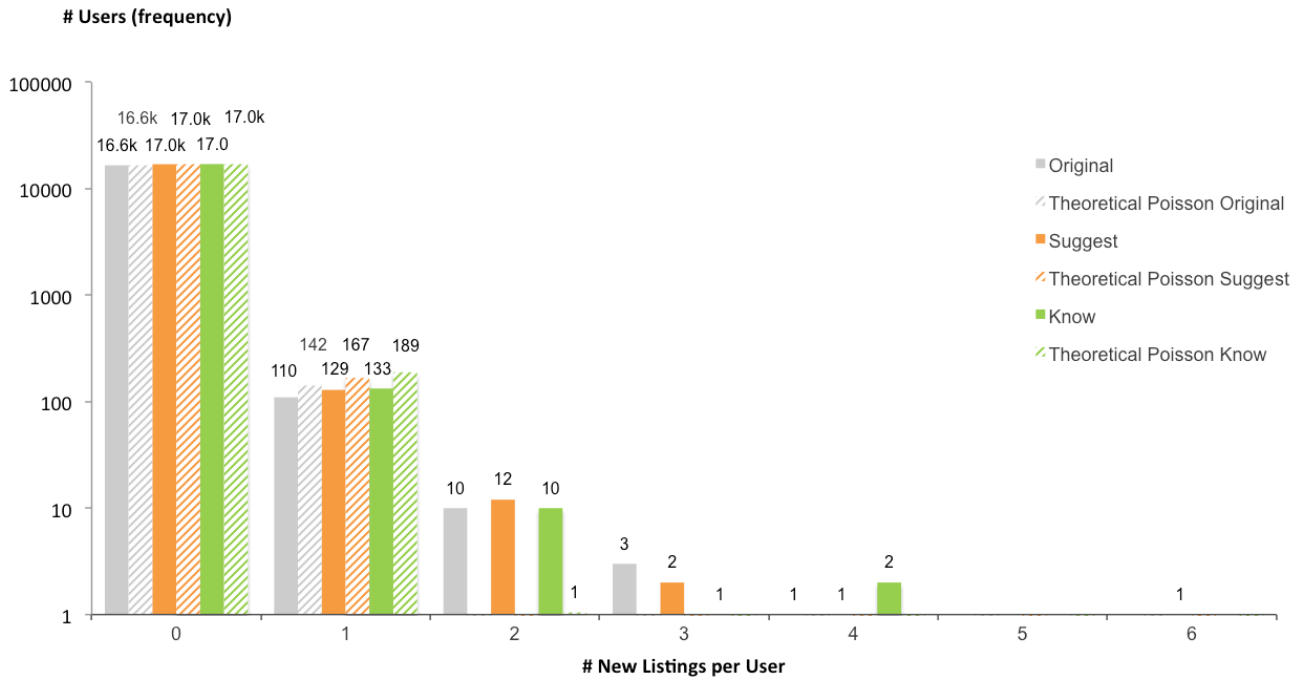
#### Poisson Distribution

It is likely that the new-listings-per-user data follows a Poisson distribution. In a Poisson process, a certain time interval, subject or area is observed and a success is counted whenever a certain event within this interval, subject or area occurs. The number of successes is always a positive integer ( $\geq 0$ ). In this case, the subject is a user in a variant in the experiment and a success is a new listing that was posted by that user. Data that follows a Poisson distribution is often referred to as 'count data', which is essentially what the new-listings-per-user data from the experiment is. For count data that follows a Poisson distribution, Poisson regression can be used. Linear regression models (such as Ordinary Least Squares or T-tests) assume that data is normally distributed around an expected value and can have either positive or negative continuous values. In the previous section, it was proven that this is not the case for new-listings-per-user data. Logistic regression is an alternative, but is applicable only to data with a Binomial distribution, thus with Boolean values only (1 or 0; true or false; yes or no). This leaves Poisson regression, used for positive integer (count) data.

The reasons to believe the new-listings-per-user data follows a Poisson distribution are threefold. First of all, the normality approximation for a sample with Poisson distribution is not dependent on the  $N$ , according to the central limit theorem, but on the parameter  $\lambda$  ( $\geq 0$ ). If  $\lambda$  is a large integer ( $>30$ ), the Poisson distribution is approximately normal. The  $\lambda$  represents the mean of Poisson distributed dataset, so would be 0.00854, 0.00985 and 0.01110 for Original, Suggest and Know respectively. So, if a Poisson distribution represents the new-listings-per-user data from the variants, a normality approximation definitely does not hold, corresponding with the conclusions made based on the QQ-plots above. The second reason is that there are no negative numbers of new listings and the number of new listings per user is always a positive integer, corresponding with a Poisson process. The third reason is because the experiment setup matches with the assumptions of a Poisson process.

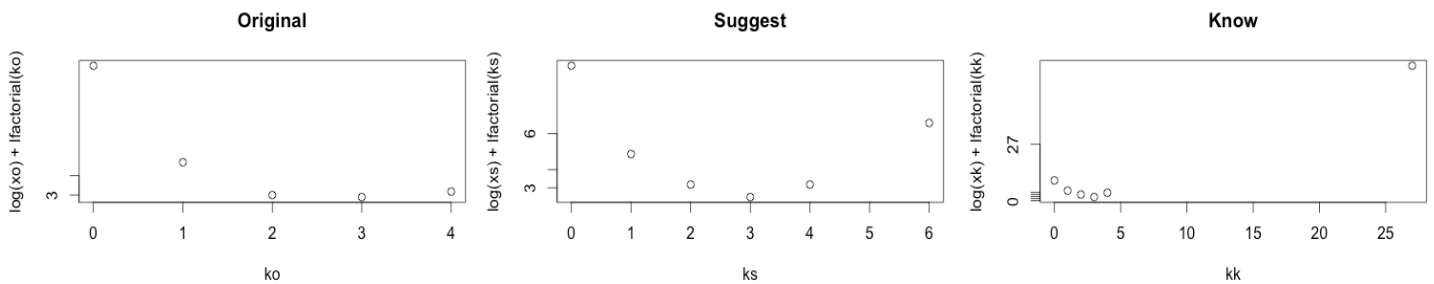
- Independence: the probability that one user in a variant posts a new listing is independent of the probability that another user in the same variant posts a new listing.
- Stationarity: the probability that one user in a variant posts a listing is the same as another user in a variant posting a listing, i.e. the differences are negligible. This does not hold on an individual level, because each user is different, etc. However, it does hold for the experiment as conducted, within each variant. The users in the Original variant are all assumed to be equally likely to post a new listing, because according to the experiment setup, there is no independent variable that differs for these users. The same goes for the users within the Suggest and Know variants. However, between the different variants, the probability is not the same, for Suggest users experience a difference website than Original users.
- Rareness: there is no chance of a users posting two new listings at exactly the same time and the probability of a new listing to be posted is proportional to the number of users that are examined.

When the means of the new-listings-per-user data of Original, Suggest and Know are used as  $\lambda$  in three theoretical Poisson functions, the graph in Figure 42 can be expanded with the theoretical frequencies as results of the Poisson functions, which was done in Figure 44. This way, the applicability of the Poisson distribution can also be visually inspection, adding to the theoretical confirmation as depicted above. The comparison of the theoretical Poisson distributions with the shows the fit is quite good for 0 or 1 new listings per user. However, 2 or more listings per user do occur in the experiment data, while the theoretical distributions cannot explain this. This observation is not enough to reject the application of a Poisson distribution to this experiment data in general, but does indicate that the theoretical Poisson distributions are not fully accurate. They will not be further used in analyses, so for this research, it is not a big issue. However, it might also indicate that the Poisson distribution is not a good fit for the new-listings-per-user data, even though the experiment seemed to comply with the assumptions in a Poisson process.



**Figure 44: Number of users with a certain number of new listings posted, split per experiment variant and theoretical Poisson distribution per variant**

Because the theoretical Poisson distributions do not fit very well, a closer look is needed. Another visual inspection method is the so-called 'Poissonness plot', which is comparable to the QQ-plot to visually inspect for a Normal distribution: the plotted points (calculated with the frequencies and post-per-user values from the data samples, where 'x' is the table containing both frequencies and values and 'k' are the values, the 'o', 's', and 'k' after k and x indicate the name of the variants) should make a straight line if the sample data fits a Poisson distribution (Hoaglin, 1980). These Poissonness plots are made in Figure 45 below for Original, Suggest and Know. As one can see, the three plots are definitely not straight lines.



**Figure 45: Poissonness plots for the new-listings-per-user data of Original, Suggest and Know**

Final checks in the form of Chi-Square goodness of fit tests yielded P-values smaller than 0.05 for Original, Suggest, Know with outlier and Know without outlier, meaning that the null-hypothesis for the goodness of fit test (the new-listings-per-user-data follows a Poisson distribution) should be rejected for all variants. See Table 13 below.

**Table 13: Chi-Square goodness of fit test to Poisson distribution for new-listings-per-user-data per variant and for all**

	$\chi^2$ value	Degrees of Freedom	P-value
All	11046.48	5	0



Original	881.2632	3	1.03E-190
Suggest	4228.586	5	0
Know with outlier	5380267	26	0
Know without outlier	4235.401	5	0

Concluding, the new-listings-per-user data very likely does not follow a Poisson distribution (even though the assumptions for the Poisson process are valid with regard to the experiment setup).

Moreover, if Poisson regression models would be used for hypothesis 1 and 2, the models would violate each other's assumptions. For example: for hypothesis 1 the assumption (according to Poisson distribution) needs to be made that the probability that a user in a variant posts a listing is equal for each user. One could argue this assumption in the first place, by stating that each user is different, so the probability is not the same. However, external but possibly influential variables on the number of new listings per user were not taken in to account in the experiment (age, gender, but also average internet usage and posting history). Therefore the assumption does hold for the experiment setup, strictly speaking. However, continuing on to hypothesis 2, a distinction is made between users who did and did not post a listing, within the same user samples (home page variants) that were used for hypothesis 1. Hypothesis 2 states there is a difference between the number of new listings per user for users who did and did not post a listing, *within* the Suggest variants. This violates the assumption that was made for hypothesis 1, because it implies that the probability for a user to post a new listing is *not* the same for all users within the Suggest variant.

### Negative Binomial Model

Now that it has been confirmed that the new-listings-per-user data does *not* follow a Normal nor a Poisson distribution, other distributions need to be inspected. A common practice within statistical analysis of count data is to examine possible over-dispersion. This occurs when the variance of Poisson data samples is greater than the mean, while true Poisson data means and variances are equal. For the new-listings-per-user data of the experiment, the variance is much higher than the mean, as can be seen in Table 12. This could be the reason that the Poisson distribution does not fit well (Cameron & Trivedi, 1998; Gardner et al., 1995; Nussbaum et al., 2008; Zeileis et al., 2008).

Over-dispersion can occur because of several reasons (Cameron & Trivedi, 1998). The first is because of heterogeneity, i.e. there are independent variables that influence the dependent variable which cause spreading in the data which cannot be predicted with a single  $\lambda$  parameter in a Poisson distribution. This could be the case, since there are listings posted by new and returning users in each variant and this variable is predicted in the hypotheses to influence the user posting behaviour. Also there are other variables, such as internet browser, source type and location which could be of influence. Moreover, click behaviour can be very irrational and there are countless more variables that were not measured in the experiment but can explain the variance in new listings per user.

A second reason for over-dispersion can be that the independence assumption does not hold, which would also lead to the negative binomial distribution as solution. This is not the case, because it can safely be assumed that all users act independent of each other and do not influence each other directly in their posting behaviour. Of course, there could be users who are friends or family, but this chance is so rare for the experiment dataset, given the random sampling, that it is negligible. Another theoretical disruption of the independence assumption could be that users post a listing based on the listings that are already online. But again, this is a theoretical situation which should not severely influence the dataset.

For count data, it is quite common to use negative binomial models in case of over-dispersion of data from a theoretically valid Poisson process (Cameron & Trivedi, 1998; Nussbaum et al., 2008). The negative binomial distribution is essentially a Poisson distribution with an extra parameter.

### Advanced: Hurdle and Zero-Inflated Models

Strictly speaking, in a Poisson distribution, each event is genuinely independent and has the same probability to occur. In the experiment, a user who posts a listing is thought to be more likely to post another listing. This is because there is an initial barrier to post (difficult, scary, not knowing what to sell and several other challenges as mentioned in chapter 3). Also, someone who posts a listing can be moving or cleaning up their house and thus have several items to post. The fact that users who have posted a listing have an increased chance of posting another listing has been confirmed by OLX in their customer lifecycle report (OLX, 2014c). In fact, there is a sort of two-step decision making process, whereby some factors influence the decision to post a first listing and other factors influence the decision to post more listings. Or, the rationale and decision making process changes from the first visit to the second visit. In order to correct for this, a modified count model has been introduced, called the 'hurdle' or 'two-step' model. Apart from the two-step decision making process, this model corrects for the excess zeros: datasets that contain more cases with zero events than was predicted by their theoretical Poisson distribution. Another similar model is the Zero-Inflated model (Cameron & Trivedi, 1998; Gardner et al., 1995; Nussbaum et al., 2008; Zeileis et al., 2008). The over-dispersion and excess zero's have now been solved by a negative binomial regression model, but it would be interesting to see if Hurdle and/or Zero-Inflated regression models fit the data better. These regression models are not readily available in SPSS and require an extra amount of statistical knowledge to interpret when running from R. Therefore, they were not used in this research.

## I.2 Comparing Two Proportions

For some parts of the analysis, two proportions need to be compared, in order to check if they differ significantly. A few examples of proportional values to be compared between Original, Suggest and Know:

- Proportion of Selling Assistant home page visits with an SA button click;
- Proportion of home page visits with an interaction (bounce rate);
- Proportion of unique users with a posted listing ( $\geq 1$ ).

These are two sample comparisons, like T-tests, but not for means but for proportions. For this, the Z-test for proportions of Independent samples can be used. This is the equivalent of a T-test for independent samples, but then to compare proportions instead of means. Data for such a proportion value has a binomial distribution, because it consists only of successes and failures. When the sample size is large enough ( $N > 30$ ) and the number of successes ( $X = N \cdot p$ ) and failures ( $N \cdot (1-p)$ ) is larger than 10, normality can be assumed according to the central limit theorem. Assumptions for the Z-test are (DasGupta, Cai, & Brown, 2001; Dan Siroker & Koomen, 2013):

1. The sampling method for each population is simple random sampling;
2. The samples are independent;
3. Each sample includes at least 10 successes and 10 failures and  $N > 30$ ;
4. Each population is at least 10 times as big as its sample.

All these assumptions are valid for the experiment setup and when the Original, Suggest and Know variants are compared. In this case, if gamification is implemented on the entire OLX mobile website, the population will be at least 10 times as big, because 10% of current traffic was allocated to the experiment. It is a relative rather than absolute metric, so it is independent of the growth or decline of the number of users.

## J EXTENSIVE RESULTS OF STATISTICAL TESTS

### J.1 Hypothesis 1

#### J.1.1 Negative Binomial Regression Model 1

This is the model in which no predictor variables are included, only experiment variant as independent variable and the number of new listings as dependent variable.

**Goodness of Fit<sup>a</sup>**

	Value	df	Value/df
Deviance	4250.073	51100	.083
Scaled Deviance	4250.073	51100	
Pearson Chi-Square	71810.223	51100	1.405
Scaled Pearson Chi-Square	71810.223	51100	
Log Likelihood <sup>b</sup>	-2731.089		
Akaike's Information Criterion (AIC)	5468.178		
Finite Sample Corrected AIC (AICC)	5468.179		
Bayesian Information Criterion (BIC)	5494.703		
Consistent AIC (CAIC)	5497.703		

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home<sup>a</sup>

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

**Omnibus Test<sup>a</sup>**

Likelihood Ratio Chi-Square	df	Sig.
2.116	2	.347

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home<sup>a</sup>

a. Compares the fitted model against the intercept-only model.

**Tests of Model Effects**

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	10333.124	1	.000
variant_id_home	2.077	2	.354

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home

**Parameter Estimates**

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-4.763	.0840	-4.928	-4.598	3216.585	1	.000
[variant_id_home=suggest ]	.143	.1141	-.081	.366	1.564	1	.211
[variant_id_home=know ]	.146	.1140	-.078	.369	1.639	1	.201
[variant_id_home=c_original ]	0 <sup>a</sup>	.	.	.	.	.	.
(Scale)	1 <sup>b</sup>						
(Negative binomial)	1 <sup>b</sup>						

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

### J.1.2 Negative Binomial Regression Model 2

This is the model in which all predictor variables from the experiment data are included.

**Goodness of Fit<sup>a</sup>**

	Value	df	Value/df
Deviance	3939.282	51079	.077
Scaled Deviance	3939.282	51079	
Pearson Chi-Square	60429.623	51079	1.183
Scaled Pearson Chi-Square	60429.623	51079	
Log Likelihood <sup>b</sup>	-2575.694		
Akaike's Information Criterion (AIC)	5199.387		
Finite Sample Corrected AIC (AICC)	5199.411		
Bayesian Information Criterion (BIC)	5411.586		
Consistent AIC (CAIC)	5435.586		

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home, browser, source, location, page\_load\_sum, variant\_id\_home \* page\_load\_sum<sup>a</sup>

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

**Omnibus Test<sup>a</sup>**

Likelihood Ratio Chi-Square	df	Sig.
312.908	23	.000

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home, browser, source, location, page\_load\_sum, variant\_id\_home \* page\_load\_sum<sup>a</sup>

a. Compares the fitted model against the intercept-only model.

**Tests of Model Effects**

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	559.607	1	.000
variant_id_home	.212	2	.899
browser	20.552	6	.002
source	4.232	3	.237
location	83.123	6	.000
page_load_sum	101.196	1	.000
variant_id_home * page_load_sum	8.365	2	.015

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home, browser, source, location, page\_load\_sum, variant\_id\_home \* page\_load\_sum

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-6.659	.6361	-7.906	-5.412	109.581	1	.000	.001	.000	.004
[variant_id_home=suggest ]	.055	.1207	-.182	.292	.207	1	.650	1.056	.834	1.338
[variant_id_home=know ]	.037	.1209	-.200	.274	.092	1	.762	1.037	.818	1.315
[variant_id_home=c_original ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[browser=unknown ]	-1.251	1.1589	-3.522	1.020	1.166	1	.280	.286	.030	2.774
[browser=ucbrowser]	-.157	.6955	-1.521	1.206	.051	1	.821	.854	.219	3.339
[browser=safari ]	.139	.5866	-1.010	1.289	.056	1	.812	1.150	.364	3.630
[browser=opera ]	.338	.7696	-1.170	1.847	.193	1	.660	1.403	.310	6.339
[browser=ie ]	-.322	.6514	-1.599	.955	.245	1	.621	.725	.202	2.598
[browser=gc ]	.476	.5841	-.669	1.621	.664	1	.415	1.609	.512	5.056
[browser=ff ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[source=search ]	-.302	.1479	-.591	-.012	4.158	1	.041	.740	.554	.988
[source=referral]	-.256	.1914	-.631	.119	1.792	1	.181	.774	.532	1.126
[source=direct ]	-.251	.1503	-.545	.044	2.784	1	.095	.778	.580	1.045
[source=campaign]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[location=PUNE ]	1.385	.5499	.307	2.463	6.345	1	.012	3.996	1.360	11.740
[location=other ]	2.089	.2399	1.619	2.559	75.840	1	.000	8.079	5.048	12.929
[location=NEWDELHI ]	2.086	.2425	1.611	2.562	74.025	1	.000	8.057	5.009	12.959
[location=MUMBAI ]	1.991	.2388	1.523	2.459	69.479	1	.000	7.320	4.584	11.689
[location=JAIPUR ]	-21.733	329494.2211	-645818.540	645775.073	.000	1	1.000	3.642E-10	.000	<sup>b</sup>
[location=COIMBATORE ]	-21.609	364605.4940	-714635.246	714592.028	.000	1	1.000	4.125E-10	.000	<sup>b</sup>
[location=CHENNAI ]	2.130	.2992	1.544	2.716	50.696	1	.000	8.416	4.682	15.126
[location=CHANDIGARH ]	1.528	.6214	.310	2.745	6.043	1	.014	4.607	1.363	15.572
[location=BENGALURU ]	-21.728 <sup>c</sup>	.	.	.	.	.	.	3.662E-10	.000	.000
[location= ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
page_load_sum	.004	.0013	.002	.006	10.036	1	.002	1.004	1.002	1.006
[variant_id_home=suggest ] * page_load_sum	.004	.0016	.001	.007	5.284	1	.022	1.004	1.001	1.007
[variant_id_home=know ] * page_load_sum	.005	.0017	.001	.008	7.497	1	.006	1.005	1.001	1.008
[variant_id_home=c_original ] * page_load_sum	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
(Scale)	1 <sup>d</sup>									
(Negative binomial)	1 <sup>d</sup>									

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home, browser, source, location, page\_load\_sum, variant\_id\_home \* page\_load\_sum

a. Set to zero because this parameter is redundant.

b. Set to system missing due to overflow

c. Hessian matrix singularity is caused by this parameter. The parameter estimate at the last iteration is displayed.

d. Fixed at the displayed value.

### J.1.3 Negative Binomial Regression Model 3

This is the same as model 2, but with all standard error scaled by ratio between the deviance and degrees of freedom (which was too small, lower than 1, probably due to some misfit issues).

**Goodness of Fit<sup>a</sup>**

	Value	df	Value/df
Deviance	3939.282	51079	.077
Scaled Deviance	51079.000	51079	
Pearson Chi-Square	60429.623	51079	1.183
Scaled Pearson Chi-Square	783565.291	51079	
Log Likelihood <sup>b,c</sup>	-2575.694		
Adjusted Log Likelihood <sup>d</sup>	-33397.927		
Akaike's Information Criterion (AIC)	5199.387		
Finite Sample Corrected AIC (AICC)	5199.411		
Bayesian Information Criterion (BIC)	5411.586		
Consistent AIC (CAIC)	5435.586		

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home, browser, source, location, page\_load\_sum, variant\_id\_home \* page\_load\_sum<sup>a</sup>

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

c. The log likelihood is based on a scale parameter fixed at 1.

d. The adjusted log likelihood is based on an estimated scale parameter and is used in the model fitting omnibus test.

**Omnibus Test<sup>a</sup>**

Likelihood Ratio		
Chi-Square	df	Sig.
4057.339	23	.000

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home, browser, source, location, page\_load\_sum, variant\_id\_home \* page\_load\_sum<sup>a</sup>

a. Compares the fitted model against the intercept-only model.

**Tests of Model Effects**

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	7256.183	1	.000
variant_id_home	2.754	2	.252
browser	266.487	6	.000
source	54.873	3	.000
location	1077.815	6	.000
page_load_sum	1312.160	1	.000
variant_id_home * page_load_sum	108.462	2	.000

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home, browser, source, location, page\_load\_sum, variant\_id\_home \* page\_load\_sum

## Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-6.659	.1767	-7.005	-6.313	1420.894	1	.000	.001	.001	.002
[variant_id_home=suggest ]	.055	.0335	-.011	.121	2.678	1	.102	1.056	.989	1.128
[variant_id_home=know ]	.037	.0336	-.029	.103	1.193	1	.275	1.037	.971	1.108
[variant_id_home=c_original]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[browser=unknown ]	-1.251	.3218	-1.882	-.620	15.114	1	.000	.286	.152	.538
[browser=ucbrowser]	-.157	.1932	-.536	.221	.664	1	.415	.854	.585	1.248
[browser=safari ]	.139	.1629	-.180	.459	.731	1	.392	1.150	.835	1.582
[browser=opera ]	.338	.2137	-.080	.757	2.508	1	.113	1.403	.923	2.133
[browser=ie ]	-.322	.1809	-.677	.032	3.171	1	.075	.725	.508	1.033
[browser=gc ]	.476	.1622	.158	.794	8.607	1	.003	1.609	1.171	2.212
[browser=ff ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[source=search ]	-.302	.0411	-.382	-.221	53.921	1	.000	.740	.682	.802
[source=referral]	-.256	.0531	-.360	-.152	23.238	1	.000	.774	.697	.859
[source=direct ]	-.251	.0417	-.333	-.169	36.100	1	.000	.778	.717	.845
[source=campaign]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[location=PUNE ]	1.385	.1527	1.086	1.685	82.272	1	.000	3.996	2.962	5.390
[location=other ]	2.089	.0666	1.959	2.220	983.387	1	.000	8.079	7.090	9.206
[location=NEWDELHI ]	2.086	.0673	1.954	2.218	959.848	1	.000	8.057	7.060	9.193
[location=MUMBAI ]	1.991	.0663	1.861	2.121	900.911	1	.000	7.320	6.427	8.336
[location=JAIPUR ]	-21.733	91502.9616	-179364.243	179320.776	.000	1	1.000	3.642E-10	.000	. <sup>b</sup>
[location=COIMBATORE ]	-21.609	101253.6196	-198475.057	198431.839	.000	1	1.000	4.125E-10	.000	. <sup>b</sup>
[location=CHENNAI ]	2.130	.0831	1.967	2.293	657.352	1	.000	8.416	7.151	9.904
[location=CHANDIGARH ]	1.528	.1726	1.189	1.866	78.356	1	.000	4.607	3.285	6.461
[location=BENGALURU ]	-21.728 <sup>c</sup>	.	.	.	.	.	.	3.662E-10	.000	.000
[location= ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
page_load_sum	.004	.0003	.003	.005	130.136	1	.000	1.004	1.003	1.005
[variant_id_home=suggest ]	.004	.0005	.003	.005	68.511	1	.000	1.004	1.003	1.005
* page_load_sum										
[variant_id_home=know ]	.005	.0005	.004	.006	97.206	1	.000	1.005	1.004	1.006
* page_load_sum										
[variant_id_home=c_original]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
* page_load_sum										
(Scale)	.077 <sup>d</sup>									
(Negative binomial)	1 <sup>e</sup>									

Dependent Variable: post\_sum

Model: (Intercept), variant\_id\_home, browser, source, location, page\_load\_sum, variant\_id\_home \* page\_load\_sum

a. Set to zero because this parameter is redundant.

b. Set to system missing due to overflow

c. Hessian matrix singularity is caused by this parameter. The parameter estimate at the last iteration is displayed.

d. Computed based on the deviance.

e. Fixed at the displayed value.

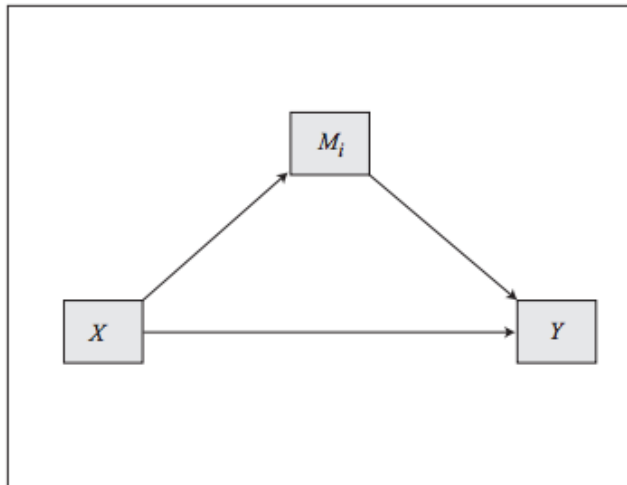


### J.1.4 OLS Regression Model with Mediation

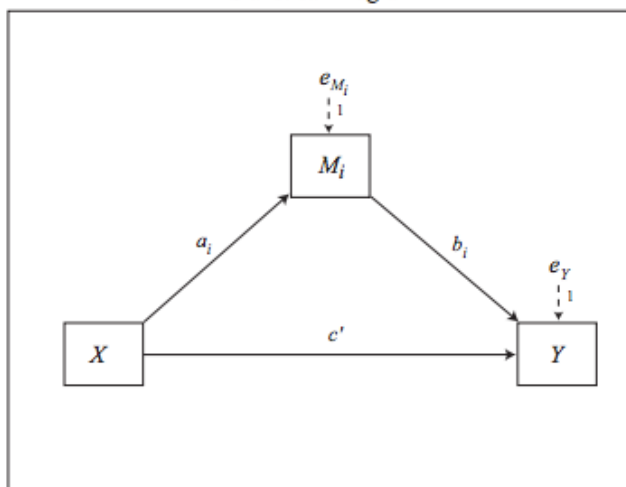
In this model, the PROCESS macro and conceptual model 4 (see image below) for SPSS by Andrew Hayes (2013) was used, in order to check the mediation effect of users' page views ('Z\_PVsa') on the relationship between the experiment variant ('SA') and the number of new listings ('Z\_Posts'). The values of the variables page views and new listings were standardised for this. The indirect effect was not significant.

**Model 4**

Conceptual Diagram



Statistical Diagram



Indirect effect of  $X$  on  $Y$  through  $M_i = a_i b_i$

Direct effect of  $X$  on  $Y = c'$

\*\*\*\*\* PROCESS Procedure for SPSS Release 2.13 \*\*\*\*\*

Written by Andrew F. Hayes, Ph.D. [www.afhayes.com](http://www.afhayes.com)  
Documentation available in Hayes (2013). [www.guilford.com/p/hayes3](http://www.guilford.com/p/hayes3)

\*\*\*\*\*

Model = 4  
Y = Z\_Posts  
X = SA  
M = Z\_PVsa

Sample size  
51103

\*\*\*\*\*

Outcome: Z\_PVs

Model Summary

R	R-sq	MSE	F	df1	df2	p
.0036	.0000	1.0000	.6760	1.0000	51101.0000	.4110

Model

	coeff	se	t	p	LLCI	ULCI
constant	-.0052	.0077	-.6742	.5002	-.0204	.0099
SA	.0077	.0094	.8222	.4110	-.0107	.0262

Covariance matrix of regression parameter estimates

	constant	SA
constant	.0001	-.0001
SA	-.0001	.0001

\*\*\*\*\*

Outcome: Z\_Posts

Model Summary

R	R-sq	MSE	F	df1	df2	p
.0824	.0068	.9933	174.5410	2.0000	51100.0000	.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	-.0073	.0077	-.9439	.3452	-.0224	.0078
Z_PVs	.0822	.0044	18.6439	.0000	.0736	.0908
SA	.0108	.0094	1.1511	.2497	-.0076	.0292

Covariance matrix of regression parameter estimates

	constant	Z_PVs	SA
constant	.0001	.0000	-.0001
Z_PVs	.0000	.0000	.0000
SA	-.0001	.0000	.0001

\*\*\*\*\* TOTAL EFFECT MODEL \*\*\*\*\*

Outcome: Z\_Posts

Model Summary

R	R-sq	MSE	F	df1	df2	p
.0054	.0000	1.0000	1.4758	1.0000	51101.0000	.2244

Model

	coeff	se	t	p	LLCI	ULCI
constant	-.0077	.0077	-.9961	.3192	-.0228	.0074
SA	.0114	.0094	1.2148	.2244	-.0070	.0299

Covariance matrix of regression parameter estimates

	constant	SA
constant	.0001	-.0001
SA	-.0001	.0001

\*\*\*\*\* TOTAL, DIRECT, AND INDIRECT EFFECTS \*\*\*\*\*

Total effect of X on Y

Effect	SE	t	p	LLCI	ULCI
.0114	.0094	1.2148	.2244	-.0070	.0299

Direct effect of X on Y

Effect	SE	t	p	LLCI	ULCI
.0108	.0094	1.1511	.2497	-.0076	.0292

Indirect effect of X on Y

	Effect	Boot SE	BootLLCI	BootULCI
Z_PVs	.0006	.0008	-.0009	.0023

Partially standardized indirect effect of X on Y

	Effect	Boot SE	BootLLCI	BootULCI
Z_PVs	.0006	.0008	-.0009	.0023

Completely standardized indirect effect of X on Y

	Effect	Boot SE	BootLLCI	BootULCI
Z_PVs	.0003	.0004	-.0004	.0011

Ratio of indirect to total effect of X on Y

	Effect	Boot SE	BootLLCI	BootULCI
Z_PVs	.0556	.7069	-.2161	1.4401

Ratio of indirect to direct effect of X on Y

	Effect	Boot SE	BootLLCI	BootULCI
Z_PVs	.0589	.9193	-.2111	1.8154

R-squared mediation effect size (R-sq\_med)

	Effect	Boot SE	BootLLCI	BootULCI
Z_PVs	.0000	.0000	.0000	.0000

Preacher and Kelley (2011) Kappa-squared

	Effect	Boot SE	BootLLCI	BootULCI
Z_PVs	.0003	.0003	.0000	.0010

\*\*\*\*\* ANALYSIS NOTES AND WARNINGS \*\*\*\*\*

Number of bootstrap samples for bias corrected bootstrap confidence intervals:  
1000

Level of confidence for all confidence intervals in output:  
95.00

----- END MATRIX -----

## J.2 Hypothesis 2

### J.2.1 Comparing Users Within Suggest

Negative Binomial Regression Model 1 – No Scaling

**Goodness of Fit<sup>a</sup>**

	Value	df	Value/df
Deviance	1169.115	16906	.069
Scaled Deviance	1169.115	16906	
Pearson Chi-Square	17655.432	16906	1.044
Scaled Pearson Chi-Square	17655.432	16906	
Log Likelihood <sup>b</sup>	-796.176		
Akaike's Information Criterion (AIC)	1628.351		
Finite Sample Corrected AIC (AICC)	1628.392		
Bayesian Information Criterion (BIC)	1767.608		
Consistent AIC (CAIC)	1785.608		

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum<sup>a</sup>

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

**Omnibus Test<sup>a</sup>**

Likelihood Ratio		
Chi-Square	df	Sig.
304.358	17	.000

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum<sup>a</sup>

a. Compares the fitted model against the intercept-only model.

**Tests of Model Effects**

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	141.467	1	.000
source	3.348	3	.341
location	22.729	6	.001
page_load_sum	31.548	1	.000
product_page	160.273	1	.000
browser	14.591	5	.012
product_page * page_load_sum	.037	1	.847

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-7.093	1.0351	-9.122	-5.065	46.957	1	.000	.001	.000	.006
[source=search ]	-.416	.2499	-.905	.074	2.769	1	.096	.660	.404	1.077
[source=referral]	-.515	.3327	-1.167	.138	2.391	1	.122	.598	.311	1.148
[source=direct ]	-.384	.2584	-.891	.122	2.213	1	.137	.681	.410	1.130
[source=campaign]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[location=PUNE ]	2.588	1.0246	.579	4.596	6.378	1	.012	13.298	1.785	99.066
[location=other ]	3.109	.7239	1.690	4.528	18.443	1	.000	22.398	5.420	92.562
[location=NEWDELHI ]	3.174	.7257	1.751	4.596	19.127	1	.000	23.897	5.763	99.093
[location=MUMBAI ]	2.975	.7241	1.556	4.395	16.884	1	.000	19.598	4.741	81.018
[location=CHENNAI ]	3.522	.7658	2.021	5.023	21.153	1	.000	33.857	7.547	151.879
[location=CHANDIGARH ]	2.601	1.2338	.183	5.019	4.444	1	.035	13.475	1.200	151.270
[location= ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
page_load_sum	.006	.0013	.003	.008	17.497	1	.000	1.006	1.003	1.008
[product_page=yes]	2.599	.2053	2.197	3.002	160.273	1	.000	13.455	8.997	20.121
[product_page=no ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[browser=ucbrowser]	-2.037	1.2539	-4.495	.421	2.639	1	.104	.130	.011	1.523
[browser=safari ]	-.528	.7428	-1.984	.928	.505	1	.477	.590	.138	2.529
[browser=opera ]	.387	1.0266	-1.625	2.399	.142	1	.706	1.473	.197	11.015
[browser=ie ]	-2.861	1.2193	-5.251	-.472	5.507	1	.019	.057	.005	.624
[browser=gc ]	-.185	.7365	-1.629	1.258	.063	1	.801	.831	.196	3.519
[browser=ff ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[product_page=yes] *										
page_load_sum	.000	.0020	-.004	.004	.037	1	.847	1.000	.996	1.004
[product_page=no ] *										
page_load_sum	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
(Scale)	1 <sup>b</sup>									
(Negative binomial)	1 <sup>b</sup>									

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

## Negative Binomial Regression Model 2 – Scaling of Standard Errors by Deviance

Goodness of Fit<sup>a</sup>

	Value	df	Value/df
Deviance	1169.115	16906	.069
Scaled Deviance	16906.000	16906	
Pearson Chi-Square	17655.432	16906	1.044
Scaled Pearson Chi-Square	255306.547	16906	
Log Likelihood <sup>b,c</sup>	-796.176		
Adjusted Log Likelihood <sup>d</sup>	-11513.104		
Akaike's Information Criterion (AIC)	1628.351		
Finite Sample Corrected AIC (AICC)	1628.392		
Bayesian Information Criterion (BIC)	1767.608		
Consistent AIC (CAIC)	1785.608		

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum<sup>a</sup>

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

c. The log likelihood is based on a scale parameter fixed at 1.

d. The adjusted log likelihood is based on an estimated scale parameter and is used in the model fitting omnibus test.

Omnibus Test<sup>a</sup>

Likelihood Ratio		
Chi-Square	df	Sig.
4401.174	17	.000

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum<sup>a</sup>

a. Compares the fitted model against the intercept-only model.

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	2045.684	1	.000
source	48.410	3	.000
location	328.678	6	.000
page_load_sum	456.206	1	.000
product_page	2317.629	1	.000
browser	210.993	5	.000
product_page * page_load_sum	.539	1	.463

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-7.093	.2722	-7.627	-6.560	679.025	1	.000	.001	.000	.001
[source=search ]	-.416	.0657	-.545	-.287	40.037	1	.000	.660	.580	.751
[source=referral]	-.515	.0875	-.686	-.343	34.577	1	.000	.598	.504	.710
[source=direct ]	-.384	.0680	-.518	-.251	32.005	1	.000	.681	.596	.778
[source=campaign]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[location=PUNE ]	2.588	.2694	2.060	3.116	92.235	1	.000	13.298	7.842	22.550
[location=other ]	3.109	.1904	2.736	3.482	266.694	1	.000	22.398	15.423	32.528
[location=NEWDELHI ]	3.174	.1908	2.800	3.548	276.581	1	.000	23.897	16.440	34.736
[location=MUMBAI ]	2.975	.1904	2.602	3.349	244.151	1	.000	19.598	13.493	28.464
[location=CHENNAI ]	3.522	.2014	3.127	3.917	305.890	1	.000	33.857	22.815	50.242
[location=CHANDIGARH ]	2.601	.3245	1.965	3.237	64.258	1	.000	13.475	7.134	25.452
[location= ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
page_load_sum	.006	.0004	.005	.006	253.013	1	.000	1.006	1.005	1.006
[product_page=yes]	2.599	.0540	2.494	2.705	2317.629	1	.000	13.455	12.104	14.957
[product_page=no ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[browser=ucbrowser]	-2.037	.3297	-2.683	-1.391	38.161	1	.000	.130	.068	.249
[browser=safari ]	-.528	.1953	-.911	-.145	7.301	1	.007	.590	.402	.865
[browser=opera ]	.387	.2700	-.142	.916	2.057	1	.151	1.473	.868	2.500
[browser=ie ]	-2.861	.3206	-3.490	-2.233	79.638	1	.000	.057	.031	.107
[browser=gc ]	-.185	.1937	-.565	.194	.915	1	.339	.831	.568	1.214
[browser=ff ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[product_page=yes] *										
page_load_sum	.000	.0005	-.001	.001	.539	1	.463	1.000	.999	1.001
[product_page=no ] *										
page_load_sum	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
(Scale)	.069 <sup>b</sup>									
(Negative binomial)	1 <sup>c</sup>									

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum

a. Set to zero because this parameter is redundant.

b. Computed based on the deviance.

c. Fixed at the displayed value.



## J.2.2 Comparing Users Within Know

### Negative Binomial Regression Model 1 – No Scaling

**Goodness of Fit<sup>a</sup>**

	Value	df	Value/df
Deviance	1280.433	16958	.076
Scaled Deviance	1280.433	16958	
Pearson Chi-Square	17754.833	16958	1.047
Scaled Pearson Chi-Square	17754.833	16958	
Log Likelihood <sup>b</sup>	-852.427		
Akaike's Information Criterion (AIC)	1740.854		
Finite Sample Corrected AIC (AICC)	1740.894		
Bayesian Information Criterion (BIC)	1880.166		
Consistent AIC (CAIC)	1898.166		

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum<sup>a</sup>

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

**Omnibus Test<sup>a</sup>**

Likelihood Ratio		
Chi-Square	df	Sig.
192.887	17	.000

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum<sup>a</sup>

a. Compares the fitted model against the intercept-only model.

**Tests of Model Effects**

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	177.283	1	.000
source	.248	3	.969
location	23.929	6	.001
page_load_sum	27.238	1	.000
product_page	108.798	1	.000
browser	9.090	4	.059
product_page * page_load_sum	8.708	1	.003

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-29.771	.4312	-30.617	-28.926	4766.715	1	.000	1.176E-13	5.051E-14	2.738E-13
[source=search ]	-.084	.2682	-.610	.441	.099	1	.753	.919	.543	1.555
[source=referral]	-.116	.3432	-.788	.557	.113	1	.736	.891	.455	1.745
[source=direct ]	-.017	.2702	-.546	.513	.004	1	.951	.984	.579	1.670
[source=campaign]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[location=PUNE ]	.748	1.0676	-1.345	2.840	.490	1	.484	2.112	.261	17.119
[location=other ]	1.710	.3860	.954	2.467	19.634	1	.000	5.531	2.596	11.785
[location=NEWDELHI ]	1.655	.3919	.887	2.423	17.841	1	.000	5.234	2.428	11.284
[location=MUMBAI ]	1.733	.3809	.986	2.479	20.688	1	.000	5.656	2.681	11.933
[location=CHENNAI ]	1.074	.5764	-.056	2.203	3.469	1	.063	2.926	.945	9.055
[location=CHANDIGARH ]	1.834	.8058	.254	3.413	5.179	1	.023	6.257	1.290	30.358
[location= ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
page_load_sum	.010	.0015	.007	.013	43.365	1	.000	1.010	1.007	1.013
[product_page=yes]	2.269	.2176	1.843	2.696	108.798	1	.000	9.673	6.315	14.816
[product_page=no ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[browser=ucbrowser]	23.762	.4453	22.890	24.635	2847.572	1	.000	20887133999.753	8726470207.651	49994139250.157
[browser=safari ]	23.170	.1733	22.831	23.510	17875.833	1	.000	11556154370.194	8228088490.838	16230343654.723
[browser=opera ]	23.001	1.0178	21.006	24.996	510.739	1	.000	9755933380.670	1327214034.096	71712801163.182
[browser=ie ]	22.628	.5240	21.601	23.655	1864.599	1	.000	6720332308.708	2406193044.016	18769427686.524
[browser=gc ]	23.587 <sup>b</sup>	.	.	.	.	.	.	17524802364.234	.000	.000
[browser=ff ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
[product_page=yes] *	-.007	.0024	-.012	-.002	8.708	1	.003	.993	.988	.998
page_load_sum										
[product_page=no ] *	0 <sup>a</sup>	.	.	.	.	.	.	1	.	.
page_load_sum										
(Scale)	1 <sup>c</sup>									
(Negative binomial)	1 <sup>c</sup>									

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum

a. Set to zero because this parameter is redundant.

b. Hessian matrix singularity is caused by this parameter. The parameter estimate at the last iteration is displayed.

c. Fixed at the displayed value.

## Negative Binomial Regression Model 2 – Scaling of Standard Errors by Deviance

**Goodness of Fit<sup>a</sup>**

	Value	df	Value/df
Deviance	1280.433	16958	.076
Scaled Deviance	16958.000	16958	
Pearson Chi-Square	17754.833	16958	1.047
Scaled Pearson Chi-Square	235144.223	16958	
Log Likelihood <sup>b,c</sup>	-852.427		
Adjusted Log Likelihood <sup>d</sup>	-11289.506		
Akaike's Information Criterion (AIC)	1740.854		
Finite Sample Corrected AIC (AICC)	1740.894		
Bayesian Information Criterion (BIC)	1880.166		
Consistent AIC (CAIC)	1898.166		

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum<sup>a</sup>

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

c. The log likelihood is based on a scale parameter fixed at 1.

d. The adjusted log likelihood is based on an estimated scale parameter and is used in the model fitting omnibus test.

**Omnibus Test<sup>a</sup>**

Likelihood Ratio		
Chi-Square	df	Sig.
2554.586	17	.000

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum<sup>a</sup>

a. Compares the fitted model against the intercept-only model.

**Tests of Model Effects**

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	2347.925	1	.000
source	3.288	3	.349
location	316.920	6	.000
page_load_sum	360.741	1	.000
product_page	1440.912	1	.000
browser	120.391	4	.000
product_page * page_load_sum	115.333	1	.000

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum

## Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.	Exp(B)	
(Intercept)	-29.771	.1185	-30.004	-29.539	63130.167	1	.000	1.176E-13	9.323E-14
[source=search ]	-.084	.0737	-.229	.060	1.307	1	.253	.919	.796
[source=referral]	-.116	.0943	-.300	.069	1.502	1	.220	.891	.741
[source=direct ]	-.017	.0742	-.162	.129	.050	1	.823	.984	.850
[source=campaign]	0 <sup>a</sup>	.	.	.	.	.	.	1	.
[location=PUNE ]	.748	.2934	.173	1.323	6.494	1	.011	2.112	1.188
[location=other ]	1.710	.1061	1.502	1.918	260.034	1	.000	5.531	4.493
[location=NEWDELHI ]	1.655	.1077	1.444	1.866	236.283	1	.000	5.234	4.238
[location=MUMBAI ]	1.733	.1047	1.527	1.938	273.997	1	.000	5.656	4.607
[location=CHENNAI ]	1.074	.1584	.763	1.384	45.947	1	.000	2.926	2.145
[location=CHANDIGARH ]	1.834	.2214	1.400	2.268	68.592	1	.000	6.257	4.054
[location= ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.
page_load_sum	.010	.0004	.009	.011	574.321	1	.000	1.010	1.009
[product_page=yes]	2.269	.0598	2.152	2.386	1440.912	1	.000	9.673	8.603
[product_page=no ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.
[browser=ucbrowser]	23.762	.1224	23.523	24.002	37713.120	1	.000	20887133999.753	16433298041.499
[browser=safari ]	23.170	.0476	23.077	23.264	236746.747	1	.000	11556154370.194	10526373395.143
[browser=opera ]	23.001	.2797	22.453	23.549	6764.205	1	.000	9755933380.670	5639180008.281
[browser=ie ]	22.628	.1440	22.346	22.911	24694.671	1	.000	6720332308.708	5067813290.206
[browser=gc ]	23.587 <sup>b</sup>	.	.	.	.	.	.	17524802364.234	.000
[browser=ff ]	0 <sup>a</sup>	.	.	.	.	.	.	1	.
[product_page=yes] *	-.007	.0007	-.008	-.006	115.333	1	.000	.993	.992
page_load_sum									
[product_page=no ] *	0 <sup>a</sup>	.	.	.	.	.	.	1	.
page_load_sum									
(Scale)	.076 <sup>c</sup>								
(Negative binomial)	1 <sup>d</sup>								

Dependent Variable: post\_sum

Model: (Intercept), source, location, page\_load\_sum, product\_page, browser, product\_page \* page\_load\_sum

a. Set to zero because this parameter is redundant.

b. Hessian matrix singularity is caused by this parameter. The parameter estimate at the last iteration is displayed.

c. Computed based on the deviance.

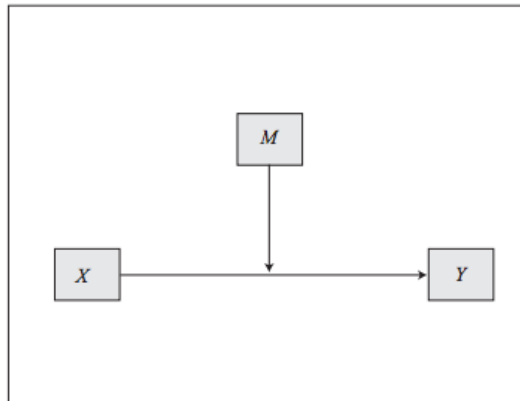
d. Fixed at the displayed value.

### J.2.3 OLS Regression Model with Moderation

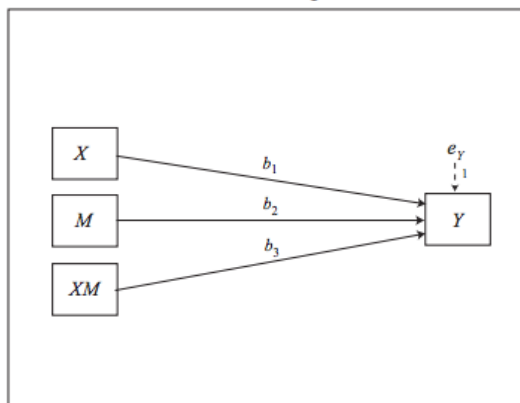
An ordinary OLS regression model with the PROCESS macro for SPSS (Hayes, 2013) and standardised variables was used, in order to confirm the effect of users clicking to the product page on the number of new listings they post. Clicking to the product page (yes or no) was tested as moderator for the effect of being in the Suggest or Know variant (yes or no, independent variable) on the number of posted listings (dependent variable), controlling for the number of page views (covariate). The moderator was a highly significant predictor.

**Model 1**

Conceptual Diagram



Statistical Diagram



Conditional effect of  $X$  on  $Y = b_1 + b_3M$

\*\*\*\*\* PROCESS Procedure for SPSS Release 2.13 \*\*\*\*\*

Written by Andrew F. Hayes, Ph.D.      www.afhayes.com  
Documentation available in Hayes (2013). www.guilford.com/p/hayes3

\*\*\*\*\*

Model = 1  
Y = Z\_Posts  
X = SA  
M = Click

Statistical Controls:  
CONTROL= Z\_PVs

Sample size  
51103

\*\*\*\*\*

Outcome: Z\_Posts

Model Summary	R	R-sq	MSE	F	df1	df2	p
	.1311	.0172	.9829	15.4036	4.0000	51098.0000	.0000

Model	coeff	se	t	p	LLCI	ULCI
constant	-.0118	.0039	-2.9917	.0028	-.0195	-.0041
Click	.0635	.0068	9.2809	.0000	.0501	.0769
SA	.0279	.0103	2.7174	.0066	.0078	.0480
int_1	2.3979	.4255	5.6353	.0000	1.5639	3.2319
Z_PVs	.0701	.0133	5.2796	.0000	.0441	.0962

Interactions:

int_1	SA	X	Click
-------	----	---	-------

\*\*\*\*\*

Conditional effect of X on Y at values of the moderator(s):

Click	Effect	se	t	p	LLCI	ULCI
-.0150	-.0080	.0089	-.9059	.3650	-.0254	.0093
.9850	2.3898	.4245	5.6297	.0000	1.5578	3.2219

\*\*\*\*\* ANALYSIS NOTES AND WARNINGS \*\*\*\*\*

Level of confidence for all confidence intervals in output:  
95.00

NOTE: The following variables were mean centered prior to analysis:  
SA      Click

NOTE: All standard errors for continuous outcome models are based on the HC3 estimator

----- END MATRIX -----

### ***Congratulations!***

You are one of the honourable few who have read my thesis until the bitter end.

You have just earned yourself 13,293,304.5 points, a badge and the number 1 position in the top 3 of 'Bauke's-thesis-readers-leaderboard'. Isn't that great?



#### **Bauke's-thesis-readers-leaderboard**

1. <your name here>
2. ...
3. ...