The Role of Prior Experience in User's Engagement with a New Recommender System

Marcelo G. Armentano
ISISTAN Research Institute
Campus Universitario
Tandil, Argentina
marcelo.armentano@
isistan.unicen.edu.ar

Silvia N. Schiaffino
ISISTAN Research Institute
Campus Universitario
Tandil, Argentina
silvia.schiaffino@
isistan.unicen.edu.ar

Analía A. Amandi
ISISTAN Research Institute
Campus Universitario
Tandil, Argentina
analia.amandi
@isistan.unicen.edu.ar

ABSTRACT

In this article we performed an online experiment to test the role of the user's prior experience with recommender systems in his/her engagement with a new recommendation technology. Our research model is based on the technology acceptance model (TAM) with two new constructs corresponding to the simplicity of the graphical user interface and the skills the user believe he/she has to use recommender systems. Our experiments confirmed the hypothesis that prior user experience plays an important role in which factors affect the user engagement with a new system.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]

General Terms

Human Factors; Experimentation

Keywords

Recommender Systems; User Engagement; TAM

1. INTRODUCTION

According to the Technology Acceptance Model (TAM) [2], the main determinants of users' engagement with an information system are the *perceived usefulness (PU)* and the *perceived ease of use (PEOU)*. PU is also seen as being directly impacted by PEOU. TAM has been applied to recommender systems in different domains such us personality-based recommendations [4], learning companion recommendations [1] and travel information [5]. We believe that, in addition to TAM constructs, the skills the user's believe he/she has to use recommender systems (Skills) and a simple design of the graphical user interface (SimpleGUI) are factors that should also be considered in the assessment of the user's engagement with the system. Finally, we hypothesize that the user's prior experience with recommender systems acts as a moderator variable. In this article, we performed an empirical user study on a movie recommender system measuring different aspects that might affect the users' engagement. Results indicate that the believed skills affect the PEOU and that a simple user interface affects the PU. However, when considering the previous user experience, these relationships among constructs are different.

2. RESEARCH MODEL

Figure 1 shows our research model. We measured the *Skills* construct with questions regarding to the ability of users to use the system, the awareness of the user self-preferences and the ability to evaluate the recommendations provided. The PEOU construct was measured by the ease of interaction, ease of use and ease of learning. The PU construct was assessed by the attractiveness, satisfaction and accuracy of the recommendations, the suitability to the users mood and tastes, the accuracy of the underlying technology, the understanding of the user preferences and receiving recommendations as good as those that the users could receive from a friend. The *SimpleGUI* construct was assessed by the absence of confusing links and irrelevant features and the ease of the registration process. Finally, the engagement was assessed by the intention to own the recommended items and the intention to return to the system. Our research hypothesis is that the paths of the model are different for users with high experience than for users with low experience.

3. METHODOLOGY

We asked students and researchers in different areas from two different universities in Argentina to use a new recommender system[1] and to fill a post treatment survey with questions associated to the variables of our model. Each question could be answered in a Likert-5 scale with 1 corresponding to "strongly disagree" and 5 corresponding to "strongly agree". The resulting dataset[2] consisted of 153 cases, collected from December 2013 to June 2014. Most users (69.9%) were male and 30.1% were female; 73.2% were in the 20-29 age group, 20.3% in the range 30-39 and 6.5% over 40 years old.

---

[1]http://pelisquemegustan.inductia.com/join
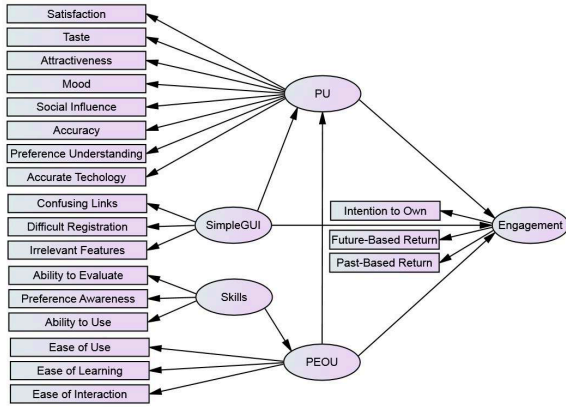[2]http://marcelo.armentano.isistan.unicen.edu.ar/datasets

Figure 1: Research model

3.1 Exploratory Factor Analysis

We performed Principal Component Analysis with Varimax rotation to extract factors from the observed variables. We found that variables loaded significantly only on one of five different factors extracted (eigenvalues values greater than 1), demonstrating sufficient *discriminant validity*. Together, they account for 73.48% of the variability in the original variables, with a *Kaiser-Meyer-Olkin* measure of sampling adequacy of 0.85 suggesting that five latent factors are representative. On the other hand, *Bartlett's test of sphericity* was highly significant ($p<0.001$) and therefore factor analysis is appropriate. The factors extracted also demonstrated sufficient *convergent validity*, as their loadings were all above the recommended minimum threshold of 0.45 for samples size of 150 [3]. Finally, the *reliability* of data can be considered to be sufficient since the Cronbach-α coefficient on different factors resulted in values greater than 0.7 (0.92 for PU, 0.78 for PEOU, 0.78 for SimpleGUI, 0.91 for Skills and 0.85 for engagement).

3.2 Confirmatory Factor Analysis

Regression Evaluation for the constructs was performed using AMOS, resulting all estimates significant at $p<0.001$ level. We performed model fit, obtaining the following indicators: cmin/df: 0.328, CFI: 1.00, RMSEA: 0.00, PCLOSE: 0.88. The model is then within the acceptable range of fitting according to the guideline thresholds proposed in [3].

To test for convergent validity we computed the Average Variance Extracted (AVE). For all factors, the AVE was over 0.5, which is the recommended threshold [3]. By comparing the square root of the AVE to all inter-factor correlations we found that all factors demonstrated adequate discriminant validity (the diagonal values are greater than the correlations). The Composite Reliability (CR) for all factors was also over the minimum threshold of 0.70 [3], indicating we have reliability in our factors.

Finally, we conducted categorical and metric invariance tests dividing the dataset according to the users' previous experience. The model fit of the unconstrained measurement models (with groups loaded separately) had adequate fit: cmin/df: 1.466, CFI: 0.939, RMSEA: 0.39, PCLOSE: 0.999. These values indicate that the model has configural invariance. After constraining the models to be equal, we found the chi-square difference test to be non-significant

Table 1: Differences for Low vs. High Experience

	Low Exp.		High Exp.		
	Est.	p	Est.	p	z-score
PEOU←Skills	0.16	0.216	0.58	***	2.15**
PU←SimpleGUI	0.14	0.015	0.10	0.162	-0.48
PU←PEOU	0.16	0.166	0.44	***	1.73*
Engage.←PU	0.81	***	0.81	***	0.00
Engage.←PEOU	0.03	0.793	0.40	***	2.40**
Engage.←SimpleGUI	0.20	0.001	0.03	0.598	-2.65***

Notes: *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.10

($p>0.05$). Thus, our measurement model meets criteria for metric invariance across the experience level in the use of recommender systems.

3.3 Hypothesis testing

Regression Evaluation was performed using AMOS and all regression weights resulted significant when considering the complete set of users involved in the experiment. To test our categorical moderation hypothesis, we produced the critical ratios for the differences in regression weights between both groups. From these critical ratios we computed p-values to determine the significance of the differences. The results are summarized in Table 1 and discussed in Section 4.

4. DISCUSSION

From our experiments, we found that there is no significant difference among groups for the effects of SimpleGUI on PU and PU on Engagement. On the other hand, for those users with low experience in the use of recommender systems the positive effect of Skills on PEOU, PEOU on PU and PEOU on Engagement is not significant while the effect is stronger and significant for users with high experience. Finally, the effect of SimpleGUI on Engagement is stronger and significant for users with low experience while it is not significant for users with high experience.

All these findings can help developers to focus their attention on factors that encourage the use of recommender systems.

5. REFERENCES

[1] H.-C. Chen, C.-C. Hsu, C.-H. Chang, and Y.-M. Huang. Applying the technology acceptance model to evaluate the learning companion recommendation system on facebook. In *IEEE 4th Int.Conf. on Technology for Education*, pages 160–163, July 2012.

[2] F. D. Davis. *A technology acceptance model for empirically testing new end-user information systems : theory and results*. PhD thesis, MIT, 1986.

[3] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson. *Multivariate data analysis*. Prentice Hall Higher Education, 2010.

[4] R. Hu and P. Pu. Acceptance issues of personality based recommender systems. In *Proc.of ACM RecSys '09*, pages 221–224, New York, NY, USA, 2009. ACM.

[5] Y.-H. Hung, P.-C. Hu, and W.-T. Lee. Improving the design and adoption of travel websites: An user experience study on travel information recommender systems. In *Proc.5th Int.Congress of Int.Assoc.of Societies of Design Research*, Tokio, Japan, Aug. 2013.