

# Users' Decision Behavior in Recommender Interfaces: Impact of Layout Design

Li Chen and Ho Keung Tsoi  
Department of Computer Science  
Hong Kong Baptist University  
Hong Kong, China

{lichen, hktsoi}@comp.hkbu.edu.hk

## ABSTRACT

Recommender systems have been increasingly adopted in the current Web environment, to facilitate users in efficiently locating items in which they are interested. However, most studies so far have emphasized the algorithm's performance, rather than from the user's perspective to investigate her/his decision-making behavior in the recommender interfaces. In this paper, we have performed a user study, with the aim to evaluate the role of layout designs in influencing users' decision process. The compared layouts include three typical ones: list, grid and pie. The experiment revealed significant differences among them, with regard to users' clicking behavior and subjective perceptions. In particular, pie has been demonstrated to significantly increase users' decision confidence, enjoyability, perceived recommender competence, and usage intention.

## Categories and Subject Descriptors

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Design, Experimentation, Human Factors.

## Keywords

Users' decision behavior, recommender system, interface layout, user study.

## 1. INTRODUCTION

Although recommender systems have been popularly developed in recent years as personalized decision support in social media and e-commerce environments, more emphasis has been placed on improving algorithm accuracy [10], and less on studying users' actual decision behavior in the recommender interfaces. On the other hand, according to user studies conducted in other areas, users will likely adapt their behavior when being presented with different information presentations. For instance, in a recent study done by Kammerer and Gerjets, the presentation of Web search engine results by means of a grid interface seems to prompt users to view all results at an equivalent level and to support their selection of more trustworthy information sources [7]. Braganza et al. also investigated the difference between one-column and multi-column layouts for presenting large textual documents in web-browsers [1]. They indicated that users spent less time scrolling and performed fewer scrolling actions with the multi-column layout.

Unfortunately, little is known about the impact of recommender interface's layout on users' decision-making behavior. There is also lack of studies that examined whether users would perceive differently, especially regarding their decision confidence and perceived system's competence, due to the change of layout. Thus, in this paper, we are particularly interested in exploring users' behavior in the recommender interface when it is presented with three layout designs: list, grid and pie. As a matter of fact, most of current recommender systems follow the list structure, where recommended items are listed one after another. The grid layout, a two-dimensional display with multiple rows and columns, has also been applied in some recommender sites to display the items. As the third alternative design, pie layout, though it has been rarely used in recommender systems, has been proven as an effective menu design for accelerating users' selection process [2]. The comparison among them via user evaluation could hence tell us which layout would be most desirable to optimize the recommender's benefits. That is, with the ideal layout design, users can be more active in clicking recommendations, be more confident in their choices, and be more likely to adopt the recommender system for repeated uses.

Concretely, we evaluated three layout designs from both objective and subjective aspects to measure users' decision performance. The objective measures include users' clicking behavior (e.g., the first clicked item's position, the amount of clicked items, etc.), and time consumption. Subjective measures include users' decision confidence, perceived interface competence, enjoyability, and usage intention. These measurements are mainly based on the user evaluation framework that we have established from prior series of user studies on recommenders [4,8,9]. We thus believe that they can be appropriately utilized as the standard to assess user behavior. Relative to our earlier work [5], this paper was for the first time to investigate the effect of basic layouts of recommender interfaces on users' decision process, which is also new in the general domain of recommender systems, to the best of our knowledge.

## 2. THREE LAYOUT DESIGNS

### 2.1 List Layout

As mentioned above, most existing recommender systems employ the standard one-dimensional ranked-order *list* style, where all items are displayed one after the other. For instance, MovieLens is a typical collaborative filtering (CF) based movie recommender system ([www.movielens.org](http://www.movielens.org)). In this system, items are ranked by their CF scores in the descending order and presented in the list format. The score represents the item's matching degree with the current user's interest.

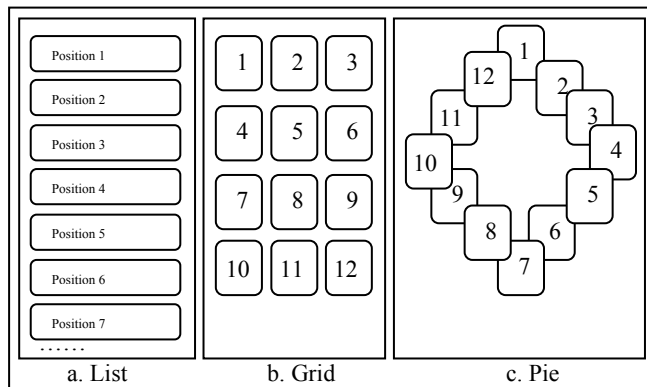
Figure 1.a shows the sample layout (where every position is for placing one item). The number of shown items varies among

existing systems. Some systems (e.g., Criticker.com) limit the number to 10 or less, while some systems (like MovieLens) give a list of items as many as possible and divide them into pages (e.g., one page displays a fixed number of items). Each item is usually described with its basic info (e.g., thumbnail image, name, rating). When users click an item, more of its details will be displayed in a separate page.

## 2.2 Grid Layout

The *grid* layout design has also been applied in some existing websites (e.g., hunch.com). In this interface, recommendations are presented in multiple rows and columns, so several items are laid out next to each other in one line. The regular presentation is to align the items horizontally (line by line). For example, as shown in Figure 1.b, the positions 1, 2, 3, ..., 12 are respectively allocated with items that are ranked 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, ..., 12<sup>th</sup> according to their relevance scores.

Because users likely shift eyes to nearby objects [6], we were interested in verifying whether the grid format would stimulate users to discover more items than in list.



**Figure 1. The three layout designs for recommender interface (the number refers to the position of a recommendation).**

## 2.3 Pie Layout

Another two-dimensional layout design is to place the items in the compass format, i.e., *pie* layout. This idea originates from the comparison of linear menu (i.e., the alphabetic ranked-order of menu choices) and pie menu [2]. In the pie menu, items are placed along the circumference of a circle at equal radial distances from the center. The distance to and size of the target can be seen as an effect on positioning time according to Fitts' law [3]. Researchers previously found that due to the decreased distance (i.e., the minimum distance needed to highlight the item as selected) and increased target size, users selected items slightly faster. The drift distance after target selection and error rates were also minimized.

We thus believe that the pie layout could offer a novel alternative and potentially more effective design to be studied. The reason is that it would support users to have a quicker overview of all displayed items, as the interface consumes greater width but less height. In addition, it would allow users to click items faster, because the mean distance between items is reduced.

When we concretely implemented this interface, we adhere to the regular clockwise direction to display the items along the circle, with the most relevant item placed at the first position (see Figure 1.c).

## 3. PROTOTYPE IMPLEMENTATION

We implemented a movie recommender system with the three layout versions. The recommending mechanism is primarily based on the hybrid of tag suggestions and tag-aware item recommendation [11]. Specifically, based on the user's initial tag profile, the system will first recommend a set of tags from other users as suggestions to enrich the new user's profile. In the mean time, a set of movie items with higher matching degree with the user's current tag profile is returned as item recommendations. If the user modifies her/his profile, the set of recommendations will be updated accordingly. More concretely, the control flow of the system works in the following four steps:

**Step 1.** To begin, the new user is asked to specify a reference product (e.g., a favorite movie) as the starting point. The product and its associated tags (as annotated by other users) are then stored in the user's profile. Alternatively, s/he can directly input one or more tag(s) for building her/his initial profile.

**Step 2. Profile-based Item Recommendation.** Based on the profile, the system generates a set of item recommendations (i.e., movies in our prototype) to the user via the weighted combination of FolkRank and content-based filtering approaches. Specifically, FolkRank transforms the tripartite graph found in the folksonomic systems into the two-dimension hyper-graph. In parallel, the content-based filtering approach rank items based on the correlation between the content of the items (i.e., title, keywords, and user-annotated tags) and the user's current profile. A tuning parameter is dynamically set to adjust the two approaches' relative weights in producing the top  $k$  recommendations.

**Step 3. Tag recommendation.** In the recommender interface, the system not only returns item recommendations, but also a set of tags to help users further enrich their profile if they need. To generate the tag recommendation, we first deployed the Latent Dirichlet Allocation (LDA), which is a dimensionality reduction technique, to extract common topics among all user tags in the database. Each topic represents a cluster, and all the extracted clusters were then applied to match with the current user's tag profile. New tags from the best matching clusters are then retrieved as recommended tags to the user. These tags' associated items are also integrated into the process of generating item recommendations in the next cycle if any of them were selected by the user. Moreover, the tag recommendations were grouped into three categories in the interface: *factual tags* (i.e., the tag describes a fact of the item, "rock"), *subjective tags* (the people's opinion, "cool") and *personal tags* (used to organize the user's own collection, e.g., "my favorites"). The grouping is automatically performed. For example, if the tag is a common keyword in the item's basic descriptions, it is treated as *factual tags*. General Inquirer<sup>1</sup>, a content analysis program, is employed to determine whether a tag is subjective. The rest of the tags that do not belong to the first two categories are then considered to be *personal tags*.

**Step 4.** If the user has done any modifications on her tag profile, it will be used to produce a finer-grained item recommendation in the next interaction cycle (returning to step 2).

The process from Step 2 to Step 4 continues till the user selects item(s) as her/his final choice(s), or quit from the system without

<sup>1</sup> <http://www.webuse.umd.edu:9090/>

selecting any recommendations. More details about the algorithm steps can be referred to [11].

To build the prototype, we crawled 998 movies and their info (including posters, names, overall ratings, number of reviewers, directors, actors/actresses, plots, etc.) from IMDB (Internet Movie Database) site. These movies’ associated tags were extracted from MovieLens for building the tag base.

Concretely, the system returns 24 movie recommendations at a time. The 24 movies are sorted in the descending order by their relevance scores, and then divided into two pages (i.e., each page with 12 movies). The switching to the second page is through the “More Movies” button. Such design could enable us to evaluate user behavior not only in a single page, but also their switching behavior across pages (i.e., whether they click the button to view more items).

The recommended movies are presented differently in the three layout versions (see Figure 3). In the *list layout*, the 12 movies in one page are displayed in the list style, where the ranked 1<sup>st</sup> is positioned at the top, followed by the ranked 2<sup>nd</sup> one (the ranked 1<sup>st</sup> one means that the movie has the highest score among the 12 movies). In the *grid layout*, three movies are displayed along one row and four in one column. More specifically, the first row shows the ranked 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> movies from left to right, the second row is with the ranked 4<sup>th</sup>, 5<sup>th</sup>, 6<sup>th</sup> movies, and so on. In the *pie layout*, the 12 movies (each with the same target size as in grid) are presented in a clockwise direction, with the ranked 1<sup>st</sup> movie at the 12 clock’s position, 2<sup>nd</sup> at the 1 clock’s position, and so forth.

In all of the three interfaces, each movie has a poster image, name, rating, number of reviews and a brief plot. More of the movie’s details can be accessed by clicking it. A separate detail page will then show the movie’s director(s), actor/actress info, detailed plot, and give links to IMDB and trailer, etc. If users like this movie, they could click the button “My Choice” at the detail page.

There is also a profile area in the three interfaces, which allows users to modify their tag profile by selecting the system-suggested ones or inputting their own. In list and grid, it is placed on the left panel, and in pie, it is in the central part.

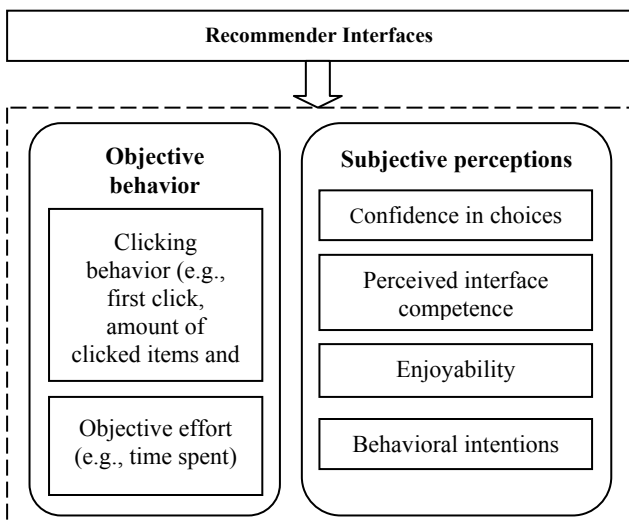


Figure 2. Objective and subjective measures in the user study.

## 4. EXPERIMENT SETUP

### 4.1 Measures

Identifying the appropriate criteria for assessing a recommender system from the user’s perspective has always been a challenging issue. Accumulated from our previous experiences on this track [4,8,9], a set of measures have been established. The framework not only includes objective interaction effort that users have spent with the system (e.g., time consumption), but also users’ perceived confidence in choices that they made in the recommender and their intention to repeatedly use the system. More specifically, in this experiment, in order to in depth identify the three layouts’ respective effects on user behavior, we assessed the following aspects (see Figure 2).

#### 4.1.1 Objective Measures

The objective measures mainly include quantitative results from analyzing users’ actual behavior in using the interface. Concretely, they cover two major aspects.

*Clicking behavior.* It has been broadly recognized that users’ clicking decisions on the recommender interface (i.e., clicking an item to view its detailed info) reflects their interest in the item. Therefore we recorded users’ clicking behavior and clicked items’ positions. The goal was to evaluate whether the clicking would be influenced by the layout, and which interface could support users to easily find interesting items. Specifically, the clicking behavior was analyzed via three variables: 1) the users’ *first clicked item’s position*, from which we could know whether users’ first click falls on the most relevant item (as predicted by the system) or not. 2) *All clicks on distinct items that a user has made* throughout her/his session of using the interface. This variable can expose the distribution of clicks over different areas on the interface. The comparison among all users could further reveal their similar clicking pattern. In addition, the total amount of clicked items could tell us how many items interested the user when s/he was confronted with the whole set of recommendations in the respective layouts. 3) *Frequency of clicking “more movies”*. Such action indicates that users switched to the next page to view more recommended items. If the frequency is higher, one possible explanation is that users felt enjoyable while using the interface and were motivated to take the effort in viewing more items, or it is because users cannot find the interesting items at the first page. Thus, this number should be analyzed in combination with other variables, especially users’ subjective opinions on the interface, so that we could more fairly attribute it to the pros or cons of the interface.

*Objective effort consumption.* Besides above mentioned analyses on users’ clicking behavior, we also recorded the time a user spent in completing the task on the specific interface. This value can be used to represent the amount of objective effort that users exerted while using the interface. In fact, it has been frequently adopted in related literatures to be an indicator of the system’s performance [10]. However, less time does not mean that users would perceive less effort taken or have better decision quality [8]. That is why we included various subjective constructs (see the next subsection) to better understand the interface’s true merits.

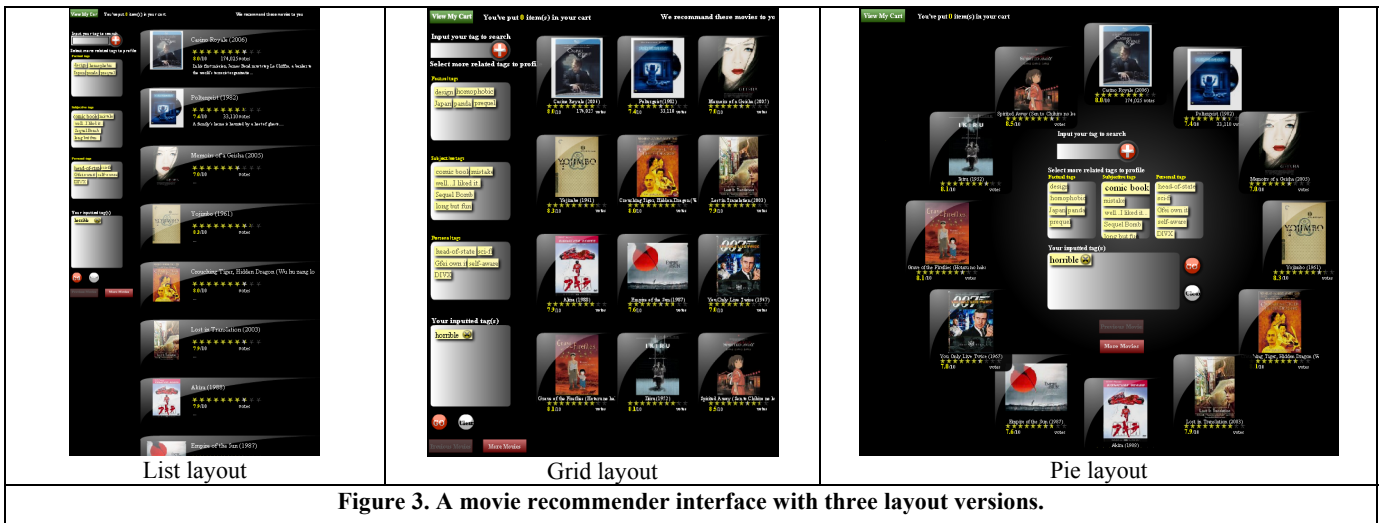


Figure 3. A movie recommender interface with three layout versions.

#### 4.1.2 Subjective Measures

Users’ decision confidence and perception of the interface were mainly obtained through the post-task survey. Actually, the subjective measures can be quite useful to expose the competence of the interface in assisting users’ decision-making and its ability in increasing users’ intention to use the system again. The variables that we have used in this experiment cover four constructs: *decision confidence*, *perceived interface competence*, *enjoyability*, and *behavioral intentions*. The perceived interface competence was qualitatively measured through multiple dimensions: users’ perception of item/tag recommendation quality, perceived ease of use of the interface in searching for info, and perceived ease of use in modifying their profile. The behavioral intention was assessed from users’ intention to use the interface again.

Table 1 lists all of the questions we used to measure these subjective variables. In the form of questionnaire, each question was required to respond on a 5-point Likert scale from “strongly disagree” (1) to “strongly agree” (5).

Table 1. Questions to measure users’ subjective perceptions

Measured variables	Question responded on a 5-point Likert scale from “strongly disagree” to “strongly agree”
<i>Decision confidence</i>	Q1: I am confident that I found the best choices through the interface.
<i>Perceived recommender interface’s competence</i>	Q2: The interface helps me find some good movies;
	Q3: This interface provides some good “tag” suggestions to help me specify criteria;
	Q4: I found it easy to use the interface to search for movies;
	Q5: I found it easy to modify my profiles in the interface.
<i>Enjoyability</i>	Q6: I felt enjoyable while using this interface.
<i>Behavioral Intention</i>	Q7: I am inclined to use this interface again.

## 4.2 Experiment Procedure and Participants

The primary factor manipulated in the experiment is *layout* as we prepared with three versions in the prototype system: list, grid, pie. To compare the three layouts, we applied the within-subjects experiment design. That is, every participant was required to evaluate all of them one by one, but the interfaces’ appearance

order was randomized in order to avoid any carryover effects (so there are six possible sequences of displaying the three layouts). To evaluate each layout, a concrete task was assigned to the user. Concretely, each layout interface was randomly associated with one scenario for the user to play the role and perform the situational task. For example, one scenario is “*This is October, the festival Halloween is coming. John is a college student, and he would like to organize an event to watch movie with his friends at his home. After discussing with his friends, they would like to watch a horror movie in this festival. John is responsible for choosing some movies as candidates. Please imagine yourself as John and use the interface to find three candidates that you would like to recommend to your friends.*” The other two scenarios were respectively for Valentine’s Day, and the military subject. In each scenario, the user was encouraged to freely use the interface to find three most suitable movies according to her/his own preferences.

The experiment was setup as an online procedure. It contains the instructions, recommender interfaces and questionnaires, so that users could easily follow and we could also automatically record all of their actions in a log file. The same administrator supervised the experiment for all participants.

A total of 24 volunteers (12 females) were recruited. 3 are with age less than 20, 1 with age above 30, and the others are between 20 and 30. Most of them are students in the university, pursuing Bachelor, Master or PhD degrees, but their studying majors are diverse. All participants had visited movie recommender sites (e.g., Yahoo movie) before the experiment, and 58.3% have even visited the indicated sites at least a few times every three months. The participants also specified the mean reasons that will motivate them to repeatedly use such a site. Among the various reasons, the ease of use of the site’s user interface was indicated as the most important factor (with the importance rate 3.83 in the range of 1 to 5). The second important factor is the site’s ability in helping them find movies that they like (3.79), followed by the site’s reputation (3.5).

## 4.3 Results

### 4.3.1 Objective Behavior

For each layout version, we first counted the number of users’ first clicks that fall on a particular position and then classified them into areas. Specifically, in one interface, each area contains three adjacent positions (e.g., 1-3 positions compose the first area, 4-6 form the second area, and so on). Areas 5 to 8 refers to the positions at the second recommendation page of the interface.

Figure 4 shows the actual distribution. In total, 8, 10, and 8 users have clicked item in the first area respectively in list, grid and pie interfaces. Then in the list and pie, there exists a linear drop from areas 1, to 2, then to 3. In area 4, the list's curve returns to the same level of area 2, but in pie it goes much higher even beyond the level of area 1. In grid, a sharp drop appears from areas 1 to 2. Then the curve rebounds and reaches to a level equivalent in areas 3 & 4. Another interesting finding is that there are 3, 2, and 1 of users' first clicks were at the second page respectively in list, grid and pie (i.e., in areas 5 to 8). To rank these areas by the amounts of first clicks, we can see that the hotter areas in list are 1, 2 & 4. In grid, they are 1, 3 & 4, and in pie, they are 1 and 4.

To further investigate the hot areas throughout a user's whole interaction session, we counted her/his total clicks made on each interface. The average numbers of items clicked by a single user are 3.96, 3.875, and 4.84 in list, grid and pie respectively. The difference between grid and pie is even marginally significant ( $p = 0.076$ ,  $t = -1.86$ , by paired samples t-test). The exact distribution of the average user's clicks among the eight areas is shown in Figure 5, from which we can see that above 50% of a user's clicks on list were in areas 1 (28.42%) and 4 (27.37%), followed by areas 3 and 2. In grid and pie, the two hotter areas are also 1 and 4, but the comparison regarding areas 2 and 3 shows that the clicks on them are more evenly distributed in pie (respectively 17.24% and 18.10%), which in fact also has higher total amount of clicks than in grid.

Moreover, the clicking distribution across pages 1 and 2 is significantly different among the three interfaces. More clicks appeared in grid's second page (24.73% accumulated from areas 5 to 8), and pie's (19.83%), against 7.37% in list. This finding suggests that grid and pie might more likely stimulate users to click the "More Movies" button for viewing more recommended items. In this regard, we further found that 50% of users have actually gone to the second page while using grid, followed by 41.7% users who did so in pie, and 25% in list ( $p = .056$  between grid and list,  $t = -2.01$ ).

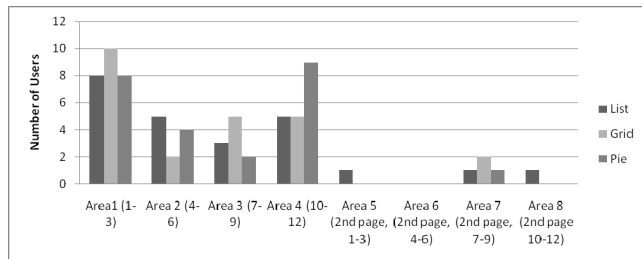


Figure 4. The distribution of users' first clicks.

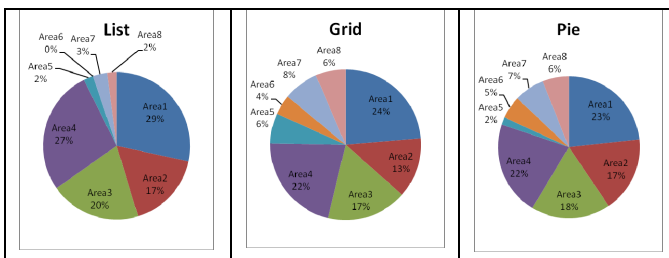


Figure 5. The distribution of an average user's whole clicks during her interaction session with an interface.

As for the total time spent on each interface, on average, it is 156.375 seconds in list, 109.875 seconds in grid, and 152.667 in

pie. Though it took longer in list and pie, the differences are not significant ( $p > 0.1$  by ANOVA and three pairs of t-test).

### 4.3.2 Subjective Perceptions

Besides measuring users' objective behavior, we were driven to further understand their subjective perceptions such as decision confidence, perceived ease of use of the recommender interface, and intention to use it again in the future, as described in Section 4.1.2.

Significant differences were found in respect of these subjective measures (see Table 2). First of all, most of users were confident that they found the best choices through pie. The mean score is 3.54 which is marginally significantly higher than the average in list (vs. 3.125,  $p = .076$ ,  $t = -1.85$ ). The grid's score is in between (3.33). Secondly, due to the change of layout, users perceived pie more competent in helping them find good movies (3.58 vs. 3.29 in grid,  $p = .09$ ,  $t = -1.77$ ; list: 3.33), easier to use (3.5 in pie against 3 in list,  $t = -2.77$ ,  $p = .01$ ; the difference between grid and list is also marginally significant: 3.375 vs. 3,  $p = .095$ ), and easier to modify their profile (3.375 in pie vs. 3.04 in list,  $t = -1.88$ ,  $p = .07$ ). Moreover, users rated higher on pie's ability in providing good tag suggestions (3.46 in pie vs. 3 in list,  $t = .241$ ,  $p = .02$ ; vs. 2.9 in grid,  $t = -2.25$ ,  $p = .03$ ). They also felt more enjoyable while using pie than list (3.42 against 2.875 in list,  $t = -2.72$ ,  $p = .01$ ; grid: 3.12). The median and mode values are also reported in Table 2.

Table 2. Users' subjective perceptions with the three layouts (L: List; G: Grid; P: Pie)

	Mean (st.d)			Median			Mode		
	L	G	P	L	G	P	L	G	P
Q1	3.125 (.85)	3.33 (.92)	3.54* <sup>L</sup> (.78)	3	3.5	4	3	4	4
Q2	3.33 (.82)	3.29 (1.04)	3.58* <sup>G</sup> (.72)	3	3.5	4	4	4	4
Q3	3 (.88)	3.375* <sup>L</sup> (.92)	3.5* <sup>L</sup> (.88)	3	3	4	3	3	4
Q4	3 (.98)	2.92 (1.02)	3.46* <sup>L,G</sup> (.88)	3	3	3.5	4	3	4
Q5	3.04 (.91)	3.17 (.92)	3.375* <sup>L</sup> (.92)	3	3	4	3	3	4
Q6	2.875 (.74)	3.17 (.96)	3.42* <sup>L</sup> (.93)	3	3	3.5	3	4	4
Q7	2.92 (.83)	3.17 (.92)	3.29* <sup>L</sup> (.95)	3	3	3.5	3	3	4

Note: Asterisks denote highly or marginally significant differences to the respective abbreviated interfaces (by paired samples t-test).

### 4.3.3 User Comments

At the end of the study, we also asked each user to give some free comments on the interfaces. 9 users explicitly praised pie. As quoted from their words, "it is easy for me to see all without scrolling the page", "easy, clear, more information", "easy to use", "no need to loop around as the movies are all in the middle", etc. Similar preference was also given to grid: "I can get a glimpse of all movies within a page", "the layout of displaying movie is good for browsing", "it lists more movies", "the item displayed clearly, and no need to scroll up or scroll down for watching the information". Thus, the obvious advantage of pie and grid, as user perceived, is that they allow them to easily see many choices without scrolling and facilitate them to browse and seek info. On the other hand, the comments to list were mainly negative (as stated by 14 users): "find the movie difficultly", "need to scroll down", "not easy to use", "I can't see all suggested movies at once", "too long inefficient take effort to scroll", etc. Therefore, the frequent reason behind users'

disliking is that the list is not easy for them to see all suggested movies and demands more effort.

## 5. CONCLUSIONS AND FUTURE WORK

In conclusion, this paper reports our in-depth studying of users' decision behavior and attitudes in different recommender interface layouts. Specifically, we compared three typical layout designs: list, grid and pie. The results revealed that in list and grid, users' first clicks largely fall in the top three positions, but in pie they also came to other areas. The distribution of an average user's whole set of clicks in an interface further showed that though the top three positions (i.e., the area 1) and the last three positions (i.e., the area 4) are commonly popular in the three layouts, the clicks are more evenly distributed in pie among all areas at its first page. Grid and pie are even more active in stimulating users to click items in the next recommendation page. From subjective measures and user comments, we found that users did prefer using pie and grid to list. Moreover, pie has been demonstrated with significant benefits in increasing users' decision confidence, perceived interface competence, enjoyability, and usage intention.

For our future work, we will conduct more user studies, including eye-tracking experiments, to track users' eye-movement behavior in the recommender interfaces. Another interesting topic will be to investigate the interaction effect from items' relevance ordering with the layout. That is, when the ordering was changed (i.e., reversed ascending order instead of regular descending order), would users' behavior be influenced or not? In fact, with the varied ordering condition, we are able to identify whether users would spontaneously evaluate the item's relevance, or their selection behavior would be largely influenced by the layout. For example, in the list interface, would they still select items at the top though they are least relevant? The relative role of layout against the relevance ordering could be hence revealed.

## 6. REFERENCES

- [1] Braganza, C., Marriott, K., Moulder, P., Wybrow, M. and Dwyer, T. Scrolling behaviour with single- and multi-column layout. In *Proc. WWW 2009*, 831-840.
- [2] Callahan, J., Hopkins, D., Weiser, M. and Shneiderman, B. An empirical comparison of pie vs. linear menus. In *Proc. CHI 1988*, ACM, 95-100.
- [3] Card, S. K., Newell, A. and Moran, T. P. *The Psychology of Human-Computer Interaction*. L. Erlbaum Assoc. Inc., Hillsdale, NJ, USA, 1983.
- [4] Chen, L. and Pu, P. Evaluating critiquing-based recommender agents. In *Proc. AAAI 2006*, 157-162, 2006.
- [5] Chen, L. and Pu, P. Eye-Tracking Study of User Behavior in Recommender Interfaces. In *Proceedings of 2010 International Conference on User Modeling, Adaptation and Personalization (UMAP'10)*, 375-380, Big Island, Hawaii, USA, June 20-24, 2010.
- [6] Halverson, T. and Hornof, A. J. A minimal model for predicting visual search in human-computer interaction. In *Proc. CHI 2007*, 431-434.
- [7] Kammerer, Y. and Gerjets, P. How the interface design influences users' spontaneous trustworthiness evaluations of web search results: comparing a list and a grid interface. In *Proc. ETRA 2010*, 299-306.
- [8] Pu, P and Chen, L. Trust-Inspiring Interfaces for Recommender Systems. *Journal of Knowledge-Based Systems (KBS)*, vol. 20 (6), 542-556, 2007.
- [9] Pu, P and Chen, L. A user-centric evaluation framework of recommender systems. In *Proc. RecSys '10 Workshop on User-Centric Evaluation of Recommender Systems and Their Interfaces (UCERSTI'10)*, 14-21, 2010.
- [10] Ricci, F., Rokach, L., Shapira, B., and Kantor, P.B. (Eds.) *Recommender System Handbook*. Springer, 2011.
- [11] Tsoi, H. K. and Chen, L. Incremental tag-aware user profile building to augment item recommendations. In *2nd Workshop on Social Recommender Systems (SRS'11) in CSCW'11*, 2011.