Relating Newcomer Personality to Survival and Activity in Recommender Systems

Raghav Pavan Karumur
GroupLens Research
University of Minnesota, Twin Cities
Minneapolis, MN USA
+1 – (612) 626 2057
raghav@cs.umn.edu

Joseph A. Konstan
GroupLens Research
University of Minnesota, Twin Cities
Minneapolis, MN USA
+1 – (612) 625 1831
konstan@umn.edu

ABSTRACT
In this work, we explore the degree to which personality information can be used to model newcomer retention, investment, intensity of engagement, and distribution of activity in a recommender community. Prior work shows that Big-Five Personality traits can explain variation in user behavior in other contexts. Building on this, we carry out and report on an analysis of 1008 MovieLens users with identified personality profiles. We find that Introverts and low Agreeableness users are more likely to survive into the second and subsequent sessions compared to their respective counterparts; Introverts and low Conscientiousness users are a significantly more active population compared to their respective counterparts; High Openness and High Neuroticism users contribute (tag) significantly more compared to their counterparts, but their counterparts consume (browse and bookmark) more; and low Agreeableness users are more likely to rate whereas high Agreeableness users are more likely to tag. These results show how modeling newcomer behavior from user personality can be useful for recommender systems designers as they customize the system to guide people towards tasks that need to be done or tasks the users will find rewarding and also decide which users to invest retention efforts in.

CCS Concepts
• Human-centered computing→Human computer interaction (HCI)→HCI design and evaluation methods→Usermodels. • Information systems→Information retrieval→Retrieval tasks and goals→Recommender Systems.

Keywords
newcomer retention; newcomer engagement; new users; personality; Big-Five Personality Traits; recommender systems;

1. INTRODUCTION
1.1 User Activity in Recommender Systems
From non-personalized feedback about items to personalized recommendations, recommender systems have become ubiquitous over the last decade and half. Within these systems, participation can take many forms. Users can browse through items (could be places, businesses, products, songs, movies, etc.), consume them (visit, buy, listen, watch etc.) or contribute items (add products, movies, videos or songs to the catalog). They can write reviews, rate or tag items, or provide other forms of feedback such videos (of products they use and how they find them), or pictures, or just view some of these forms of feedback. Through these explicit and implicit actions, users provide their preferences. Recommender systems learn these preferences and make suggestions.

Under-contribution [9, 12], lack of diversity in contribution [29, 52, 55], and early user dropouts [18, 29, 39, 57] are challenges that are historically plaguing many online communities, and recommender systems are no exception. Indeed, these challenges have a much greater impact on the performance of recommender systems, particularly during the cold-start phase [1, 44], for they also hinder their ability to gather sufficient (and diverse) information about users as well as items and make appropriate recommendations. In this work, we are interested in the properties of users that might help us understand their specific preferences for activity.

Our goal is to identify attributes that recommender systems operators could use to evaluate these properties and determine the users who are likely to contribute and ways they are likely to contribute to select and guide them to activities/experiences they are more likely to find fulfilling. We hope that this improves overall user satisfaction, activity, and retention, simultaneously increasing the annotations about the products.

1.2 Personality
Personality is known to explain variation in user preferences and behaviors across a variety of online [7, 15, 34, 47, 56] and offline contexts [8, 31, 32], and prior research suggests that it is a stable attribute across the human lifespan [14, 41]. Personality is commonly represented using the Five Factor Model [16, 26, 35, 51]. Therefore, we are interested in knowing how personality can model newcomer usage of a recommender.

1.3 Research Questions
In this paper, we measure usage using three classes of metrics: metrics about retention, metrics about early time investment and metrics about distribution and level of activity. Specifically, we ask the following questions:

RQ1. How is personality related to newcomer retention?

RQ2. How is personality related to newcomer investment?

RQ3. How is personality related to newcomer intensity of engagement?

UMAP'16, July 13–17, 2016, Halifax, NS, Canada

DOI: http://dx.doi.org/10.1145/2930238.2930246

We extract various high-level measures characterizing survival, engagement intensity, and level of activity of newcomers in MovieLens, a movie recommender system (movielens.org). We show how these measures correlate with users' personality, as measured by a standard Big Five-Factor Model questionnaire.

1.4 Findings and Contributions

In this paper, we find that personality information can be used to model newcomer retention, investment, intensity of engagement and distribution of activity in a community movie recommender system. In particular, we find trend evidence for the following:

- Introverts and low Agreeableness user users are more likely to survive into second and subsequent sessions compared to their respective counterparts.
- Introverts and low Conscientiousness users are a significantly more active population compared to their respective counterparts.
- High Agreeableness users are more likely to tag whereas low Agreeableness users are more likely to rate.
- High Openness users and high Neuroticism users contribute significantly more tags compared with their counterparts, but their counterparts browse and bookmark more.

2. BACKGROUND AND RELATED WORK

2.1 Newcomer Motivations, Activity and Retention

Newcomer retention and activity have been studied in a variety of online communities such as Wikis [39], Q&A sites [17, 57], Online Role Playing Games [30], social networks [10, 20] and recommender systems [21, 29]. Some of the prior work identified factors such as desires to volunteer online, help others, be social, gain reputation, develop in one's career, improve one's skills, have fun/intellectual stimulation, or having prior experience in the area motivate contribution (and lead to greater retention) in these communities. [21, 28, 37, 54]. Others also showed improved participation and engagement by use of personalized early interventions in the context of social networks [20]. However, in order to make accurate recommendations on multiple categories which is the usual case in a typical recommender system, it seems essential to capture the fundamental nature of each individual. Prior research has shown that personality can account for individual differences in attitudes, motivations, experiences, and emotions [35]. We therefore explore personality to model both retention and user activity in the system.

2.2 Personality and The Big Five Model

Research on personality traits in social psychology and computer-mediated communication since the 1990s has shown that personality can predict user preferences and behaviors in all kinds of contexts, ranging from media [32], to activities such as reading books and attending concerts [32], to appreciation for arts such as music and paintings [41, 58], to job success [8] and marital satisfaction [31], and to the amount of internet and social media usage [7, 15, 34, 43, 46]. A lot of studies focused on understanding how internet usage varied among people with different personality type and we summarize them below under each personality type.

The Big Five Model on Personality, also known as the Five Factor Model is a well-researched and widely accepted model of personality traits and is commonly used in studies examining personality and human behavior [16, 26, 35, 51]. This model has been found to be reliable after testing across multiple languages and cultures [45]. The Five Dimensions of this Model, often abbreviated using the acronym OCEAN are:

Openness (to experience): High Openness people tend to be characterized by higher creativity, imagination and ability to ideate. They possess greater intellectual curiosity and appreciation for novelty or variety in experiences and diversity in interests. Low Openness users are more down-to-earth and conservative.

Prior work found positive correlations for Openness and use of internet for entertainment [53] and games [48, 53]. This may be due to their proclivity for new experiences and variety and curiosity. It was also found in [46] that high Openness users stayed online longer. Others found that Openness to experience was positively related to the use of social networking sites and features such as instant messaging [15]. High Openness is associated with an interest in more complex and exciting recreational activities [32].

Conscientiousness: High Conscientiousness people tend to be highly disciplined, organized, consistent, cautious, and dutiful in their behavior, whereas those with low Conscientiousness tend to be more impulsive, creative, easy-going, and flexible.

Several works report that high Conscientiousness is negatively correlated with general internet use and time spent online on entertainment, leisure and social networking sites [11, 33, 42]. On the other hand, high Conscientiousness is positively correlated with time spent on academic/work related sites [33]. Some researchers [7, 34] reason that conscientious people tend to have less interest in activities related to entertainment such as playing games or listening to music as they involve less planned use of time and are more spontaneous activities, which is opposed to their nature of being cautious and self-disciplined, possessing impulse control and having planned behavior [23]. Also, in [24], low Conscientiousness people were found to rate more items, whereas high Conscientiousness people were found to rate only the required number of items, and such cautious behavior is again characteristic of high Conscientiousness users. Others found that Conscientiousness was negatively related to ability to undertake difficult or unconventional activities [32].

Extroversion: Extroverts tend to be more sociable, out-going, energetic and desire the company of others and stimulation in external environments. Introverts are more reserved, self-absorbed, low-key, and seek environments in which stimulation is much lower.

Some researchers claim that Extroverts tend to prefer face-to-face interactions while Introverts tend to prefer use of online channels for self-expression [3]. Amiel et al found high Extroversion to be negatively associated with comfort in online communication [4]. Anolli et al. found a negative relationship between Extroversion and use of online chat [5]. Whereas in [13], Extroversion was negatively associated with addiction to gaming, Teng [48] found that Extroverts were significantly more into gaming compared to Introverts. Others have found positive associations between high Extroversion and the use of internet for communication and emails [50, 56] as well as more direct face-to-face friendships [50]. Yet others have also found Extroversion to be positively correlated with social network usage [43, 15]. Some found that Extroverts do a lot of liking, commenting and expressing their appreciation or sympathy for others, befriending a lot of people [7]. Others suggest that Extroverts may use the internet for more networking and Introverts may use it to escape their offline personas [38].

Agreeableness: High Agreeableness persons tend to be more cooperative, submissive, flexible, adaptable, tolerant, and empathize with others, whereas low Agreeableness persons are more competitive, challenging and tend to exercise their authority over others.

Some works did not find any relationship for Agreeableness with performance and internet use [8, 34]. Others found high Agreeableness to be negatively related to the time spent online [33], and activities such as playing online games [13, 40]. While Agreeableness was negatively associated with ability to undertake unconventional and difficult activities [32], high Agreeableness users were found to be associated with higher number of tags in [7]. It was found in [24] that high Agreeableness users tend to give ratings that are more positive.

Neuroticism: Users high in Neuroticism tend to be more sensitive, insecure, pessimistic, self-conscious, and are more susceptible to anger, frustration, anxiety, hopelessness and negative emotions. They are more likely to experience stress and depression. People with low Neuroticism, on the other hand, tend to be calmer and more emotionally stable.

Because High Neuroticism users are susceptible to a lot of negative emotions, use of the internet could provide venues to alleviate such emotions, get rid of insecurity/loneliness and find a sense of belonging. A lot of studies found high levels of Neuroticism of users to be associated with higher use and a greater amount of time spent on the internet, in particular on social networks [2, 4, 6, 11, 15, 34, 38, 42, 56]. Some researchers also found activities of leisure such as playing music or watching movies to be attractive for users with high Neuroticism [47, 56]. At the same time, other researchers found that High Neuroticism users are less likely to use the internet to seek information [2, 53]. One reason for this may be their insecurity and inability to trust any source of information. Another might be due to their nature of lacking hope and being susceptible to frustration.

Some of prior work has connected personality to rating behaviors [19, 25] in recommender systems, but we are aware of no work that specifically highlighted relationship between personality and newcomer survival, time investment, level and distribution of early user activity in a system. In this work, we are specifically interested in using personality to model newcomer retention and level of activity since newcomer survival and activity are intricately connected to community success [9, 12, 18, 29, 39, 57].

3. RESEARCH METRICS

RQ1. ***How is personality related to newcomer retention?*** We measure retention using the following metrics:

- Number of sessions at the end of first month, and at the end of the first four months.
- Odds of returning for a 2nd, 5th or 10th session[1].
- Time to first return.
- Average return time (time between sessions) during the first four months

RQ2. ***How is personality related to newcomer investment (time committed to early sessions)?*** To answer this question, we measure:

- Length of first session[2].
- Average session length for first four months of activity.

RQ3. ***How is personality related to newcomer intensity of engagement?*** We define level of activity to be number of ratings, number of tags applied, number of items the person adds to their wish list, proportion of tags to ratings, number of page views and so forth. We now measure:

- Level of activity for first-session.
- Average level of activity per session for the first 4 months.
- Aggregate (total amount of) activity for the first four months.

We look at a variety of metrics to address the three research questions. We recognize that the underlying constructs and metrics have overlap. For example, we categorize metrics such as frequency of logging in as retention, but they can also be measures of intensity of engagement. Our goal is to understand user behavior characterized by these metrics. So we have chosen a single organization for our investigation, and report the resulting data to allow others to draw further conclusions.

In the next section, we discuss the structure and properties of the MovieLens platform. We frame the hypotheses for user behavior in a system like MovieLens based on existing knowledge of personality types. We then present our findings, summarize them and draw implications from them before we conclude the paper with limitations and future directions.

4. PLATFORM, STUDY DESIGN AND METHODOLOGY

4.1 MovieLens

MovieLens (movielens.org) is a standalone movie recommendation engine which provides an opportunity for its users to express preferences through rating, tagging and wishlisting movies, while allowing them to view movie details at different levels (summary of plot, trailers, posters, etc). With more than 200,000 registered users worldwide, and an average of 50 new user registrations every day, MovieLens is a suitable platform for studying user engagement, participation, retention and commitment in recommender systems.

MovieLens is primarily used for obtaining movie recommendations based on individual taste preferences. Rating is much more common than tagging, both because ratings build user personalization profiles and because the site design permits ratings at every movie display (with a simple click) while tagging requires visiting a detail page and typing. Clicking on a movie brings up a “movie details page” with plot and cast information, the tagging interface, and various other ways to interact with the movie. Users can add movies to a wishlist anywhere they can rate them, but wishlists are a not a widely-used feature. Very rarely, some users suggest movies to be added to MovieLens through an interface for suggesting movies. MovieLens runs several recommendation algorithms, which it calls “The Peasant”, “The

[1] We choose these sessions to be consistent with prior work [29].

[2] First session in MovieLens is considerably different from other sessions as most users provide a majority of ratings during this session.

Bard", "The Warrior", and "The Wizard" and provides different kinds of recommendations depending on what the user selects as their primary recommender. Occasionally, users change their recommenders too. Our data also suggests that occasionally, users view the posters and watch the trailers on the movie details page. Since rating, tagging and wishlisting movies are the three primary activities on MovieLens, and findings on these activities are generalizable to other recommender systems, we mostly focus our analyses on these three activities. However, we do report results on the number of movie detail pages a user visits and the total number of activities the user performs (which may include all the above activities) as well, for completeness.

4.2 Dataset

In order to collect personality information for improving recommendations, Tien Nguyen of GroupLens Research administered a questionnaire based on [22] to MovieLens users during the summer of 2015. Users were asked to respond to questions assessing their personality on a Likert Scale with responses ranging from 1 (Strongly Disagree) to 7 (Strongly Agree). Based on these answers, a score for each of the five personality dimensions was computed for each user on the scale 1-7. We use the results of this survey to study retention, early time investment and activity level of new users.

Table 1. Counts of users in low and high personality types

Personality	#low users	#high users
Openness	62	430
Conscientiousness	33	228
Extroversion	222	87
Agreeableness	34	113
Neuroticism	59	213

We pick 1008 of these users, who registered between 01 July 2015 and 01 October 2015 and extract their activity log for four months along with their personality scores on the scale 1-7 for this study. MovieLens makes it optional for users to enter any profile information and so only a very small fraction of users have some information about their gender and age. We are therefore unable to report summary statistics about age groups, gender, and location of these users.

Finding effect sizes that are small is a known challenge in personality related research methods. In order to circumvent this problem, increase the sensitivity of statistical analyses used, and ensure comparability of results some researchers [43, 46] divide the personality dimensions into thirds in terms of percentiles and compare the users scoring in the higher third with the users scoring in the lower third. We realized that these approaches might have the possibility of users with similar scores (such as a score of 5 on Openness) coming in two different thirds (in this example, the middle third as well as the upper third). So, we partition the users such that those scoring less than or equal to 2 on each dimension are the low personality type, and those scoring greater than or equal to 6 are the high personality type and those with no strong preferences (scoring between 2 and 6) are the medium personality type. Most results reported in the next section are based on a comparison between the users in the low and high personality types. However, since we had too few low Openness users based on this approach to draw statistically significant conclusions, in order to explore the effect of Openness trait in a useful way, we set the threshold for low Openness at 3.5. Since 4 on the Likert scale corresponds to 'Neither Agree Nor Disagree', 3.5 for Openness has the same directional effect as 2. However, since our goal is to also optimize the sensitivity of our analyses, we retained the lower threshold of 2 for the remaining four personality types. We report the counts of users with low and high personality types in Table 1.

4.3 Hypotheses

Based on the existing knowledge about personality and user behavior, we frame the following hypotheses for newcomer behavior in the context of MovieLens[3]:

4.3.1 Hypotheses for Openness:

Because Openness is characterized by a tendency to seek variety, and a system like MovieLens offers a diverse collection of movies for users to keep returning, we expect high Openness users to last longer. Because Openness is positively associated with use of internet for entertainment and games [48, 53] and MovieLens does not offer movies to watch, we expect high Openness users to invest shorter durations of time in their visits, maybe just enough to find movies for watching. Because creative activities excite high Openness users [32] and tagging exercises one's creativity we expect high Openness users to tag more. Because high Openness users have greater curiosity and a desire for entertainment, we expect them to have already watched a lot of movies and therefore add less movies to their wish lists compared to low Openness users. Because curiosity is characteristic of Openness users, we expect them to visit more movie detail pages.

O1: Openness is positively correlated with likelihood of retention.

O2: Openness is negatively correlated with time investment per session.

O3: Openness is positively correlated with tagging movies.

O4: Openness is negatively correlated with wishlisting movies.

O5: Openness is positively correlated with visiting movie detail pages.

4.3.2 Hypotheses for Conscientiousness:

Because Conscientiousness is characterized by self-discipline and planned behavior, we expect high conscientious users to be more judicious with the amount of time they spend on a site aimed at entertainment. So, we expect lower activity and lower number of movie detail views from high Conscientiousness users who are less spontaneous and easy-going. Prior work [24] found evidence for negative correlation between Conscientiousness and rating items, and Conscientiousness and ability to undertake difficult activities. So we have the following hypotheses in relation to Conscientiousness:

C1: Conscientiousness is negatively correlated with likelihood of retention.

C2: Conscientiousness is negatively correlated with time investment per session.

C3: Conscientiousness is negatively correlated with rating movies.

C4: Conscientiousness is negatively correlated with tagging movies.

---

[3] We do not state all possible combinations of hypotheses for each personality type because nothing we know of their nature suggests an expected behavior for certain actions for some personality types.

C5: Conscientiousness is negatively correlated with wishlisting movies.

C6: Conscientiousness is negatively correlated with visiting movie detail pages.

C7: Conscientiousness is negatively correlated with aggregate activity per session.

4.3.3 Hypotheses for Extroversion:

Prior work suggests that extroverts primarily enjoy environments which stimulate them and so would show positive associations in online environments that are social, and help them network or compete with others, but otherwise have negative correlations with online activity in standalone systems like MovieLens. So we make the following hypotheses:

E1: Extroversion is negatively correlated with likelihood of retention.

E2: Extroversion is negatively correlated with time investment per session.

E3: Extroversion is negatively correlated with rating movies.

E4: Extroversion is negatively correlated with tagging movies.

E5: Extroversion is negatively correlated with wishlisting movies.

E6: Extroversion is negatively correlated with visiting movie detail pages.

E7: Extroversion is negatively correlated with aggregate activity per session.

4.3.4 Hypotheses for Agreeableness:

Because high Agreeableness is associated with a tendency to trust others [27], we expect more consumption behavior from high Agreeableness users. Because low Agreeableness persons tend to exercise their authority over others, we expect them to actively critique and thus contribute to activities such as rating and tagging movies. Since MovieLens is primarily a rating system, we expect low Agreeableness users to stay longer and offer their critiques. So, we have the following hypotheses in relation to Agreeableness:

A1: Agreeableness is negatively correlated with likelihood of retention.

A2: Agreeableness is negatively correlated with time investment per session.

A3: Agreeableness is negatively correlated with rating movies.

A4: Agreeableness is negatively correlated with tagging movies.

4.3.5 Hypotheses for Neuroticism:

Neuroticism is associated with insecurity and loneliness and a tendency to seek a sense of belonging. So prior work found Neuroticism to be positively related to time spent on social networks and sites with leisure activities such as playing games or watching movies. Since MovieLens is only a movie recommender, we don't necessarily expect any relation to time spent online. Since high Neuroticism users are insecure, there may be a tendency to exercise their opinion on a group of people. So we expect positive correlation with activities such as rating and tagging which annotate the system's items. Since high Neuroticism users often change their mood, it may be hard to understand their wishlisting behavior and we hypothesize that low Neuroticism users or emotionally stable users have higher activity on tasks such as wishlisting movies. High Neuroticism users are known to be not good at information-seeking, a behavior that may be likely due to their inability to trust any source of information [2, 53]. We therefore expect negative correlation to browsing pages about movie details:

N1: Neuroticism is positively correlated with rating movies.

N2: Neuroticism is positively correlated with tagging movies.

N3: Neuroticism is negatively correlated with wishlisting movies.

N4: Neuroticism is negatively correlated with visiting movie detail pages.

4.4 Method

To validate the hypotheses, we compute several metrics at several points in time. Due to space constraints, we report only a few of them that typified our results in this paper.

We use the term 'session' to mean a normal login period that begins with the user signing in and ends with the user logging out or with the expiration of the cookie. However, since most users multitask (use multiple tabs and switch between them), they make it harder to record their true session length as there is no explicit logout action in the MovieLens data log. So we computed session lengths explicitly as the differences between their first recorded activity and their last recorded activity per unique session ID.

The samples in the low and high groups, although independent are not necessarily normally distributed. So, we use the Wilcoxon-Mann-Whitney-test to determine whether the users in the low and the high personality type groups differ significantly in terms of their behavior in relation to the metrics listed in the research questions section. In the cases where one of the groups has a lot of zeros for the metric under consideration (this is mostly the case with the number of tags or the number of movies the user adds to their wishlist), we step away from comparing low and high personality groups and use the personality scores on the original 1-7 scale. We employ the Poisson, Negative Binomial, Zero-inflated Poisson, or Zero-inflated Negative Binomial models, as appropriate, subsequently testing the assumptions for each, to draw conclusions about effect sizes. Our interpretations will therefore follow two different patterns, one directly making a comparison between high and low personality type and the other talking about the change in the metric score associated with an increase/decrease in the particular personality score. We report results that are significant (at 0.05 level) and marginally significant (at 0.1 level) in the Results section.

5. RESULTS

First we combine the findings for the three research questions and report the results grouped by each personality type.

Table 2. Summary Statistics for some of the metrics

Metric	Min	1st Q	Median	3rd Q	Max
Metrics related to newcomer retention (RQ1)					
Number of Sessions during first month	1	2	5	11	120
Number of Sessions during first four months	1	3	7	19	451
Return time for second session (in seconds)	0	8863	54960	253500	10190000
Average return time between sessions (in seconds)	0	151200	334000	780700	10190000
Metrics related to newcomer investment (RQ2)					
First session length (in seconds)	19	860	1945	3907	35860
Average session length (in seconds)	45.25	587.8	963.9	1456	7218
Metrics related to newcomer intensity of engagement (RQ3)					
Number of ratings in first session	0	28	62	134	1372
Total number of movie detail page views in first session	1	18	41	87	1753
Total number of activities during first session	1	59	119	250	3143
Total number of ratings for the first four months	0	61	143	305	6364
Total number of activities for the first four months	1	158	352	731	9833
Total number of movie detail views for the first four months	1	65	162	360	4689
Average number of ratings per session during the first four months	0	8	16	35	516
Average number of movie detail page views per session	1	11	18	31	266
Average number of activities per session	1	22	39	74	679

In Table 2, we report the five summary statistics for some of the measures we use in the results section. In this table, the minimum values for first return time and average return time are zero. Return times have been computed by subtracting the beginning time of a session from the ending time of the previous session. However, a very small proportion of users logged in simultaneously from another device while using MovieLens from one device and for these cases, our approach yields negative return times. In order to resolve this issue, we consider these users to return in "no time" and assign zeros. Also, 44 users did not return after the first session. We exclude these users for the results reported on first and average return times. The user who had the longest inter-session time had only 2 sessions resulting in the same maximum value of 10190000 sec for average return time between sessions and return time for second session.

Openness We find a trend of high Openness users having a 21% higher odds of returning for the fifth session compared to low Openness users ($p < 0.1$[4]). We also find a trend of high Openness users having sessions that are 7.2 minutes shorter than low Openness users during the first session ($p < 0.1$). A unit increase in Openness score on the scale ranging from 1 to 7 is associated with a 21% increase in the expected number of tags from them during the first session ($p < 0.05$) and a 28.3% increase in the expected number of tags from them per session on an average for all the sessions during the first four months ($p < 0.05$) supporting our hypothesis O3. We also find that a unit increase in Openness score on the scale ranging from 1 to 7 is associated with a 156% increase in the odds of producing both nonzero ratings as well as tags on the aggregate during the four month period ($p < 0.05$) and a 177% increase in the odds of producing both nonzero ratings as well as tags per session on an average during the first four months ($p < 0.05$). We find a trend of high Openness users adding an average of 58.4% of total number of movies added by low Openness users to their wish lists during the first session ($p < 0.1$).

Conscientiousness We find that low Conscientiousness users return by a median of 39.2 hours earlier for the next session on an average for all session return times during the first four months ($p < 0.05$) and also a trend of returning 5.4 hours earlier for the second session ($p < 0.1$) compared to high Conscientiousness users supporting our hypothesis C1 that low Conscientiousness users show more likelihood of retention compared to their counterparts. We find that low Conscientiousness users last longer per session by a median of 8.6 minutes on an average for all sessions during the first four months compared to high Conscientiousness users ($p < 0.05$) confirming our hypothesis C2 on time investment per session. Low Conscientiousness users rate a median of 42 more movies during the first session ($p < 0.05$), 7 more movies on an average per session for all sessions ($p < 0.05$) and 63 more movies on the aggregate for the first four months ($p < 0.05$) compared to high Conscientiousness users. These findings support our hypothesis C3 on rating movies. We do not find statistically significant difference between number of tags produced by users in the high and low Conscientiousness groups. A unit increase in Conscientiousness is associated with a 13% decrease in the number of movies wishlisted on an average per session for all sessions ($p < 0.05$) supporting our hypothesis C5. We find a trend of low Conscientiousness users viewing a median of 15 additional movie detail pages during the first session ($p < 0.1$) and a statistically significant median of 8 additional movie detail pages per session on an average for all sessions during the first four months ($p < 0.05$) compared to their counterparts. This supports our hypothesis C6 on visiting movie detail pages. We find that low Conscientiousness users perform a

[4] We note analyses with $0.05 < p < 0.1$ to provide trend information that may be useful to guide future work, but not as statistically significant results.

Table 3. Summary of findings (selected results listed for each hypothesis)

Hyp	Results	Data	Summary
RQ1. ***How is personality related to newcomer retention?***			
O1	Not Supported	High Openness users have 21% higher odds of returning ($p < 0.1$)	Marginally significant for fifth session*
C1	Supported	Low Conscientiousness users return 39.2 hours earlier ($p < 0.05$)	Significant per session on average
E1	Supported	Introverts have 33% higher odds of returning ($p < 0.05$)	Significant for fifth and tenth sessions
A1	Supported	Low Agreeableness users return earlier for a second session ($p < 0.05$)	Significant for second session
RQ2. ***How is personality related to newcomer investment (time committed to early sessions)?***			
O2	Not Supported	Sessions for High Openness users are 7.2 minutes shorter ($p < 0.1$)	Marginally significant for first session*
C2	Supported	Low Conscientiousness users last 8.6 minutes longer ($p < 0.05$)	Significant per session on average
E2	Supported	Introverts last 3.6 minutes longer ($p < 0.05$)	Significant per session on average
A2	Not Supported		Not Significant
RQ3. ***How is personality related to newcomer intensity of engagement and distribution of activity?***			
O3	Supported	21-28% more tags per unit increase in Openness score ($p < 0.05$)	Significant for first session, first four months
O4	Not Supported	Low Openness users wishlist 1.6 times more movies ($p < 0.1$)	Marginally significant for first session*
O5	Not Supported		Not Significant
C3	Supported	+42 in first session, +7 per session on average, +63 in all ($p < 0.05$)	Significant for all mentioned periods
C4	Not Supported		Not Significant
C5	Supported	13% less tags/session per unit increase in Conscientiousness ($p < 0.05$)	Significant per session on average
C6	Supported	+15 in first session, +8 per session on average ($p < 0.05$)	Significant for all mentioned periods
C7	Supported	+65 in first session, +18 per session on average ($p < 0.05$), +121 in all	Significant for mentioned periods
E3	Supported	+26 in first session, +52 in all ($p < 0.05$)	Significant for first session, first four months
E4	Supported	29% less tags/session per unit increase in Extroversion ($p < 0.05$)	Significant per session on average
E5	Not Supported	+1 additional movie ($p < 0.1$)	Marginally significant for first four months*
E6	Supported	+30 in first session, +6 per session on average, +81 in all ($p < 0.05$)	Significant for all mentioned periods
E7	Supported	+67 in first session, +10 per session on average, +156 in all ($p < 0.05$)	Significant for all mentioned periods
A3	Not Supported	+25 during first session, +45 in all ($p < 0.1$)	Marginally significant results found*
A4	Not Supported	24% more tags/session per unit increase in Agreeableness ($p < 0.05$)	Significant per session on average
N1	Not Supported	62% higher odds of nonzero ratings/session per unit increase ($p < 0.1$)	Marginally significant per session on average*
N2	Supported	16% more tags per unit increase in Neuroticism ($p < 0.05$)	Significant for first session, first four months
N3	Supported	26% decrease in wishlists per unit increase in Neuroticism ($p < 0.05$)	Significant per session on average
N4	Not Supported		Not Significant

* We saw marginally significant effects ($p < 0.1$) at this amount and we report them as trend evidence; these might deserve further investigation.

median of 65 more activities during the first session ($p < 0.05$), 18 more activities on an average per session for all sessions ($p < 0.05$) and a trend of 121 more activities on the aggregate for the first four months ($p < 0.1$) compared to high Conscientiousness users. These findings support our hypothesis C7 on overall activeness of low Conscientiousness users.

Extroversion Introverts visit more frequently by a median of 1 additional session during the first month ($p < 0.05$). We also find a trend of introverts visiting more frequently by a median of 1 additional session on the aggregate four month period ($p < 0.1$) compared to extroverts. Introverts have 34.5% higher odds of returning for the fifth session ($p < 0.05$) and 33.5% higher odds of returning for the tenth session ($p < 0.05$) compared to extroverts. We find a trend of Introverts returning a median of 3.2 hours earlier than extroverts for a second session ($p < 0.1$). All these confirm our hypothesis E1 that Introverts are more likely to retain in the community compared to extroverts. Introverts last for a median of 215 seconds more on an average per session for all sessions during the first four months compared to extroverts, supporting our hypothesis E2 on investment. Introverts rate a median of 26 more movies during the first session ($p < 0.05$) and 52 more movies on the aggregate for the first four months ($p < 0.05$) compared to extroverts, supporting our hypothesis E3 on relationship between Extroversion and rating movies. A unit increase in Extroversion on the score ranging from 1 to 7 is associated with a 40% decrease in the expected number of tags during the first session ($p < 0.05$) and a 29% decrease in the expected number of tags per session on an average for all the sessions during the 4 month period ($p < 0.05$). These findings support E4. We find a trend of Extroverts wishlisting an average of about 55.4% of the total number of movies wishlisted by Introverts during the first session ($p < 0.1$) and Introverts wishlisting a median of 1 additional movie on the aggregate during the entire four month period compared to extroverts ($p < 0.1$). Introverts view a median of 30 additional movie detail pages during the first session ($p < 0.05$), 6 additional movie detail pages on an average per session for all sessions ($p < 0.05$) and 81 additional movie detail pages on the aggregate for the first four months ($p < 0.05$) compared to extroverts supporting our hypothesis E6. Introverts perform a median of 67 additional activities ($p < 0.05$) during the first session, 10 additional activities per session on an average for all sessions during the first four months ($p < 0.05$) and 156 additional activities on the aggregate for the first four months ($p < 0.05$) compared to extroverts, supporting our hypothesis E7.

Agreeableness Low Agreeableness users show a trend of visiting more frequently (by a median of 3 sessions more) during the first month ($p < 0.1$) and having a 35% higher odds of returning for the fifth session ($p < 0.1$) compared with high Agreeableness users. We find that low Agreeableness users return for the second session 4.7 hours earlier than high Agreeableness users ($p < 0.05$). We find a trend of low Agreeableness users rating a median of 25 more movies during the first session ($p < 0.1$) and a median of 45 additional movies on the aggregate during the first four months ($p < 0.1$) compared to high Agreeableness users. A unit increase in Agreeableness is found to be associated with a 24.3% increase in the expected number of tags per session on average for all sessions during the first four months ($p < 0.05$). Here we find a direction opposite to the assertion we made for hypothesis A4. One reason for this might be that these users are mostly producing tags similar to what others have produced before just by adding existing tags, which is characteristic of Agreeableness users (to agree with others). This may also be a reason why we do not find any statistically significant relationship between Agreeableness and early time investment. Both high and low Agreeableness users might be investing in different activities (rating and tagging). Bachrach et al (2012) find Agreeableness to be a hard trait to predict using Facebook profile features and report very low R^2 for their model (0.01) [7]. Others [8, 34] do not find any relationship between Agreeableness and internet use. So, it is not surprising that many of our results are only significant at 0.1 instead of 0.05.

Neuroticism We find a trend of a unit increase in Neuroticism being associated with a 61.5% increase in the odds of having both nonzero ratings and tags per session on an average during the first four months ($p < 0.1$). A unit increase in Neuroticism on scale with scores ranging from 1 to 7 is associated with a 16.5% increase in the expected number of tags during the first session ($p < 0.05$). This finding supports our assertion in hypothesis N2 on the relationship between Neuroticism and tagging activity. A unit increase in Neuroticism is found to be associated with an average decrease of 26.4% in the number of movies wishlisted per session for all sessions during the first four months ($p < 0.05$). Low Neuroticism users wishlist a median of 2 additional movies on the aggregate for the first four months compared to high Neuroticism users ($p \sim 0.05$). These findings support our hypothesis N3. We do not find any statistically significant results to support our assertion on visiting movie detail pages. This may again be due to opposite behaviors on rating and tagging, and wishlisting.

We summarize and report selected findings grouped by the research questions in Table 3.

6. DISCUSSION

The above results suggest that different personality types use the system differently. Specifically, we find that users with certain personality types (low Extroversion, low Agreeableness) have a higher likelihood of returning to the community compared to their counterparts; users with certain other personality types (low Extroversion and low Conscientiousness) are more active in a system like MovieLens compared with their counterparts; users with some other personality types show different activity preferences (low Agreeableness users are more likely to rate and high Agreeableness users are more likely to tag); and low and high personality types can show a preference towards consumption vs contribution (ex: high Openness users and high Neuroticism users contribute more compared to their counterparts). All in all, our results show that the challenges of newcomer churn and activity levels can be approached by making use of their personality information.

6.1 Implications

Our findings show that there is value in using a stable trait such as personality in deciding how to adapt a recommender system and customize interaction for specific personality types, which features to present to them or how to nudge them towards various existing features, who to recruit at cold-start (e.g., personality types that contribute more annotations), who to recruit for specific tasks (e.g., rating vs tagging), whether to invest particular efforts in them, or how to retain them.

7. LIMITATIONS AND FUTURE WORK

In this paper, we investigate the relationship between newcomer retention and activity, and their personality. We expand the theory on personality traits and online behavior by contributing our hypotheses and findings of user activity in one recommender system, MovieLens.

7.1 Limitations

MovieLens has the common features of a standalone recommender system with primarily anonymous features. It is not representative of all recommender systems. In particular, it is not a social system. There are limitations in the kind of data that we have and the kind of activities people can do on MovieLens.

7.2 Future Work

One future direction would be to exploit this idea in a wider variety of systems (e.g., that are not standalone, or those which are not anonymous) with different types of social affordances. High Conscientiousness users might use Amazon differently. Extroverts might use social systems differently. We leave all such investigations to future work.

There is also future work to be done in customizing the interface to match personality where it is known. Tkalcic and Chen [49] explore other ways in which personality can be used to improve performance of recommender systems such as determining whether or not to present novel, diverse items, improving performance of collaborative filtering algorithms, improving group recommendations and so forth. We focus here on issues of newcomer retention and feature usage which were not explored earlier using personality, but we wish to explore some of these in future.

We had few low Openness users in our dataset. So, in order to explore the effects of Openness trait in a useful way, we set a different lower threshold for Openness. Future work should explore whether finding few low Openness users is endemic to recommender systems or just an artifact of MovieLens. Also, in this work we analyzed personality traits in isolation from each other based on their theoretical independence. Future work, however, should explore ways in which the combination of traits found in each individual can be used to look at relationships with user retention, investment, intensity of engagement, and distribution of activity in various domains.

8. ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under the grant IIS-13-19382. We thank Tien T. Nguyen of GroupLens Research for providing us this dataset and the MovieLens users who took the personality survey. We also thank the anonymous reviewers for their valuable comments.

9. REFERENCES

[1] Adomavicius, G. and Tuzhilin, A. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17, 6 (June 2005) 734-749. DOI= 10.1109/TKDE.2005.99.

[2] Amichai-Hamburger, Y. and Ben-Artzi, E. 2000. The relationship between extraversion and neuroticism and the different uses of the Internet. *Computers in human behavior*, 16, 4, (July 2000), 441-449.

[3] Amichai-Hamburger, Y., Wainapel, G. and Fox, S. 2002. "On the Internet no one knows I'm an introvert": Extroversion, neuroticism, and Internet interaction. *Cyber Psychology & Behavior*, 5, 2 (April 2002), 125-128.

[4] Amiel, T. and Sargent, S.L. 2004. Individual differences in Internet usage motives. *Computers in Human Behavior*, 20, 6 (Oct 2004), 711-726.

[5] Anolli, L., Villani, D. and Riva, G. 2005. Personality of people using chat: An on-line research. *Cyber Psychology & Behavior*, 8, 1 (Feb 2005), 89-95.

[6] Armstrong, L., Phillips, J.G. and Saling, L.L. 2000. Potential determinants of heavier Internet usage. *International Journal of Human-Computer Studies*, 53, 4 (Oct 2000), 537-550.

[7] Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P. and Stillwell, D. 2012. Personality and patterns of Facebook usage. In *Proceedings of the 4th Annual ACM Web Science Conference* (Evanston, Illinois, June 22-24, 2012).WebSci '12. ACM, New York, NY, 24-32. DOI= http://dx.doi.org/10.1145/2380718.2380722.

[8] Barrick, M.R. and Mount, M.K. 1991. The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, 44, 1 (March 1991), 1-26.

[9] Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., and Kraut, R. E. 2004. Using social psychology to motivate contributions to online communities. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (Chicago, Illinois, Nov 06-10, 2004). CSCW '04. ACM, New York, NY, 212-221. DOI= http://dx.doi.org/10.1145/1031607.1031642.

[10] Burke, M., Marlow, C. and Lento, T. 2009. Feed me: motivating newcomer contribution in social network sites. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Boston, MA, April 04-09, 2009).CHI '09. ACM, New York, NY, 945-954. DOI= http://dx.doi.org/10.1145/1518701.1518847.

[11] Butt, S. and Phillips, J.G. 2008. Personality and self-reported mobile phone use. *Computers in Human Behavior*, 24, 2 (March 2008), 346-360.

[12] Butler, B.S. 2001. Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures. *Information Systems Research*, 12, 4 (Dec 2001), 346–362.

[13] Charlton, J.P. and Danforth. D.W. Validating the distinction between computer addiction and engagement: Online game-playing and personality *Behavior & Information Technology*, 29, 6 (Oct 2009), 601–613.

[14] Cobb-Clark, D. A. and Schurer, S. 2012. The stability of big-five personality traits. *Economics Letters* 115, 1 (April 2012), 11–15. doi:10.1016/j.econlet.2011.11.015.

[15] Correa, T., Hinsley, A.W., and Gil de Zúñiga, H. 2010. Who interacts on the Web?: The intersection of users' personality and social media use. *Comput. Hum. Behav.* 26, 2 (March 2010), 247-253. DOI= http://dx.doi.org/10.1016/j.chb.2009.09.003.

[16] Costa Jr, P.T. and McCrae, R.R. 1992. Neo personality inventory–revised (neo-pi-r) and neo five-factor inventory (neo-ffi) professional manual. *Odessa, FL: Psychological Assessment Resources*.

[17] Dror, G., Pelleg, D., Rokhlenko, O., and Szpektor, I. 2012. Churn prediction in new users of Yahoo! answers. In *Proceedings of the 21st international conference companion on World Wide Web* (Lyon, France, April 16-20, 2012). WWW '12. ACM, New York, NY, 829-834. DOI= http://doi.acm.org/10.1145/2187980.2188207.

[18] Ducheneaut. N. 2005. Socialization in an Open Source Software Community: A Socio-Technical Analysis. *Comput. Supported Coop. Work*, 14, 4 (Sep 2005), 323-368.

[19] Elahi, M., Braunhofer, M., Ricci, F., & Tkalcic, M. 2013. Personality-based active learning for collaborative filtering recommender systems. In *AI* IA 2013: Advances in Artificial Intelligence,* Springer International Publishing. 360-371.

[20] Freyne, J., Jacovi, M., Guy, I., and Geyer, W. 2009. Increasing engagement through early recommender intervention. In *Proc RecSys '09*. ACM, New York, NY, 85-92. DOI= http://dx.doi.org/10.1145/1639714.1639730

[21] Fuglestad, P. T., Dwyer, P. C., Moses, J. F., Kim, J. S., Mannino, C. A., Terveen, L., and Snyder, M. 2012. What Makes Users Rate (Share, Tag, Edit…)? Predicting Patterns of Participation in Online Communities. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (Seattle, Washington, Feb 11-15, 2012). CSCW '12. ACM, New York, NY, 969-978. DOI= http://dx.doi.org/10.1145/2145204.2145349.

[22] Gosling, S. D., Rentfrow, P.J., and Swann, W.B. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality* 37, 6 (Dec 2003), 504–528.

[23] Hogan, J. and Ones, D.S. 1997. Conscientiousness and Integrity at Work. *Handbook of Personality Psychology* (1997), 849–870.

[24] Hu, R. and Pu, P. 2013. Exploring Relations between Personality and User Rating Behaviors. In *UMAP Workshops*.

[25] Hughes, D.J., Rowe, M., Batey, M. and Lee, A. 2012. A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. *Computers in Human Behavior*, 28, 2 (March 2012), 561-569.

[26] John, O.P. The Big Five factor taxonomy: Dimensions of personality in the natural language and in questionnaires. *Handbook of personality: Theory and research*, 14 (1990) 66–100.

[27] Judge, T.A. and Ilies, R. 2002. Relationship of personality to performance motivation: a meta-analytic review. *Journal of applied psychology*, 87, 4 (Aug 2002), 797-807.

[28] Kairam S.R, Wang, D.J., and Leskovec, J. 2012. The life and death of online groups: predicting group growth and longevity. In *Proceedings of the fifth ACM International Conference on Web Search and Data Mining* (Seattle, Washington, Feb 8-12, 2012).WSDM '12. ACM, New York, NY, 673-682. DOI= http://doi.acm.org/10.1145/2124295.2124374.

[29] Karumur R.P., Nguyen T.T, and Konstan J.A. 2016. Early Activity Diversity: Assessing Newcomer Retention from First-Session Activity. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, Feb 27 – Mar 02, 2016).CSCW '16. ACM, New York, NY, 595-608. DOI= http://dx.doi.org/10.1145/2818048.2820009.

[30] Kawale, J., Pal, A. and Srivastava, J. 2009. Churn prediction in MMORPGs: A social influence based approach. In *12th International Conference on Computational Science and Engineering* (Vancouver, BC, Aug 29-31, 2009) CSE'09. IEEE, NJ, 423-428. DOI= http://dx.doi.org/10.1109/CSE.2009.80.

[31] Kelly, E.L. and Conley, J.J. 1987. Personality and compatibility: a prospective analysis of marital stability and marital satisfaction. *Journal of personality and social psychology*, 52, 1 (Jan 1987), 27-40.

[32] Kraaykamp, G. and Van Eijck, K., 2005. Personality, media preferences, and cultural participation. *Personality and individual differences*, 38, 7 (May 2005), 1675-1688.

[33] Landers, R.N. and Lounsbury, J.W. 2006. An investigation of Big Five and narrow personality traits in relation to Internet usage. *Computers in Human Behavior,* 22, 2 (March 2006), 283–293.

[34] Mark, G. and Ganzach, Y. 2014. Personality and Internet usage: A large-scale representative study of young adults. *Computers in Human Behavior,* 36 (July 2014), 274–281.

[35] McCrea, R. and John, O. An introduction to the five-factor model and its applications. *Journal of personality*, 60, 2 (June 1992), 175–215.

[36] McCrea, R.R. and Costa Jr, P.T. A five-factor theory of personality. L.A. Pervin, O.P. John (Eds.), *Handbook of personality: Theory and research*, Guilford, NY (1999), 139–153.

[37] Nov, O. 2007. What motivates Wikipedians? *Communications of the ACM*, 50, 11 (Nov 2007), 60-64. DOI= http://doi.acm.org/10.1145/1297797.1297798.

[38] Orchard, L.J. and Fullwood, C. 2010. Current perspectives on personality and Internet use. *Social Science Computer Review*, 28, 2 (May 2010), 155–169.

[39] Panciera, K., Halfaker, A. and Terveen, L., 2009, May. Wikipedians are born, not made: a study of power editors on Wikipedia. In *Proceedings of the ACM 2009 international conference on supporting group work* (Sanibel Island, Florida, May 10-13, 2009), GROUP '09, ACM, New York, NY, 51-60. DOI= http://dx.doi.org/10.1145/1531674.1531682.

[40] Phillips, J., Butt, S., and Blaszczynski, A. 2007. Personality and Self-Reported Use of Mobile Phones for Games. *Cyber Psychology & Behavior*, 9, 6 (Jan 2007), 753–758.

[41] Rentfrow, P.J., and Gosling, S.D. 2003. The do re mi's of everyday life: the structure and personality correlates of music preferences. *J Pers Soc Psychol*. 84, 6, (Jun 2003), 1236-1256.

[42] Ryan, T. and Xenos, S. 2011. Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. *Computers in Human Behavior,* 27, 5 (Sep 2011), 1658–1664.

[43] Ross, C., Orr, E. S., Sisic, M., Arseneault, J. M., Simmering, M. G., and Orr, R. R. 2009. Personality and motivations associated with Facebook use. *Computers in Human Behavior,* 25, 2 (March 2009), 578-586.

[44] Schein A.I., Popescul, A., Ungar, L.H., and Pennock, D. M. 2002. Methods and metrics for cold-start recommendations. *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval* (Tampere, Finland, Aug 11-15, 2002), SIGIR '02, ACM, New York, NY, 253-260. DOI= http://dx.doi.org/10.1145/564376.564421.

[45] Schmitt D., J. Allik, R. McCrae, and V. Benet-Martinez. 2007. The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, 38, 2, (March 2007), 173-212.

[46] Schrammel, J., Köffel, C. and Tscheligi, M. 2009. Personality traits, usage patterns and information disclosure in online communities. In *Proceedings of the 23rd British HCI Group annual conference on people and computers: Celebrating people and technology* (Cambridge, UK, Sep 01-05, 2009). BCS-HCI '09. British Computer Society, Swinton, UK, 169-174.

[47] Swickert, R.J., Hittner, J.B., Harris, J.L. and Herring, J.A. 2002. Relationships among Internet use, personality, and social support. *Computers in human behavior*, 18, 4, (July 2002), 437–451.

[48] Teng, C. 2008. Personality Differences between Online Game Players and Nonplayers in a Student Sample. *CyberPsychology & Behavior,* 11, 2, (April 2008), 232–234.

[49] Tkalcic, M. and Chen L. 2015. Personality and Recommender Systems. In *Recommender Systems Handbook*, Springer US. 715–739.

[50] Tosun, L. P., and Lajunen, T. 2010. Does Internet use reflect your personality? Relationship between Eysenck's personality dimensions and Internet use. *Computers in Human Behavior*, 26, 2, (March 2010), 162–167.

[51] Tupes, E.C., and Christal, R.E. 1992. Recurrent personality factors based on trait ratings. *Journal of Personality*, 60, 2, (April 2006), 225–251.

[52] Turner T.C., Smith, M. A., Fisher, D., and Welser, H. T. 2005. Picturing Usenet: Mapping computer mediated collective action. In *Computer-Mediated Communication*. 10, 4, (June 2006), 00-00.

[53] Tuten T.L., and Bosnjak M. 2001. Understanding Differences In Web Usage: The Role Of Need For Cognition And The Five Factor Model Of Personality. *Social Behavior and Personality: an international journal soc behav pers,* 29, 4 (January 2001), 391–398.

[54] Wasko M.M, and Faraj, S. 2000. "It is what one does": why people participate and help others in electronic communities

of practice. *Journal of Strategic Information Systems* 9, 2-3 (Sep 2000), 155-173.

[55] Welser, H.T., Gleave, E., Fisher, D., and Smith, M. 2007. Visualizing the signatures of social roles in online discussion groups. *Journal of social structure*, 8, 2, (July 2007), 1-32.

[56] Wolfradt, U., and Doll, J. 2001. Motives of adolescents to use the Internet as a function of personality traits, personal and social factors. *Journal of Educational Computing Research*, 24, 1, (Jan 2001), 13-27.

[57] Yang, J., Wei, X., Ackerman, M.S. and Adamic, L.A. 2010. Activity Lifespan: An Analysis of User Survival Patterns in Online Knowledge Sharing Communities In Proceedings of the Fourth International AAAI Conference on Web and Social Media (Washington, DC, May 23-26, 2010). ICWSM '10. AAAI, Palo Alto, CA, 186-193.

[58] Zuckerman, M., Ulrich, R.S., and McLaughlin, J. 1993. Sensation seeking and reactions to nature paintings. *Personality and Individual Differences,* 15, 5, (Nov 1993), 563-576.