Development of a Reading Material Recommender System Based On a Design Science Research Approach

Evren Eryilmaz

March 15th, 2018

Overview

- Motivation and problem identification
- Objective of my software
- Design and development
- Demonstration
- Evaluation
- Communication
- Comments & questions

Motivation

- Collaboration between software developers and business users can be instrumental to the success of software development projects.
- Effective collaboration is an important interpersonal skills for an entry-level software developer' professional growth within an organization (Aasheim et al., 2009).
- Asynchronous online discussion can facilitate a natural setting for collaboration in virtual teams.

Aasheim, C. L., Li, L., & Williams, S. (2009). Knowledge and skill requirements for entry-level information technology workers: A comparison of industry and academia. *Journal of Information Systems Education*, *20*(3), 349-356.

Problem Identification

Many students perceive online discussions more confusing compared to face-to-face discussions because

- They feel being overwhelmed by a large number of messages

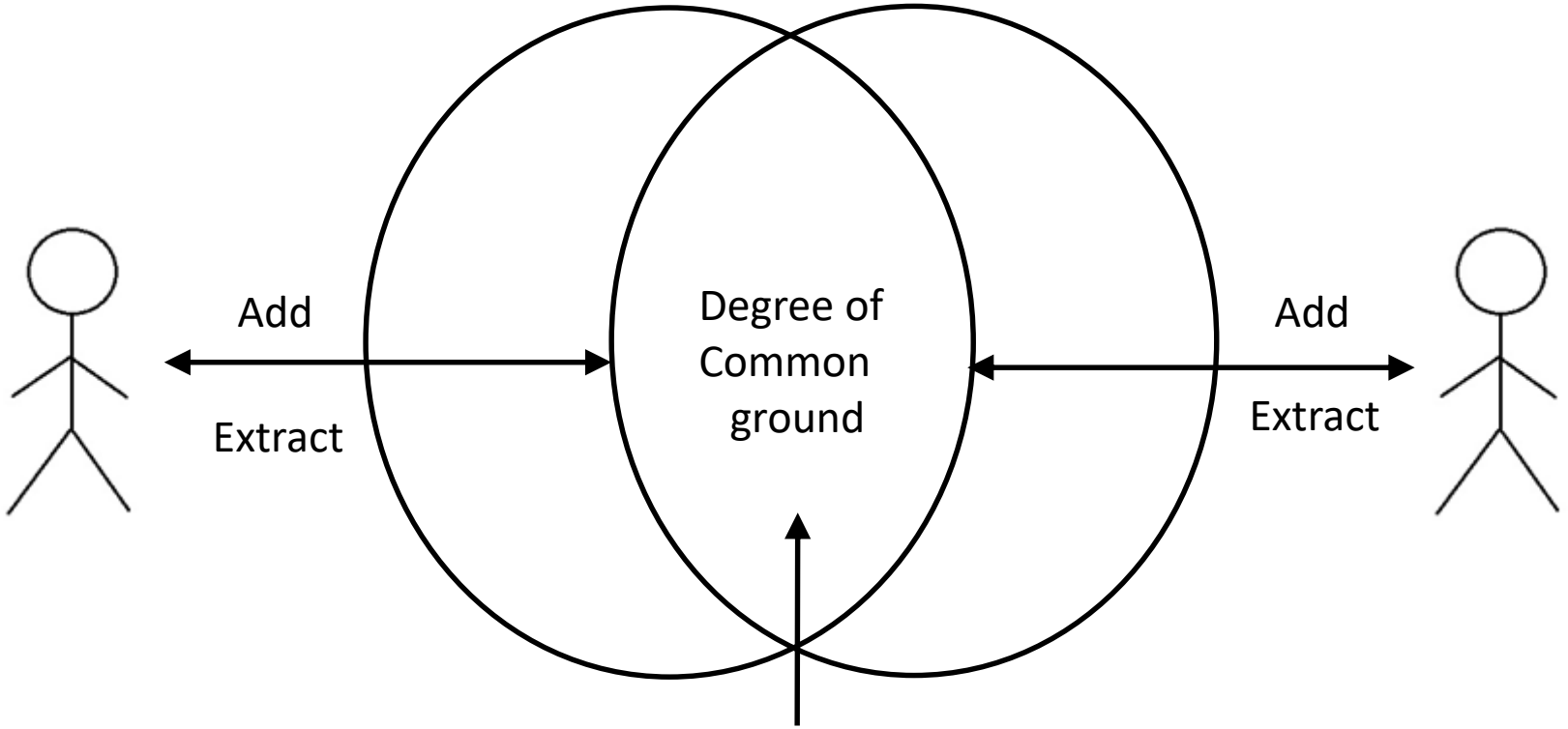
Peters, V. L., & Hewitt, J. (2010). An investigation of student practices in asynchronous computer conferencing courses. *Computers & Education*, *54*(4), 951-961.

Objective of my software

- Draw asynchronous online discussion participants' attention to the most important parts of overwhelmingly large discussions.

Qiu, M., & McDougall, D. (2015). Influence of group configuration on online discourse reading. *Computers & Education*, *87*, 151-165.

Design



A recommendation functionality with high predictive accuracy and perceived usefulness

Design

- Collaborative Filtering
- Content-based Filtering
- Knowledge-based filtering
- Hybrid approaches

Abel, F., Bittencourt, I. I., Costa, E., Henze, N., Krause, D., & Vassileva, J. (2010). Recommendations in online discussion forums for e-learning systems. *IEEE transactions on learning technologies*, *3*(2), 165-176.

Design

1. Students' interests change over time depending on their level of understanding of a subject
2. The system needs to generate precise recommendations with a small amount of input

Development

Cosine Similarity

$$Sim(a,b) = \frac{\vec{r}_a \vec{r}_b}{||\vec{r}_a|| \, ||\vec{r}_b||} = \frac{\sum_p r_{a,p} r_{b,p}}{\sqrt{\sum_p r^2_{a,p}} \sqrt{\sum_p r^2_{b,p}}}$$

a,b : students

$r_{a,p}$: rating of student a for message p

Development

Pearson Correlation Coefficient

$$sim(a,b) = \frac{\sum_{p \in I} (r_{a,p} - \bar{r}_a)(r_{b,p} - \bar{r}_b)}{\sqrt{\sum_{p \in I} (r_{a,p} - \bar{r}_a)^2} \sqrt{\sum_{p \in I} (r_{b,p} - \bar{r}_b)^2}}$$

a,b : students

$r_{a,p}$: rating of student a for message p

I : set of messages, rated both by a and b

Development

Constrained Pearson Correlation Coefficient

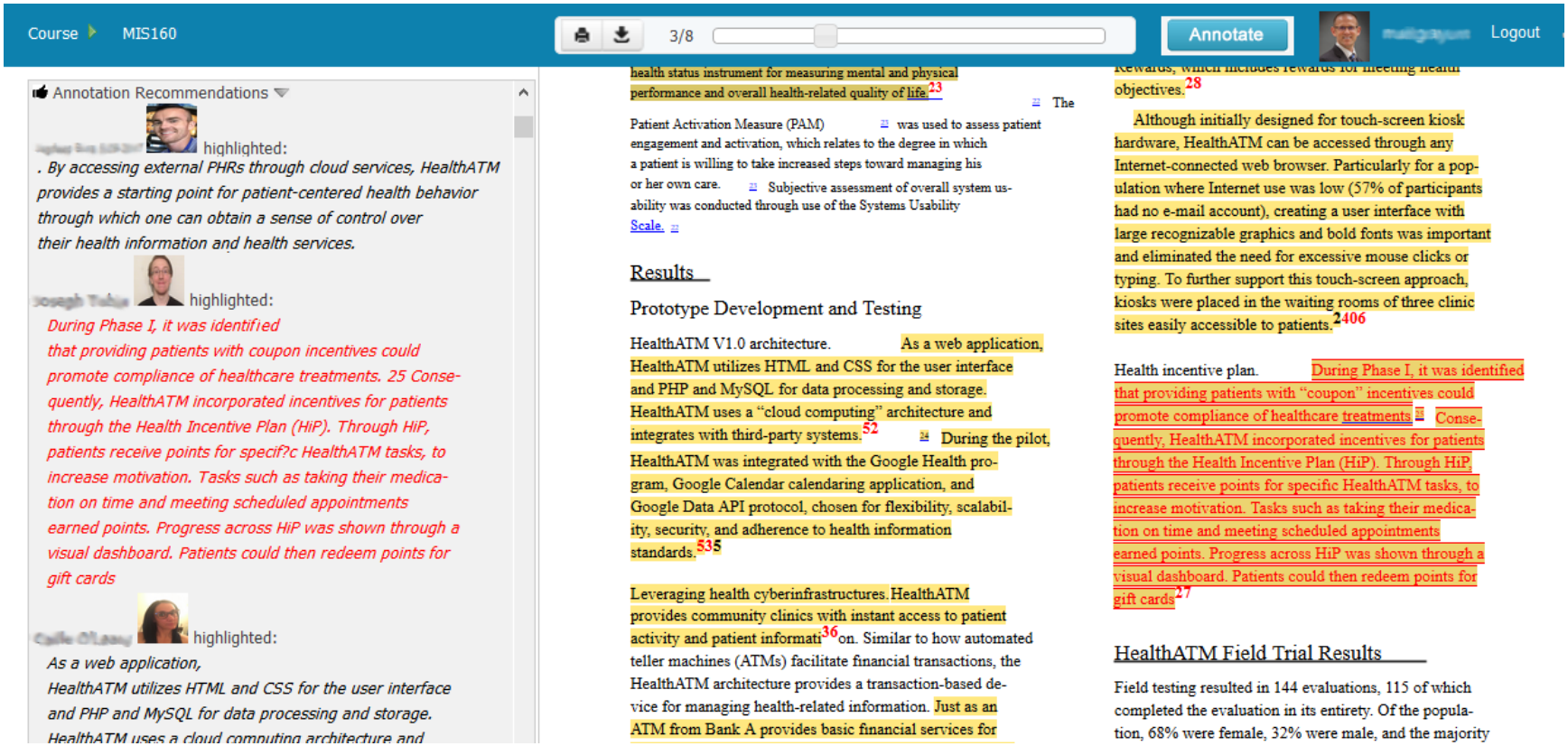
$$sim(a,b) = \frac{\sum_{p \in I} (r_{a,p} - r_{med})(r_{b,p} - r_{med})}{\sqrt{\sum_{p \in I} (r_{a,p} - r_{med})^2} \quad \sqrt{\sum_{p \in I} (r_{b,p} - r_{med})^2}}$$

Development

$$Pred(a,p) = \bar{r}_a + \frac{\sum_{b \in N} sim(a,b)(r_{b,p} - \bar{r}_b)}{\sum_{b \in N} |sim(a,b)|}$$

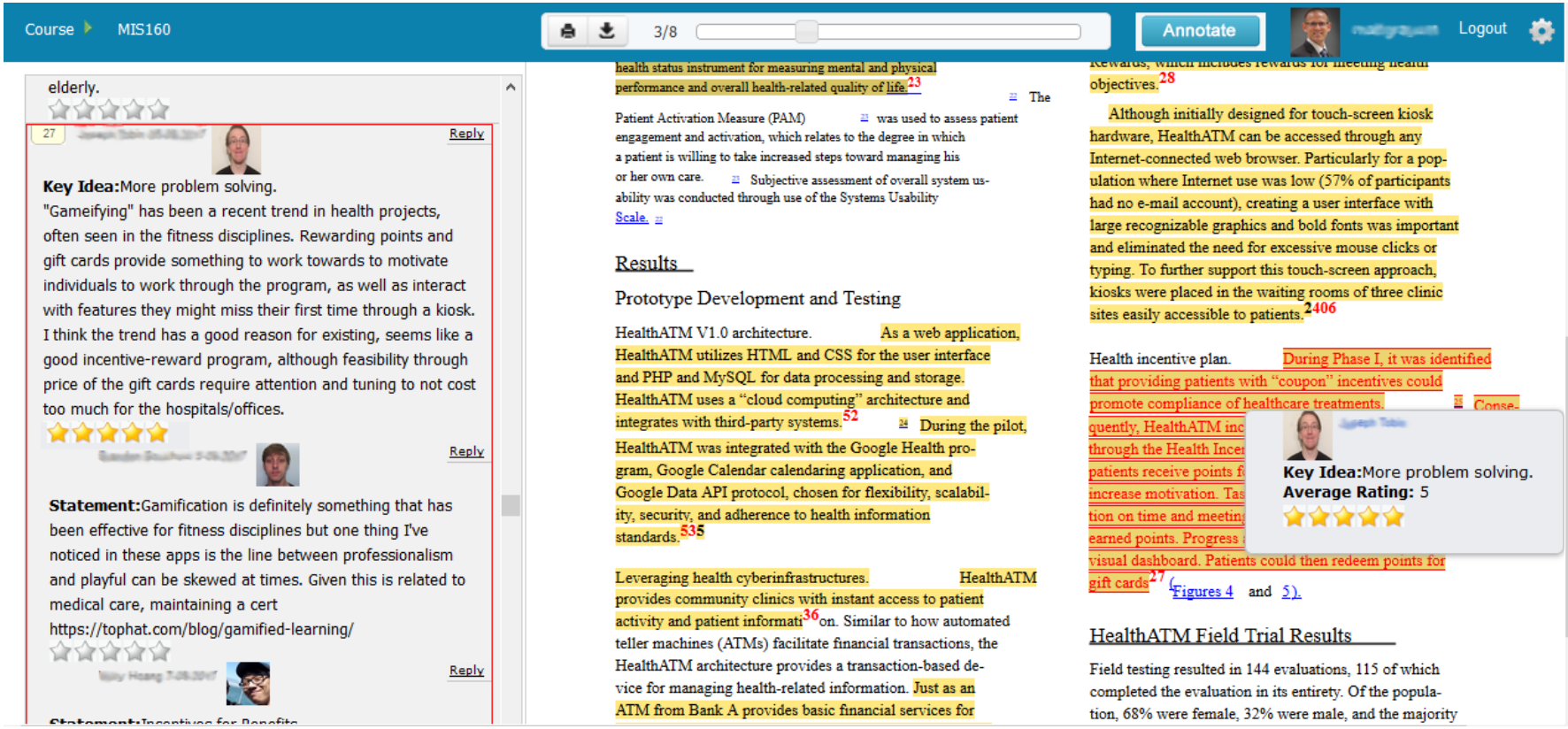
Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In *The adaptive web* (pp. 291-324). Springer, Berlin, Heidelberg.

Development



Eryilmaz, E. & Thoms, B., & Canelon, J. (Accepted). How Design Science Research Helps Improving Learning Efficiency in Online Conversations. Communications of the Association of Information Systems.

Development



Demonstration

- Experiment 1: Is there any difference in the predictive accuracy and perceived usefulness of the developed recommendation functionalities?

Evaluation-Predictive Accuracy

$$\sqrt{1/n \sum_{u,i} (p_{u,i} - r_{u,i})^2}$$

Recommendation Functionality	Root Mean Squared Error
Cosine Similarity	1.73
Pearson Correlation Coefficient	1.21
Constrained Pearson Correlation Coefficient	0.87

Evaluation-Perceived Usefulness

Q1: The recommendations were exactly what I was looking for

Recommendation Functionality	Average	Standard Deviation
Cosine Similarity	3.62	0.78
Pearson Correlation Coefficient	4.06	0.60
Constrained Pearson Correlation Coefficient	4.44	0.61

$F(2,99) = 12.90,\ p < 0.001^{***}$

Evaluation-Perceived Usefulness

Q1: The recommendations were exactly what I was looking for

Comparison pair	Tukey HSD Q statistic	Tukey HSD p-value
Cosine Similarity vs Pearson Correlation Coefficient	3.85	0.02*
Cosine Similarity vs Constrained Pearson Correlation Coefficient	3.33	0.05*
Pearson Correlation Coefficient vs Constrained Pearson Correlation Coefficient	7.18	0.001**

Evaluation-Perceived Usefulness

Q2: I was surprised by the recommendations

Recommendation Functionality	Average	Standard Deviation
Cosine Similarity	4.09	0.65
Pearson Correlation Coefficient	4.23	0.67
Constrained Pearson Correlation Coefficient	4.35	0.64

$F(2,99) = 1.39,\ p = 0.25$

Evaluation-Perceived Usefulness

Q3: The recommendations helped me to read instructional materials more effectively

Recommendation Functionality	Average	Standard Deviation
Cosine Similarity	4.15	0.68
Pearson Correlation Coefficient	4.29	0.70
Constrained Pearson Correlation Coefficient	4.38	0.55

$F(2,99) = 1.15, p = 0.32$

Evaluation-Perceived Usefulness

Q4: The recommendations prompted me to read postings on the forum

Recommendation Functionality	Average	Standard Deviation
Cosine Similarity	4.15	0.71
Pearson Correlation Coefficient	4.29	0.82
Constrained Pearson Correlation Coefficient	4.38	0.61

$F(2,99) = 11.82,\ p < 0.001^{***}$

Evaluation-Perceived Usefulness

Q4: The recommendations prompted me to read postings on the forum

Comparison pair	Tukey HSD Q statistic	Tukey HSD p-value
Cosine Similarity vs Pearson Correlation Coefficient	3.56	0.04*
Cosine Similarity vs Constrained Pearson Correlation Coefficient	6.88	0.001**
Pearson Correlation Coefficient vs Constrained Pearson Correlation Coefficient	3.32	0.05*

Evaluation-Perceived Usefulness

Q5: The recommendations prompted me to write on the forum

Recommendation Functionality	Average	Standard Deviation
Cosine Similarity	3.89	0.76
Pearson Correlation Coefficient	4.09	0.51
Constrained Pearson Correlation Coefficient	4.25	0.45

$F(2,99) = 3.53, p = 0.03^*$

Evaluation-Perceived Usefulness

Q5: The recommendations prompted me to write on the forum

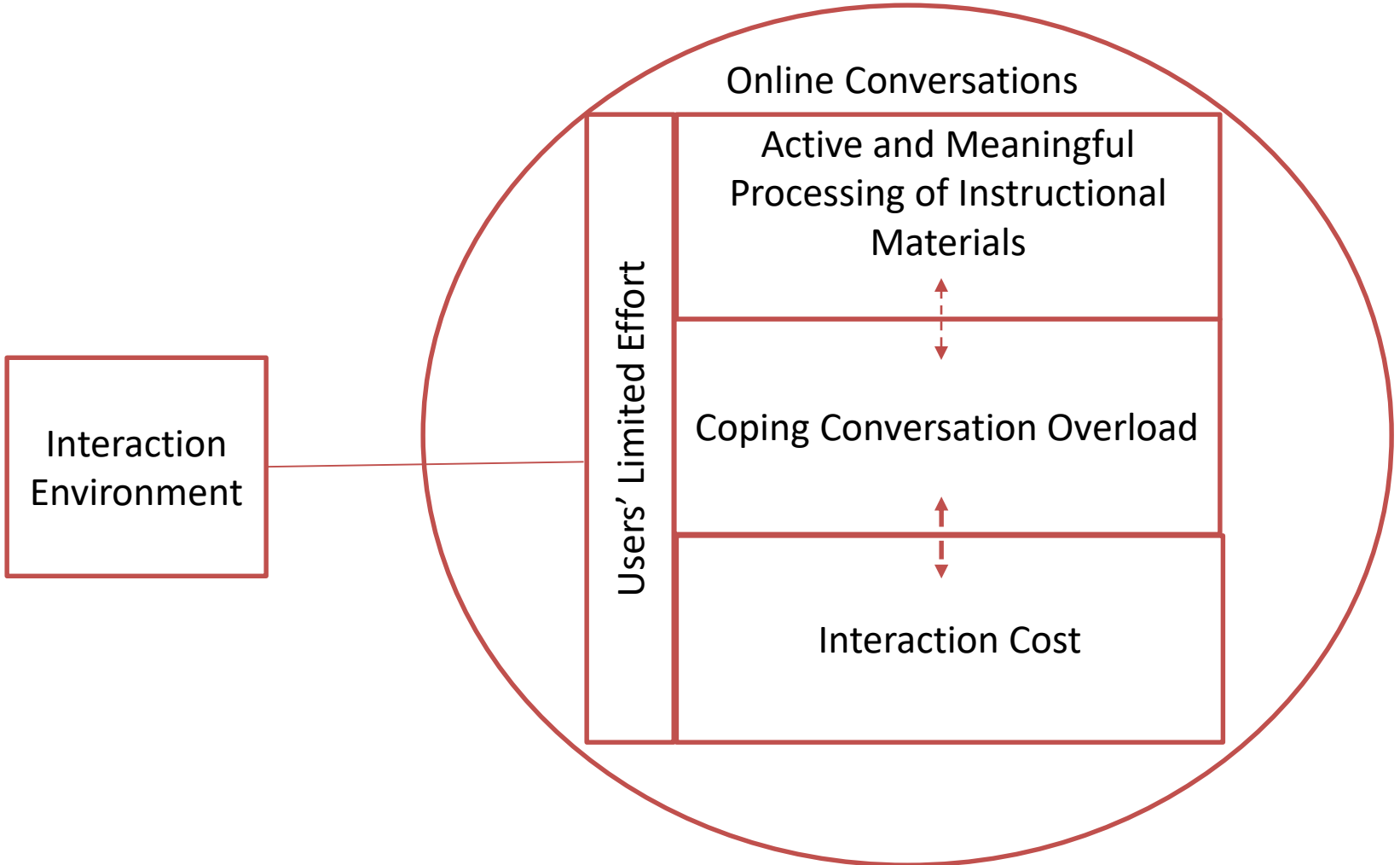
Comparison pair	Tukey HSD Q statistic	Tukey HSD p-value
Cosine Similarity vs Pearson Correlation Coefficient	2.02	0.33
Cosine Similarity vs Constrained Pearson Correlation Coefficient	3.76	0.02*
Pearson Correlation Coefficient vs Constrained Pearson Correlation Coefficient	1.73	0.44

Evaluation-Conversation Overload Coping Strategies

Q1: In an average week, what percentage of the week's messages do you read?

	Control Software	Constrained Pearson Correlation Coefficient		
Choices	%	%	X^2	P
0-20%	0.15	0.09	0.56	0.45
21-40%	0.35	0.12	5.23	0.02*
41-60%	0.32	0.12	4.19	0.04*
61-80%	0.15	0.5	9.68	0.002**
81-100%	0.03	0.18	3.99	0.05*

Communication



Thank You for Your Time

Your Comments and Questions are welcomed.

Have a great spring break!