

Table 3. User feedback. Average scores and p-values between the baseline (B) and our system (O), for each question and group of questions (bold font). Responses on a Likert scale: strongly disagree = -2, disagree = -1, neither agree or disagree = 0, agree = 1, strongly agree = 2. We reversed scores for questions where disagree was better (marked with *), so higher numbers are always better. Questions marked in gray showed no statistically significant difference between the two systems.

Question	B	O	p-value
I. Quality of Recommended Items	0.2	0.9	$2 \cdot 10^{-6}$
1. The keywords displayed to me matched the search objective	0.5	1.2	$6 \cdot 10^{-4}$
2. The articles displayed to me matched the search objective	0.2	1.2	$2 \cdot 10^{-5}$
3. The system produced good results	-0.2	1.1	$3 \cdot 10^{-6}$
4. The system helps me discover new articles	0.7	1.3	0.03
5. The articles displayed to me are similar to each other*	0.0	-0.2	0.33
II. Interaction Adequacy	-0.4	0.8	$2 \cdot 10^{-5}$
6. The search streams provides an adequate way for me to conduct my search	-0.7	1.0	$3 \cdot 10^{-5}$
7. The keyword feedback provides an adequate way for me to refine my search	0.0	0.6	0.04
III. Interface Adequacy	0.9	1.1	0.14
8. The design of the system interface components (how to conduct a search, how to give feedback, and scrolling through the articles) is clear and adequate	0.9	1.1	0.33
9. The layout of the system interface (the positioning of the keywords, radar and article list) is attractive and adequate	0.9	1.0	0.16
IV. Perceived Ease of Use	-0.4	0.8	$8 \cdot 10^{-6}$
10. I easily found the articles that were relevant to my search	-0.5	0.8	$3 \cdot 10^{-6}$
11. It is easy to learn to tell the system what to search for	-0.4	1.1	$3 \cdot 10^{-5}$
12. I feel in control of telling the system what I want	-0.4	0.8	$1 \cdot 10^{-3}$
13. I understood why the articles were recommended to me	-0.1	0.7	$4 \cdot 10^{-3}$
V. Attitude	-0.3	1.0	$1 \cdot 10^{-5}$
14. Overall, I am satisfied with the system	-0.3	1.0	$1 \cdot 10^{-5}$
VI. Behavioral Intentions	0.3	0.7	$5 \cdot 10^{-4}$
15. If a system such as this exists, I would use it to find scientific articles	-0.2	1.1	$3 \cdot 10^{-6}$
16. I think I would use this system frequently if given the opportunity	-0.4	0.6	$2 \cdot 10^{-4}$
17. I found the system unnecessarily complex*	0.5	0.7	0.08
18. I thought the system was easy to use	0.4	0.8	0.05
19. I think that I would need the support of a technical person to be able to use the system*	0.8	0.6	0.43
20. I found the various functions in this system were well integrated	0.6	0.8	0.16
21. I thought there was too much inconsistency in the system*	-0.1	0.5	0.01
22. I would imagine that most people would learn to use this system very quickly	0.3	0.5	0.16
23. I found the system very cumbersome to use*	0.4	0.6	0.38
24. I felt very confident using the system	-0.1	0.4	$7 \cdot 10^{-3}$
25. I needed to learn a lot of things before I could get going with this system*	0.5	0.6	0.58