

Appendix A: Exploratory Factor Analysis

Anita Crescenzi, Austin R. Ward, Yuan Li, Rob Capra. 2021. Supporting metacognition during exploratory search with the OrgBox. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21), July 11–15, 2021, Virtual Event, Canada. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3404835.3462955>

This appendix contains

- the full text of the 13 metacognition items on the post-task questionnaire (also shown in Table 4 in the paper),
- overall descriptive statistics (M , SD , min , max) and correlations between items (Pearson's r)
- results of exploratory factor analysis
- a description of the composite variable creations and descriptive statistics

Metacognition items on Post-task Questionnaire

Participants indicated their agreement with the following statements on a 1-7 point scale with labeled endpoints (1=Strongly disagree, 7=Strongly agree).

The OrgDoc/OrgBox helped me to...

- ...keep track of how broadly I was covering the topic. (*track broad*)
- ...keep track of how deeply I was covering some aspects of the topic. (*track deep*)
- ...organize the information about the topic. (*org info*)

- ...keep track of my progress towards completing the task. (*track progress*)
- ...decide when I had met the task goal. (*met goal*)
- ...evaluate whether my strategy for the task was working. (*evaluate strategy*)
- ...change my approach to the task. (*change approach*)
- ...ask myself if I had included important facets of the topic. (*important facets*)

- ...figure out how the information about the topic fits together. (*info fits together*)
- ...determine facets of the topic. (*determine facets*)

- ...understand what the task was asking me to do. (*understand task*)
- ...develop a plan for approaching the task. (*plan task*)
- ...decide where to begin my search. (*begin search*)

Descriptive statistics

Appendix Table 1. Post-task questionnaire metacognition items: descriptive statistics (Mean, SD, Min, Max), pairwise correlations (Pearson's r with uncorrected statistical significance. ‡ $p < .001$, † $p < .01$, * $p < .05$), and composite variable (after exploratory factor analysis)

Item		Descriptive Statistics				Correlations													Composite
#	Name	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	Name
(1)	<i>track broad</i>	5.23	1.26	2	7	1													monitor
(2)	<i>track deep</i>	4.96	1.35	2	7	† .44	1												monitor
(3)	<i>org info</i>	5.58	1.20	2	7	‡ .53	‡ .46	1											monitor
(4)	<i>track progress</i>	5.08	1.61	1	7	* .30	† .37	.24	1										evaluate
(5)	<i>met goal</i>	4.42	1.47	1	7	.20	† .46	.11	‡ .61	1									evaluate
(6)	<i>eval. strategy</i>	4.73	1.30	1	7	.12	* .34	* .34	† .38	‡ .47	1								evaluate
(7)	<i>change approach</i>	4.31	1.49	1	7	.19	.14	.18	.27	* .34	‡ .63	1							evaluate
(8)	<i>important facets</i>	5.15	1.37	2	7	† .38	* .34	* .32	‡ .52	* .34	‡ .51	† .44	1						evaluate
(9)	<i>info fits together</i>	4.85	1.53	2	7	† .44	* .33	† .41	† .42	.27	.25	* .32	† .38	1					understand topic
(10)	<i>determine facets</i>	4.79	1.40	2	7	† .38	.25	† .45	† .38	.21	.27	.19	‡ .54	‡ .75	1				understand topic
(11)	<i>understand task</i>	3.83	1.55	1	7	.18	.04	.10	* .36	.25	.17	.24	.22	‡ .53	‡ .50	1			plan
(12)	<i>plan task</i>	4.65	1.54	1	7	.15	.26	.05	0.27	.25	† .42	* .36	† .44	† .44	† .37	‡ .48	1		plan
(13)	<i>begin search</i>	3.31	1.52	1	6	.08	.03	.05	† .37	* .33	.25	.18	* .31	.26	.28	‡ .54	* .37	1	plan

Exploratory factor analysis

Before exploratory factor analysis, two tests were run to ascertain whether or not it was advisable to pursue factor analysis: Barlett's test of sphericity and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy.

Barlett's test of sphericity. $\chi^2(78)=265.57, p<.001$

H_0 items are not correlated.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy. $KMO=.729$

Then exploratory factor analysis was conducted using principal factor with a pre-specified three factor solution.

Oblique, promax rotation to allow for correlated factors.

Pattern matrix of promax-rotated loadings

Structural matrix of unrotated loadings after principal factors extraction.

Requirements for item retention: (1) loading $\geq .4$, (2) no crossloadings $> .32$, (3) communality $\geq .4$

Factor loadings $> .4$ are shaded.

Appendix Table 2: Pattern matrices after principal factors extraction

Item Name	Rotated loadings			Communality	Unrotated loadings		
	F1	F2	F3		F1	F2	F3
<i>track broad</i>	0.00	0.00	0.68	0.51	0.50	0.44	-0.10
<i>track deep</i>	0.34	-0.23	0.57	0.59	0.51	0.39	0.24
<i>organize info</i>	0.02	-0.13	0.77	0.61	0.49	0.55	-0.03
<i>track progress</i>	0.46	0.24	0.16	0.61	0.66	-0.07	0.13
<i>met goal</i>	0.65	0.06	0.03	0.62	0.57	-0.13	0.35
<i>evaluate strategy</i>	0.78	-0.02	0.04	0.71	0.62	-0.12	0.47
<i>change approach</i>	0.61	0.10	-0.04	0.59	0.53	-0.18	0.32
<i>important facets</i>	0.46	0.18	0.27	0.68	0.71	0.02	0.15
<i>info fits together</i>	-0.06	0.57	0.46	0.76	0.73	0.10	-0.39
<i>determine facets</i>	-0.12	0.58	0.49	0.78	0.71	0.13	-0.44
<i>understand task</i>	0.00	0.82	-0.11	0.65	0.55	-0.38	-0.41
<i>plan task</i>	0.33	0.51	-0.10	0.57	0.58	-0.32	-0.06
<i>begin search</i>	0.24	0.57	-0.23	0.46	0.45	-0.41	-0.12
Variance/Eigenvalues	3.18	3.14	3.03		4.55	1.16	1.08
Variance explained	0.44	0.43	0.42		0.62	0.16	0.15

Appendix Table 3: Correlation matrix of promax rotated factors

	F1	F2	F3
F1	1		
F2	0.36	1	
F3	0.37	0.40	1

Composite variables

Composite variables were created by averaging the values of the items in the factors.

$$\text{Monitor} = (\text{track broad} + \text{track deep} + \text{organize info}) / 3$$

$$\text{Evaluate} = (\text{track progress} + \text{met goal} + \text{evaluate strategy} + \text{change approach} + \text{important facets}) / 5$$

$$\text{Plan} = (\text{understand task} + \text{plan task} + \text{begin search}) / 3$$

$$\text{Understand topic} = (\text{info fits together} + \text{determine facets}) / 2$$

Appendix Table 4: Descriptive statistics (Mean, SD, min, max, skew, kurtosis), and correlation matrix for composite variables. Note: kurtosis = 3 for normal distribution

Item		Descriptive statistics						Correlations			
#	Name	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	Skew	Kurtosis	(1)	(2)	(3)	(4)
(1)	Monitor	5.26	1.03	3	7	-.19	2.14	1.00			
(2)	Evaluate	4.74	1.08	1.4	6.8	-.55	3.77	† .45	1.00		
(3)	Understand topic	4.82	1.37	2	7	-.19	2.43	‡ .50	‡ .46	1.00	
(4)	Plan	3.93	1.23	1	6.7	-.28	3.05	.16	‡ .49	‡ .53	1