A Survey on Temporal Web Search Experience

Hideo Joho[1], Adam Jatowt[2], and Roi Blanco[3]
[1]University of Tsukuba, Japan
[2]Kyoto University, Japan
[3]Yahoo! Research, Barcelona, Spain

Outline

- Motivation
- Data Collection Method
- Gross Statistics
- Correlation → Breakdown
- (Some) Qualitative Results
- Discussion
- Summary

Motivation

- Much of research on temporal search focuses on log analyses or system development
- A broader and deeper understanding of people's temporal information seeking behaviour can facilitate the community
  - Morris [11] for collaborative search
 - Ed and Chi [5] for social search
- Let's ask people!

Data Collection Method 1

- Critical Incident Analysis
  - Asking people to describe a prominent experience
  - Usually carried out via interviews
  - Evans & Chi [5] applied to questionnaires
- 18 Questions
  - Most recent web search experience
  - 15 closed and 3 open questions (Broad scope)
  - Contexts, Information Needs, Process, Outcomes
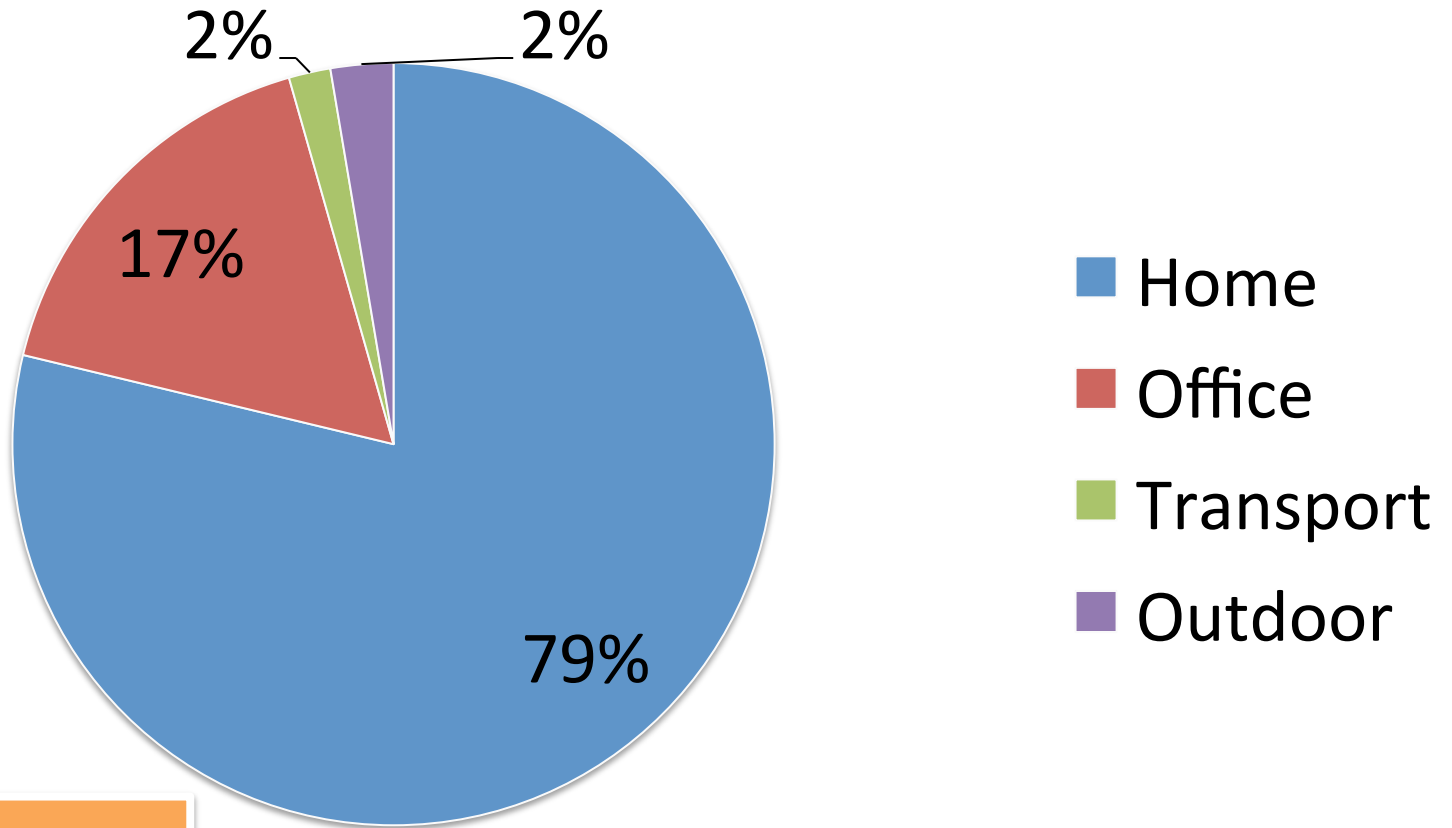
Data Collection Method 2

- 110 Participants in Japan
  - 11 Females + 11 Males / 20s – 60s Blocks
  - Filtered by a search experience within 3 days
  - 48% single, 52% married
  - Self-reported occupations
    - Office Workers (36), Homemakers (21) Part-time workers (15), University Students (8), Others (16), Managers (2)
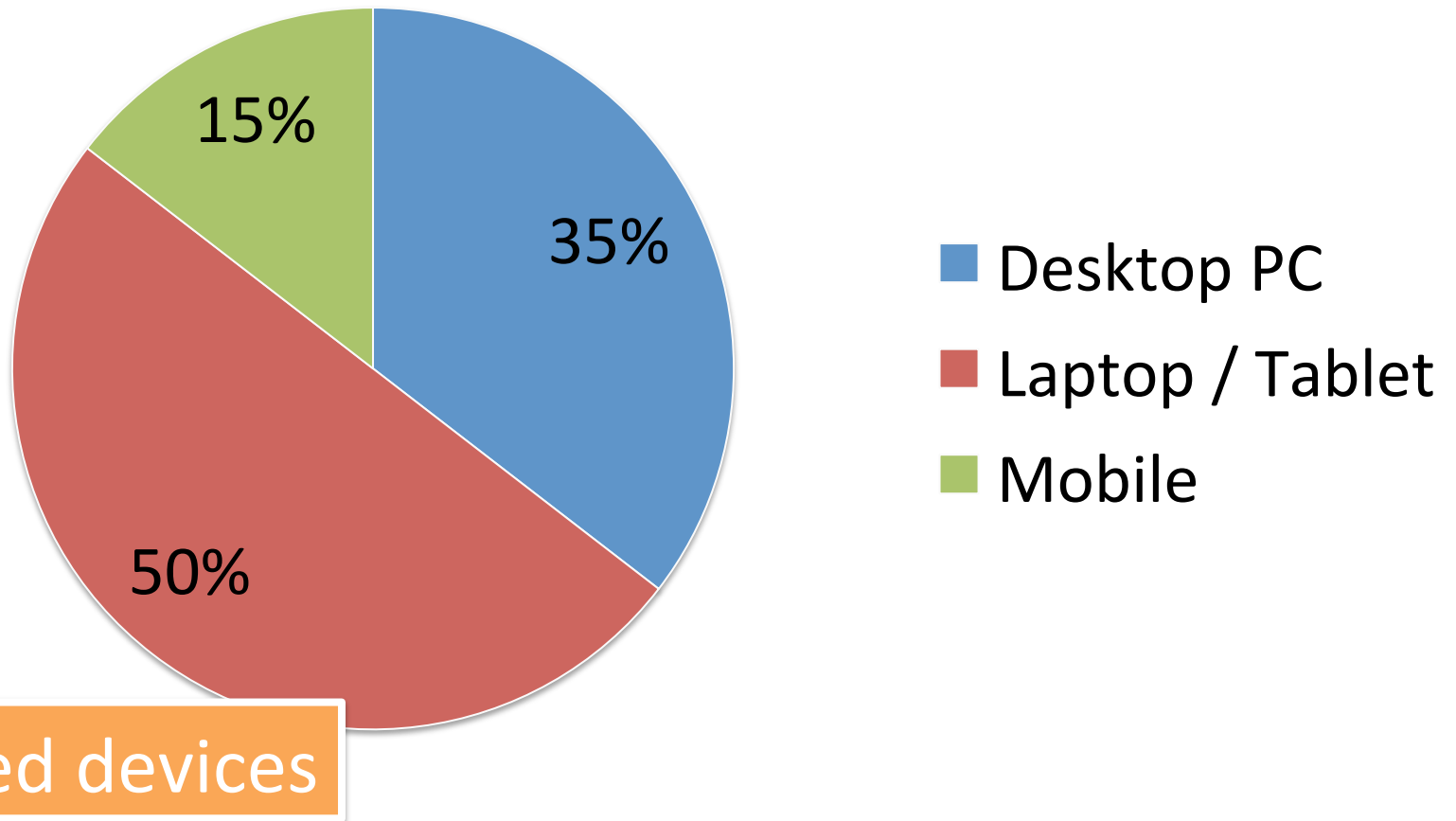- Mostly quantitative analyses so far

Context: When

Legend:
- < 1 hour
- < 1 day
- Yesterday
- < 2 days
- > 2 days

43%, 38%, 15%, 3%, 1%
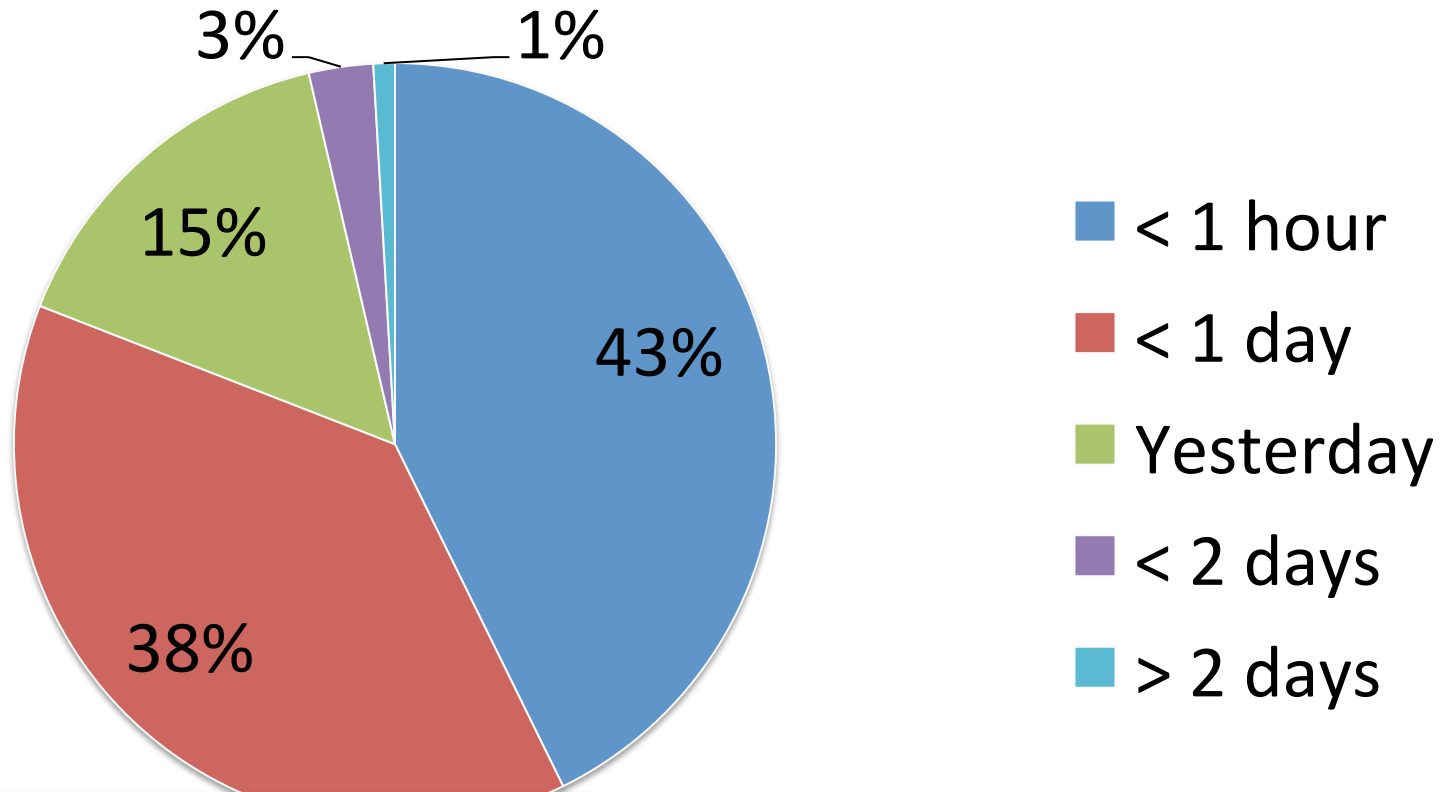
Mostly fresh memories

Context: Where

Very indoor

Context: How

35%

50%

15%

Mixed devices

- Desktop PC
- Laptop / Tablet
- Mobile

Info Needs: Target Time

Recency is important

Info Needs: Seasonal Interests

- Long Term — 24%
- Seasonal — 21%
- Recent — 41%
- Other — 14%
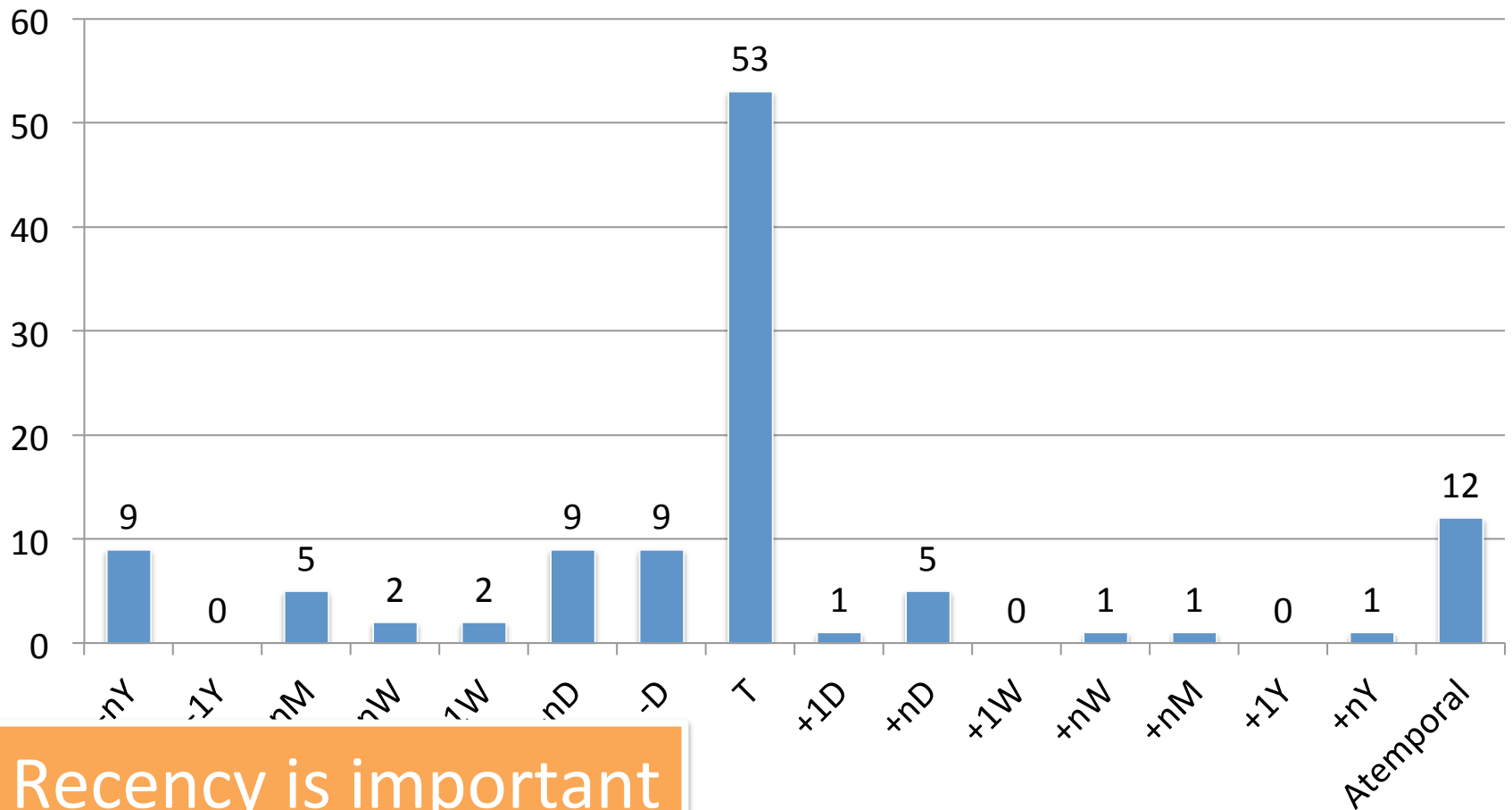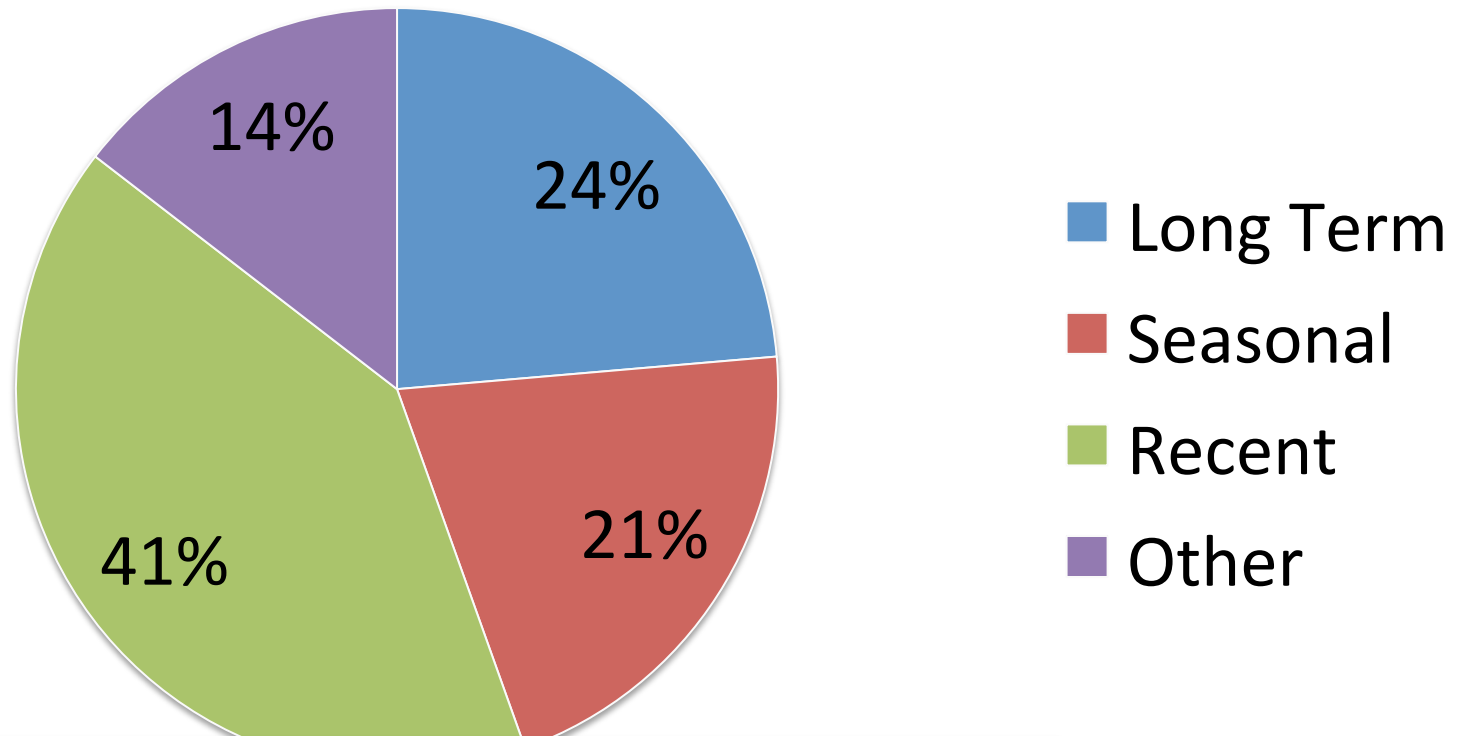
76% is some kind of temporal needs

Info Needs: Fresh, Re-find, Tech

Table 5: Freshness, re-findability, and technicality of information needs.

Freshness	Frequency	Percentage
Freshness was important	67	60.9%
Freshness was not so important	43	39.1%
Total	110	100.0%

Re-finding	Frequency	Percentage
Have searched before	72	65.5%
Have never searched before	38	34.5%
Total	110	100.0%

Technicality	Frequency	Percentage
Searching technical information	16	14.5%
Searching general information	94	85.5%
Total	110	100.0%

Process: Search Time

- < 5 min
- 5 - 10 min
- 10 - 30 min
- 30 min - 1h
- > 1h
- N/A

45%
31%
14%
3%
3%
4%
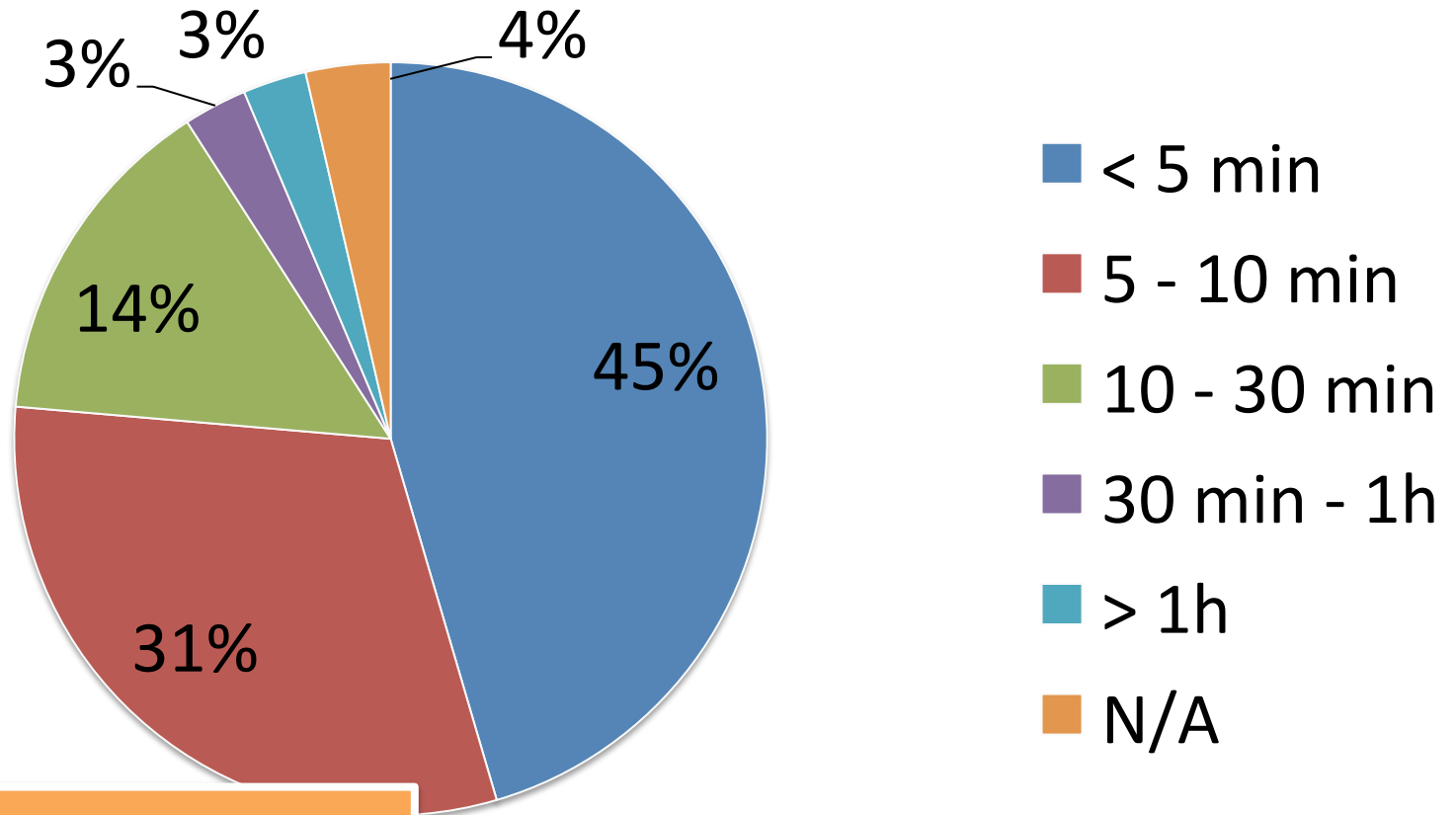
Nothing unusual?

Process: Difficulty

- Website — 9%
- Query — 24%
- Search Results — 35%
- None — 32%
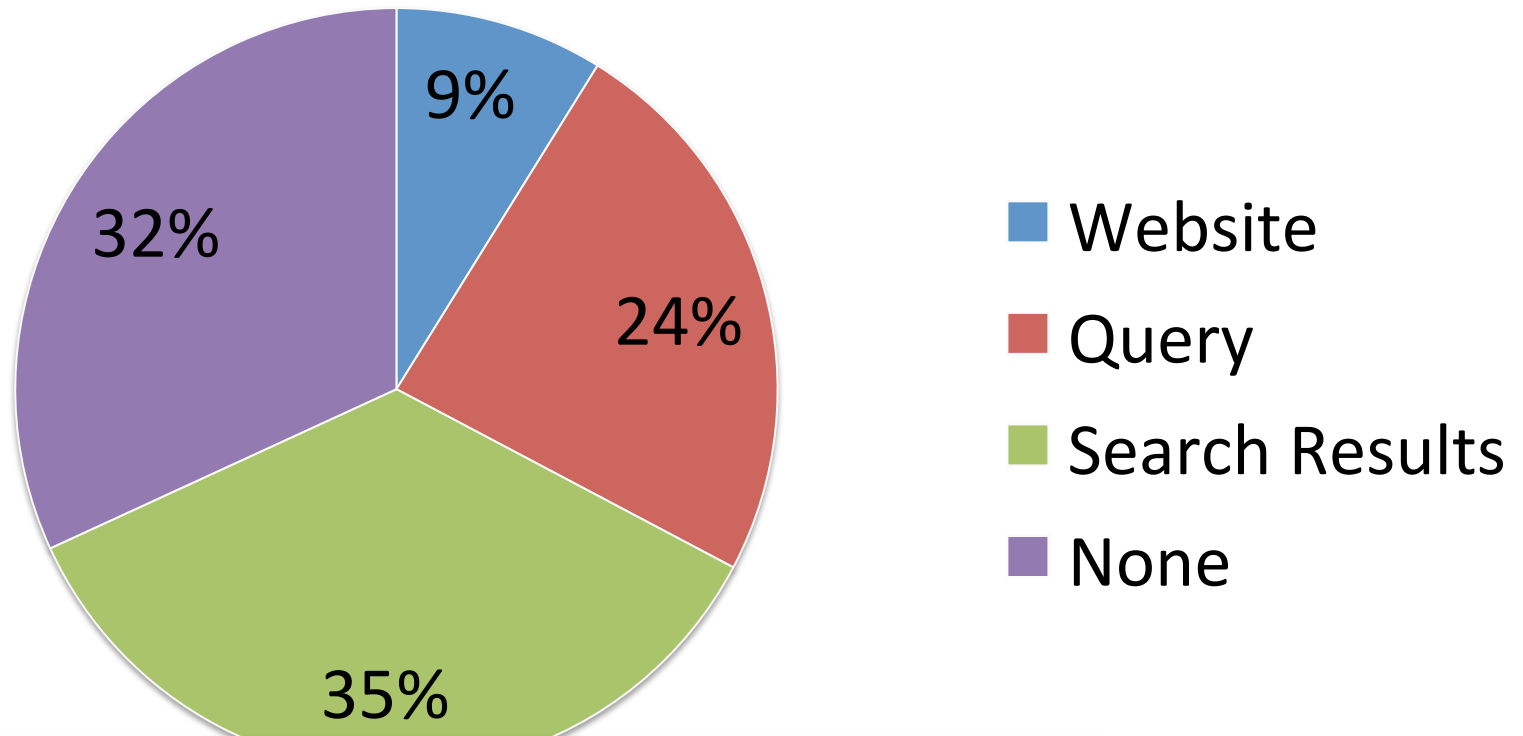
2/3 still have a difficulty (in general)

Outcome: Found?

2%

23%

75%

- Yes
- Partially
- No
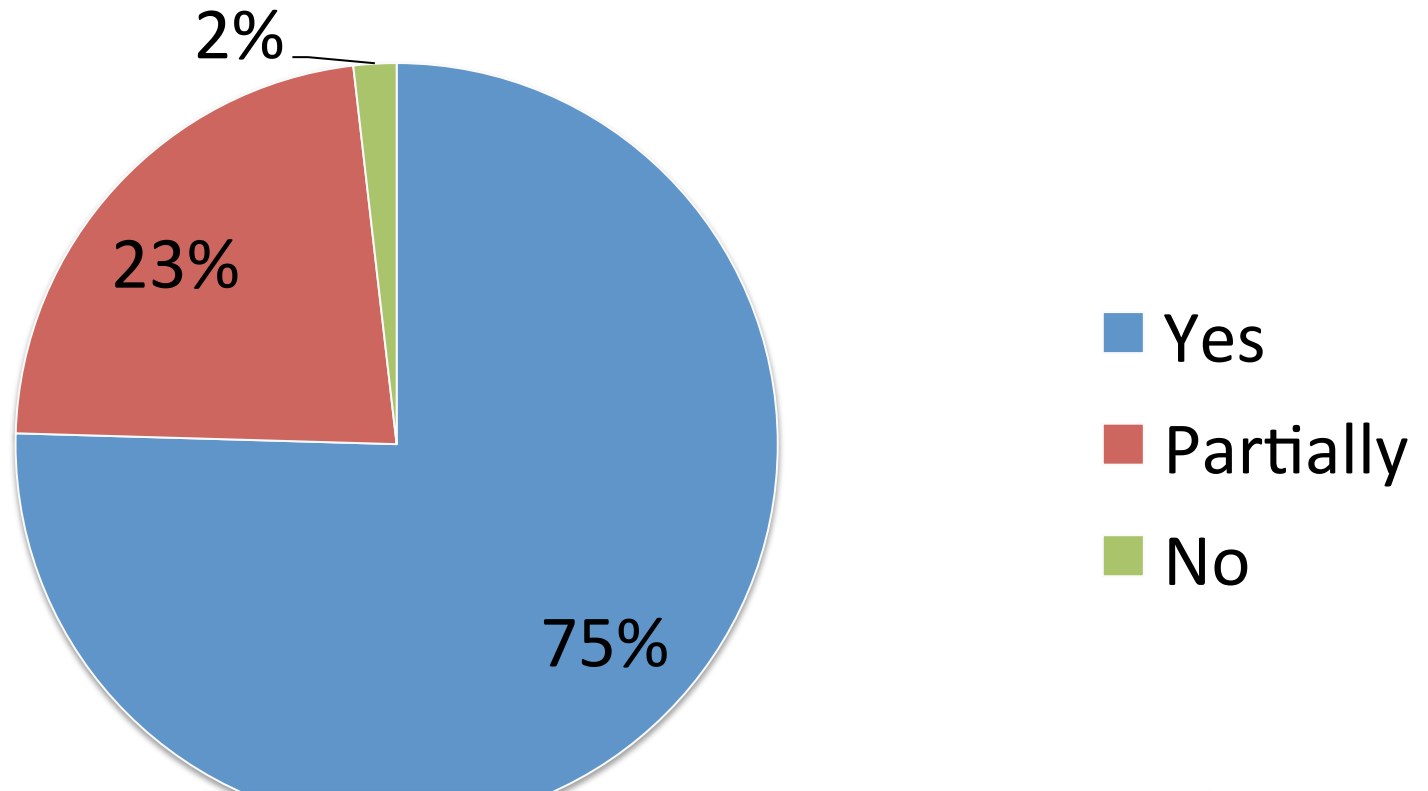
Yet managed to find relevant information

Outcome: Satisfaction

2%

13%

15%

70%

- Very Satisfied
- Satisfied
- Either
- Not Satisfied
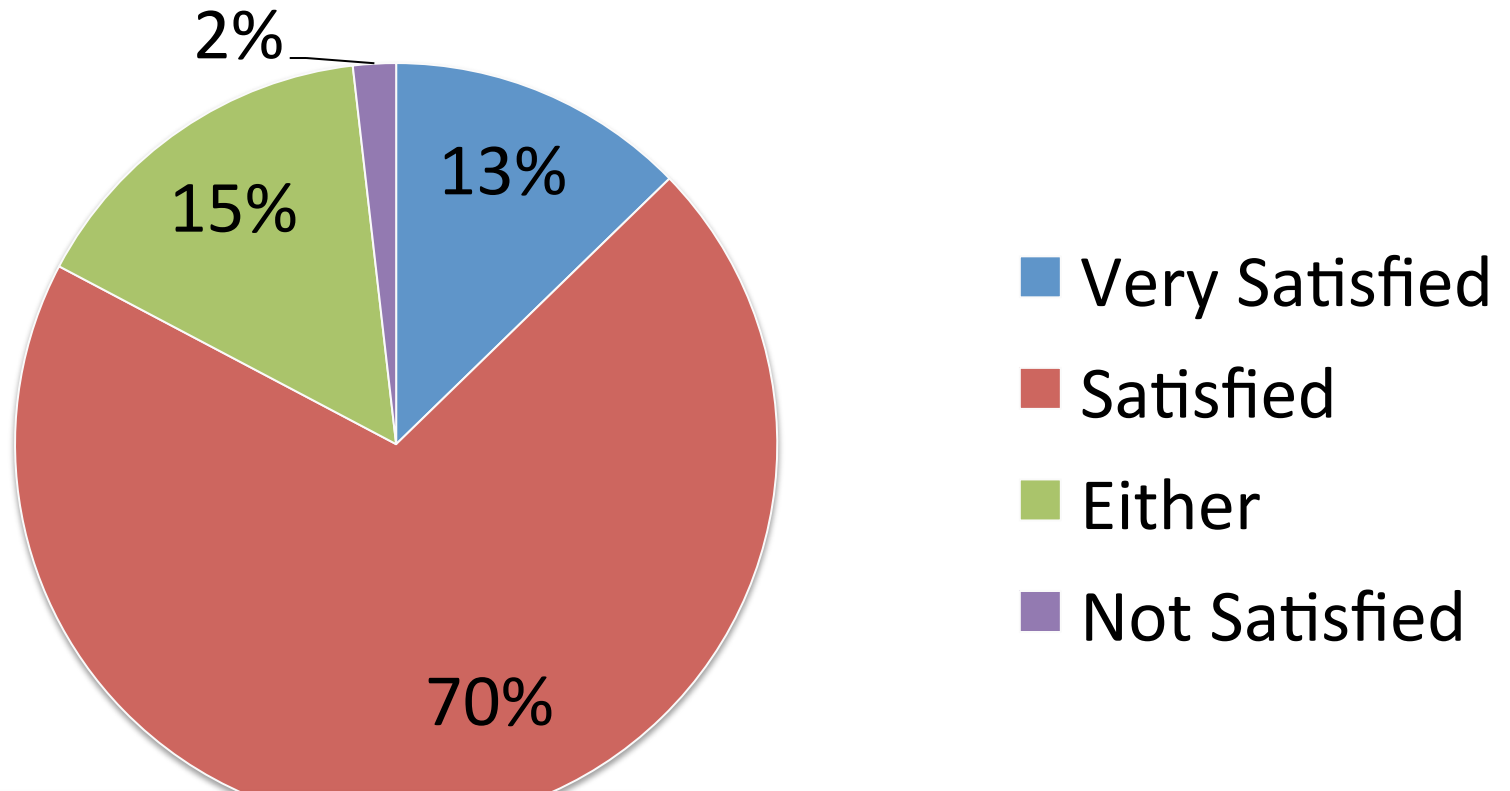
Still 17% are not happy

Outcome: Information Use

35% was not used within a day

Bar chart of information use by time category:
- 0 (Now): 28
- < 1D: 18
- < nD: 26
- < 1W: 2
- < nW: 9
- < 1M: 2
- < nM: 2
- < 1Y: 1
- < nY: 1
- No Plan: 21
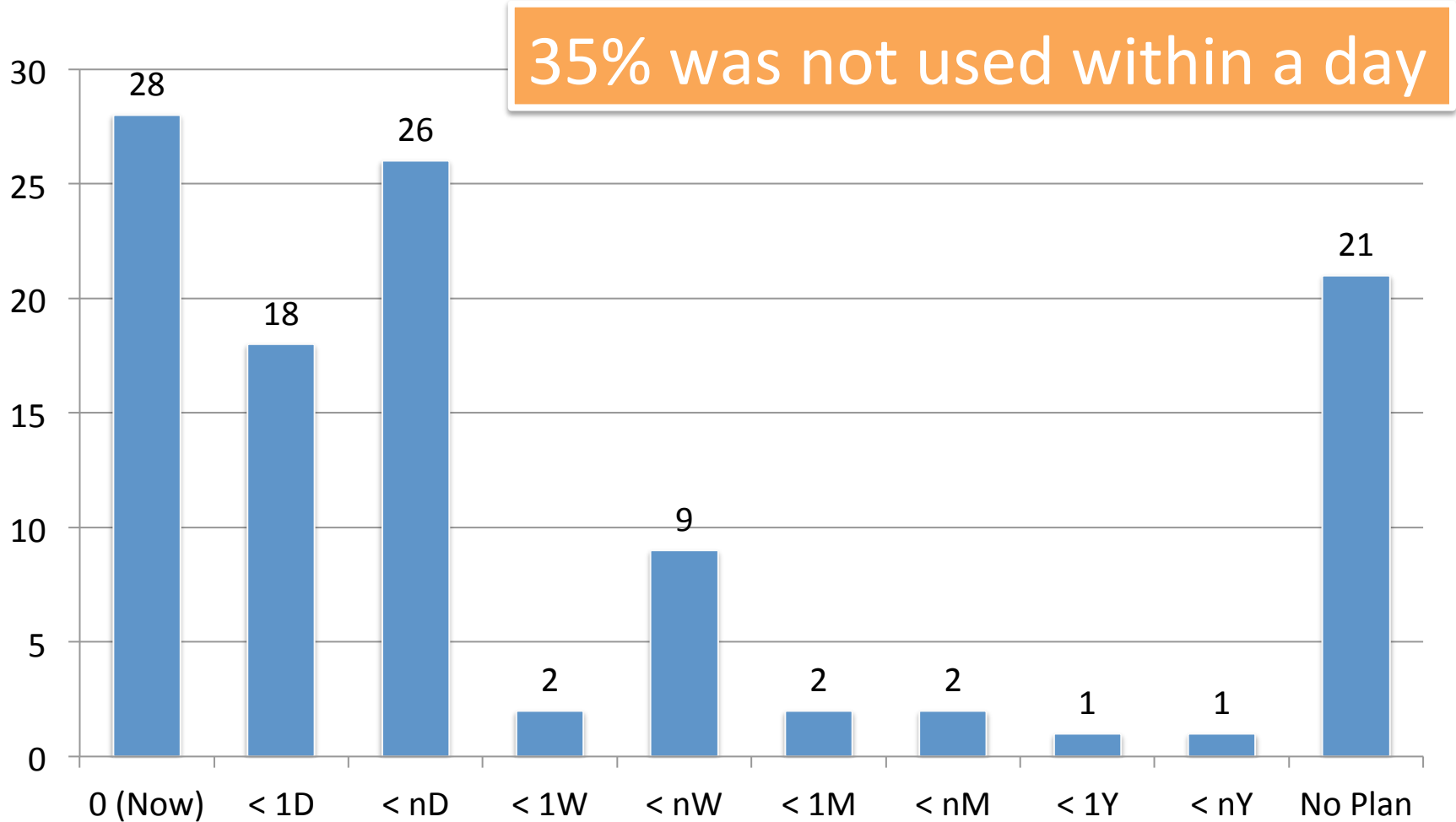
Correlation

**Table 10: Gross Correlation analyses by Spearman. Question number are based on Appendix A. Significant correlation ($p \le .05$) is highlighted.**

	Q2 Device	Q3 Location	Q5 Target	Q6 Season	Q7 Specific	Q8 Fresh	Q9 Re-find	Q12 Clarity	Q13 Use	Q14 Time	Q15 Found
Q2 Device to search	1.00										
Q3 Location of search	-0.04	1.00									
Q5 Target time of information	-0.14	**0.22**	1.00								
Q6 Seasonal needs	**-0.22**	0.08	**0.20**	1.00							
Q7 Specificity of needs	0.08	-0.01	**-0.20**	**-0.20**	1.00						
Q8 Freshness of needs	-0.07	-0.06	-0.04	0.11	**0.30**	1.00					
Q9 Re-findability	-0.08	0.04	0.06	**0.18**	-0.03	**0.24**	1.00				
Q12 Clarity of needs	0.05	0.03	-0.04	-0.10	-0.09	-0.05	-0.07	1.00			
Q13 Information Use	-0.08	-0.14	-0.07	-0.04	0.01	-0.09	-0.08	**0.24**	1.00		
Q14 Time taken to search	-0.02	-0.14	0.03	-0.08	0.06	-0.11	**-0.22**	**0.27**	0.16	1.00	
Q15 Rel doc found	-0.12	0.00	-0.03	0.05	0.13	0.11	-0.14	**0.39**	**0.19**	**0.25**	1.00
Q18 Satisfaction	-0.15	0.10	-0.02	-0.02	0.11	-0.07	-0.06	**0.39**	0.09	**0.26**	**0.47**
	Device Q2	Location Q3	Target Q5	Season Q6	Specific Q7	Fresh Q8	Re-find Q9	Clarity Q12	Use Q13	Time Q14	Found Q15

Focusing on some of significant pairs

Location vs. Target Time

Recent information when outside or stops

No future information sought from offices
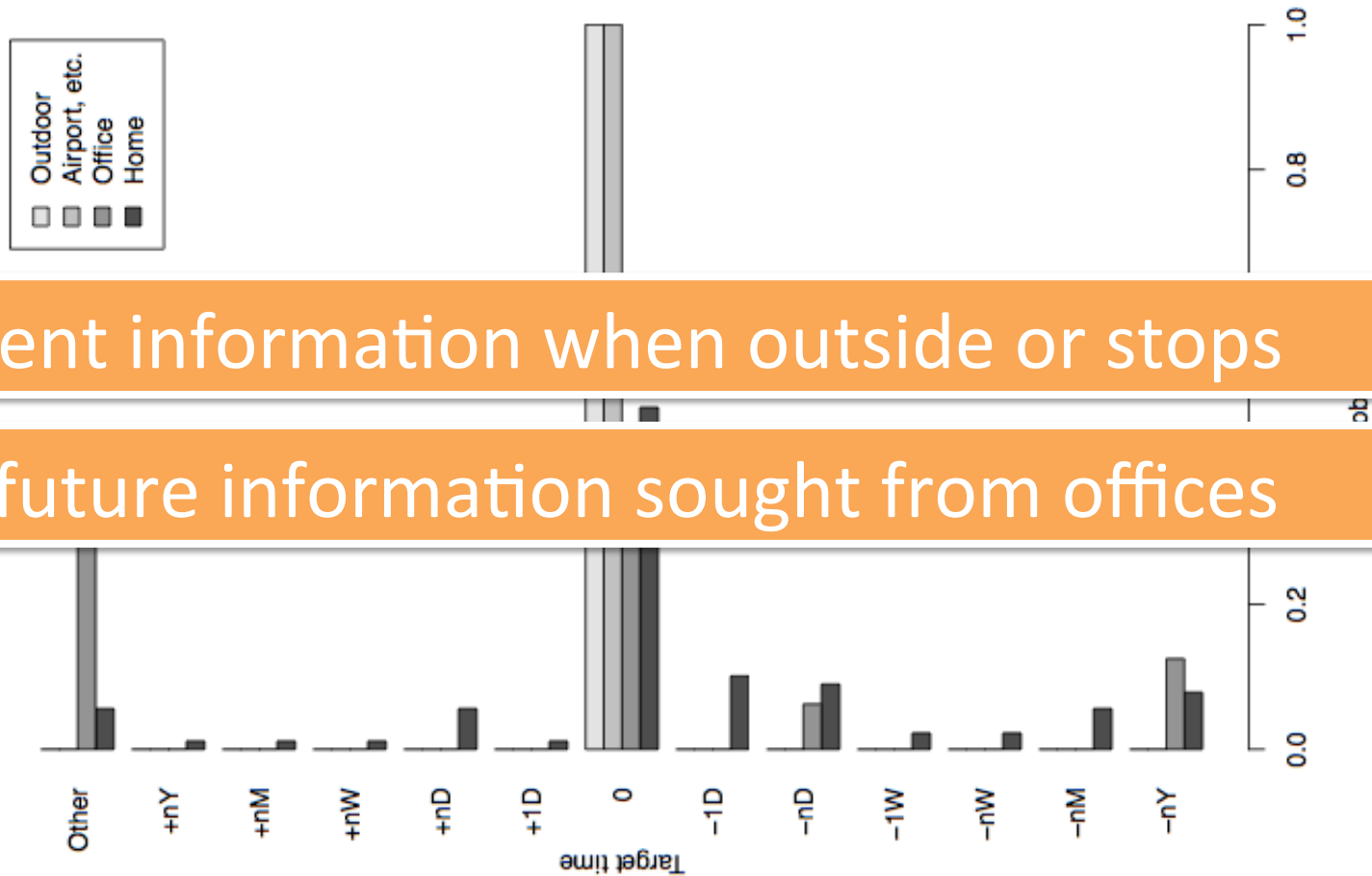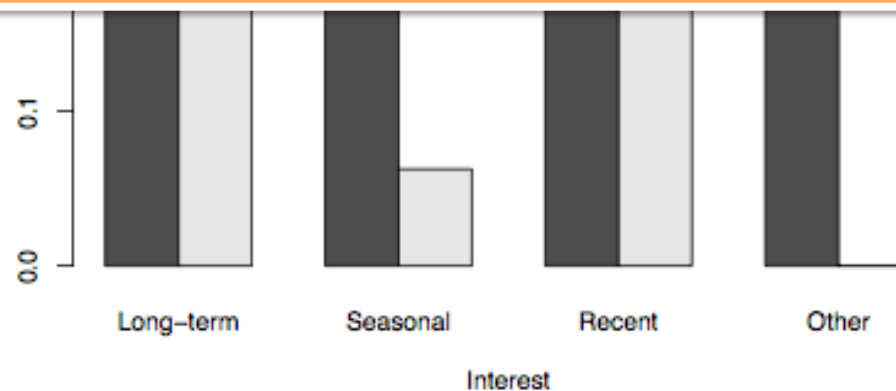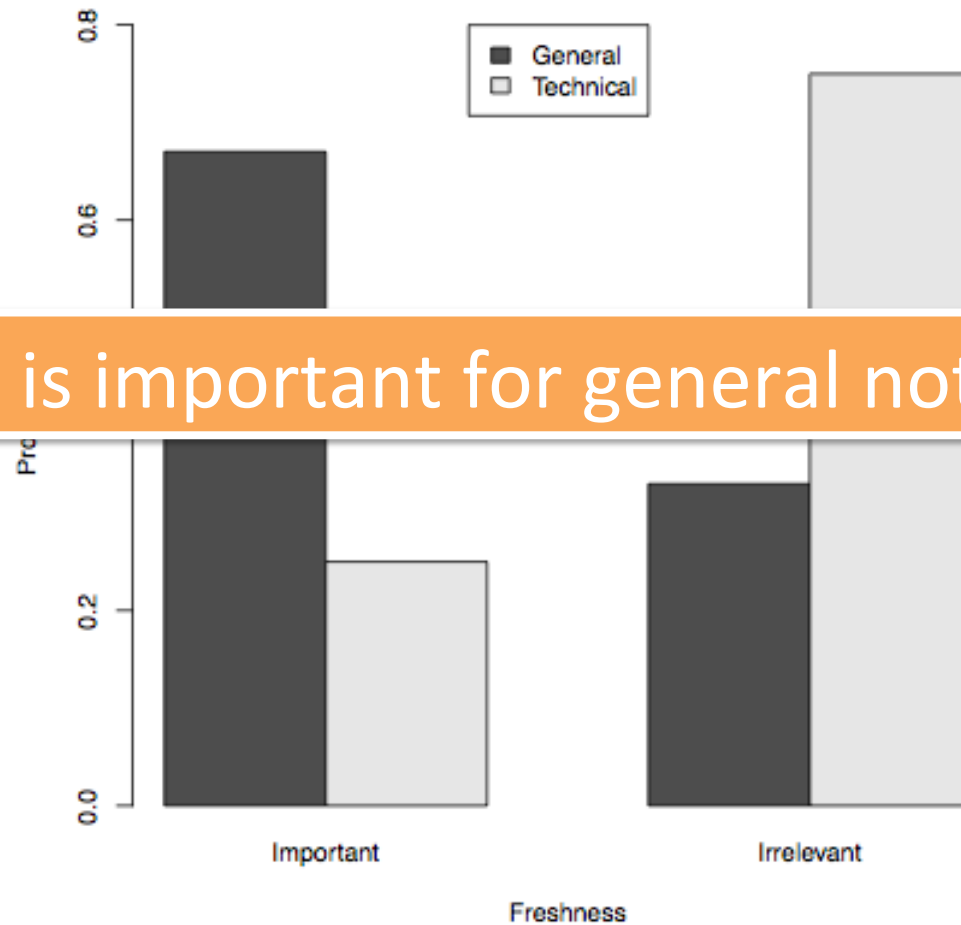
Technicality vs. Seasonal

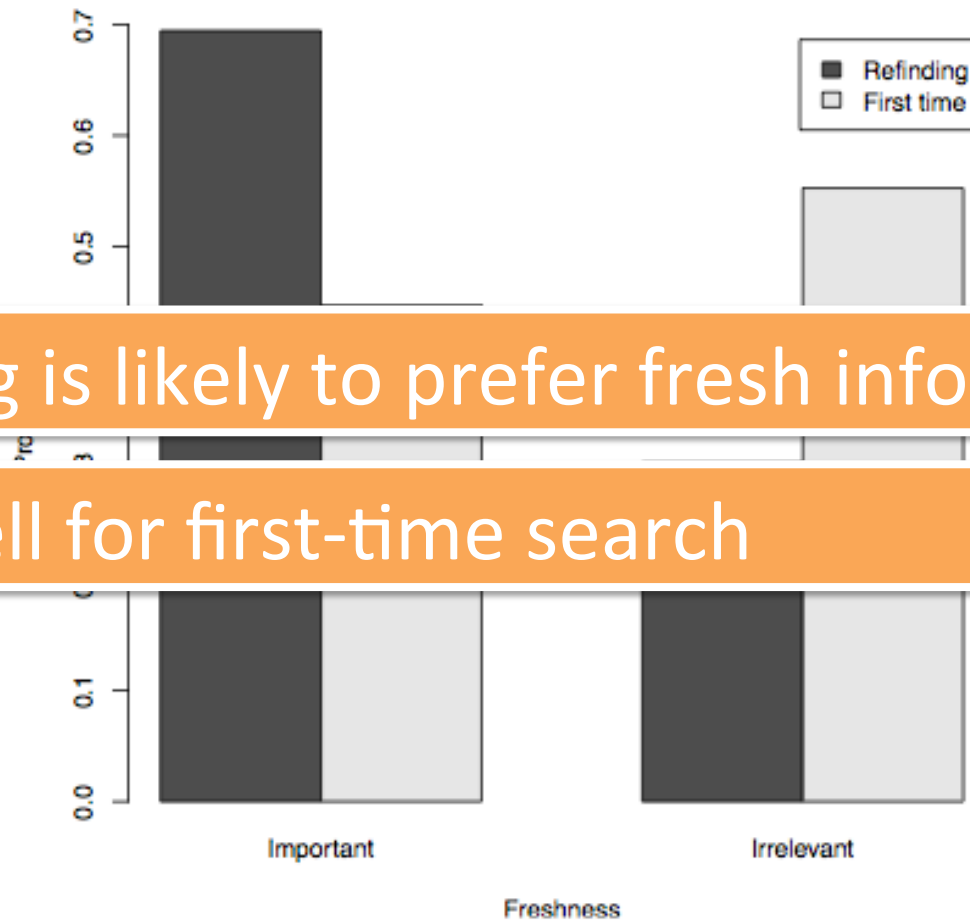**Technical information tends to be long-term or recent interests**

Freshness vs. Technicality

Freshness is important for general not technical

Re-finding vs. Freshness

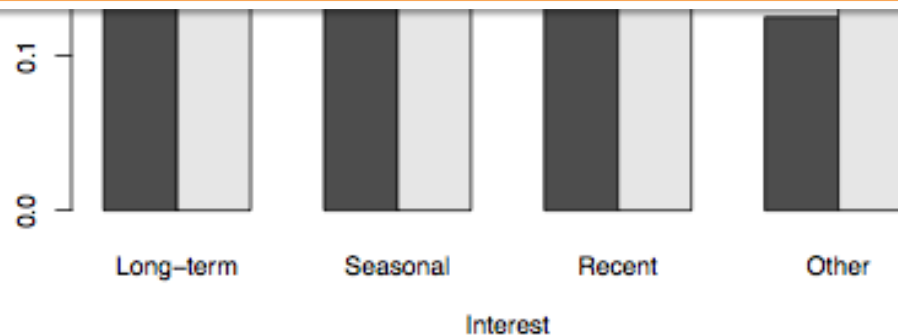Re-finding is likely to prefer fresh information

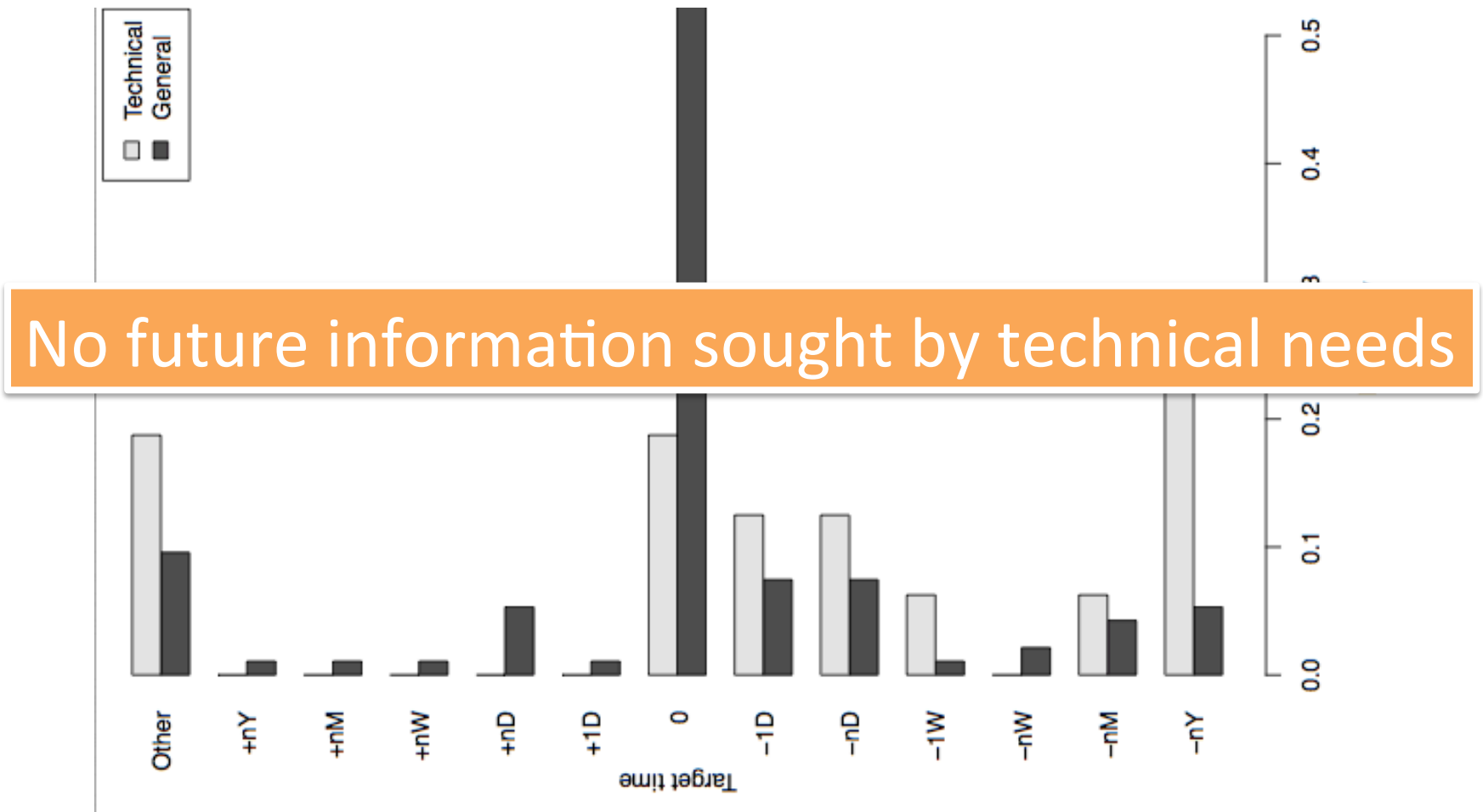Cannot tell for first-time search

Re-finding vs. Seasonal

First time search is likely to prefer recent info

Cannot tell for re-finding search

Technicality vs. Target Time

No future information sought by technical needs

Qualitative analyses

- A lot of TV is going on
- A lot of Internet browsing is going on
- Weak information needs
- Future temporal expression can be difficult
  - Project X is supposed to complete by the end of January 2015.
  - Temporal incidents in fictions

Summary

- Questionnaire-based CIA was carried out to gain broader & deeper understanding of temporal information seeking behaviour

- Results reinforcing the importance of recency needs

- Seasonal interests, technicality, target time, re-finding, and freshness can all interplay to affect people's temporal search behaviour.

- How can we deal with the gap between search and use?

- Smaller proportion of future information search is natural cause or lack of advancement?

References

1. O. Alonso, R. Baeza-Yates, J. Strötgen, M. and Gertz. Temporal information retrieval: Challanges and opportunities. In *TWAW 2011 Workshop,* 2011.

2. K. Berberich, S. Bedathur, O. Alonso, and G. Weikum. A Language Modeling Approach for Temporal Information Needs. In *ECIR 2010,* pages 13–25, 2012.

3. K. Berberich, S. Bedathur, T. Neumann, and G. Weikum. A Time Machine for Text Search. In *SIGIR 2007,* pages 519–526, 2007.

4. R. Campos, G. Dias, and A.M. Jorge. An Exploratory Study on the impact of Temporal Features on the Classification and Clustering of Future-Related Web Documents. In *EPIA 2011.*

5. B. M. Evans and E. H. Chi. An elaborated model of social search. *Inf. Process. Manage.,* 46(6): 656–678, 2010.

6. R. Jones and F. Diaz. Temporal profiles of queries. *ACM Transactions on Information Systems,* 25(3):656–678, 2007.

7. N. Kanhabua, and K. Nørvåg. Determining Time of Queries for Re-Ranking Search Results. In *ECDL 2010..*

8. A. Khodaei, and O. Alonso. Temporally-Aware Signals for Social Search. In *TAIA 2012 Workshop,* 2012.

9. A. Kulkarni, J. Teevan, K.M. Svore and S.T. Dumais. Understanding temporal query dynamics. In *WSDM 2011,* pages 167–176.

10. D. Metzler, R. Jones, F. Peng, and R. Zhang. Improving search relevance for implicitly temporal queries. In *SIGIR 2009,* 2009.

11. Morris, M.R. A survey of collaborative web search practices. In *SIGCHI 2008,* pages 1657–1660, 2008.

12. S. Nunes, C. Ribeiro and G. David. Use of temporal expressions in web search. In *ECIR 2008,* pages 580–584, 2008.

13. M. Shokouhi. Detecting seasonal queries by time-series analysis. In *SIGIR 2011,* pages 1171–1172, 2011.

14. J. Teevan and E. Adar. Information re-retrieval: repeat queries in yahoo's logs. In *SIGIR 2007,* pages 151–158, 2007.