

# Using Principal Component Analysis to Better Understand Behavioral Measures and their Effects

School of Information and Library Science  
University of North Carolina at Chapel Hill



Anita Crescenzi



Jaime Arguello



# Motivation



# Important Questions in IIR

- How are search behaviors influenced by task characteristics?
  - ▶ developing tools to support users
- How do search behaviors influence post-task perceptions?
  - ▶ evaluating systems
  - ▶ behavioral measures → search experience

# Prior Work

- How are search behaviors influenced by task characteristics?
  - ▶ **Type:** known-item vs. exploratory search
  - ▶ **Goal:** specific vs. amorphous
  - ▶ **Product:** factual vs. intellectual
  - ▶ **Complexity:** simple (memorization) vs. complex (evaluation/decision-making)

# Prior Work

- How do search behaviors influence post-task perceptions?
  - ▶ Difficulty
  - ▶ Frustration
  - ▶ Time pressure
  - ▶ Engagement (intellectual and emotional investment)

# A Common Approach (with limitations)...

- Compute a wide-range of behavioral measures and examine how they differ based on task characteristics or perceptions

## Predicting Search Task Difficulty

Jaime Arguello

University of North Carolina at Chapel Hill  
jarguello@unc.edu

**Abstract.** Search task difficulty refers to a user's assessment about the amount of effort required to complete a search task. Our goal in this work is to learn predictive models of search task difficulty. We evaluate features derived from the user's interaction with the search engine as well as features derived from the user's level of interest in the task and level of prior knowledge in the task domain. In addition to user-interaction features used in prior work, we evaluate features generated from scroll and mouse-movement events on the SERP. In some situations, we may prefer a system that can predict search task difficulty early in the search session. To this end, we evaluate features in terms of whole-session evidence and first-round evidence, which excludes all interactions starting with the second query. Our results found that the most predictive features were different for whole-session vs. first-round prediction, that mouseover features were effective for first-round prediction, and that level of interest and prior knowledge features did not improve performance.

[Arguello, ECIR 2014]

# A Common Approach (with limitations)...

- 44 behavioral measures derived from...
  - ▶ Queries
  - ▶ Clicks
  - ▶ Bookmarks
  - ▶ Mouseover events
  - ▶ Scroll events
  - ▶ Session duration

Whole-Session Analysis		
	easy	difficult
<b>Query Features</b>		
NumQueries	1.810 (1.462)	2.373 (1.641) <sup>▲</sup>
AvgQueryLength	5.398 (2.980)	5.779 (3.702)
NumQueryTerms	9.073 (8.251)	12.448 (10.333) <sup>▲</sup>
UniqueQueryTerms	6.504 (3.666)	8.091 (5.039) <sup>▲</sup>
TokenTypeRatio	1.315 (0.628)	1.471 (0.590) <sup>▲</sup>
AvgStopwords	0.201 (0.212)	0.204 (0.196)
AvgNonStopwords	0.799 (0.212)	0.796 (0.196)
NumAOLQueries	0.286 (0.705)	0.295 (0.731)
NumQuestionQueries	0.286 (0.573)	0.336 (0.625)
<b>Click Features</b>		
NumClicks	3.263 (2.481)	4.618 (3.292) <sup>▲</sup>
AvgClicks	2.161 (1.739)	2.425 (2.033)
AvgClickRank	2.704 (1.737)	3.701 (3.517) <sup>▲</sup>
AvgTimeToFirstClick	8.613 (8.278)	8.351 (7.062)
NumViews	2.815 (2.055)	3.793 (2.623) <sup>▲</sup>
AvgViews	1.901 (1.507)	2.040 (1.703)
AvgViewRank	2.697 (1.795)	3.713 (3.499) <sup>▲</sup>
NumPageClicks	0.092 (0.450)	0.282 (0.937) <sup>▲</sup>
NumAbandon	0.294 (0.779)	0.378 (0.755) <sup>▲</sup>
PercentAbandon	0.078 (0.178)	0.106 (0.196) <sup>▲</sup>
<b>Bookmark Features</b>		
NumBook	2.336 (1.559)	2.722 (1.509) <sup>▲</sup>
AvgBook	1.620 (1.258)	1.548 (1.238)
AvgBookRank	2.713 (1.865)	3.900 (3.793) <sup>▲</sup>
NumQueriesWithBook	1.359 (0.790)	1.651 (0.905) <sup>▲</sup>
PercentQueriesWithBook	0.875 (0.229)	0.814 (0.257) <sup>▼</sup>
NumQueriesWithoutBook	0.451 (1.020)	0.722 (1.205) <sup>▲</sup>
PercentQueresWithoutBook	0.125 (0.229)	0.186 (0.257) <sup>▲</sup>
NumClicksWithoutBook	0.927 (1.521)	1.896 (2.821) <sup>▲</sup>
PercentClicksWithoutBook	0.184 (0.242)	0.275 (0.279) <sup>▲</sup>
NumViewsWithoutBook	0.479 (0.996)	1.071 (2.103) <sup>▲</sup>
PercentViewsWithoutBook	0.105 (0.193)	0.176 (0.253) <sup>▲</sup>
<b>Mouse Features</b>		
TotalMouseovers	23.039 (32.056)	42.602 (52.086) <sup>▲</sup>
AvgMouseovers	12.307 (13.160)	16.185 (15.026) <sup>▲</sup>
MaxMouseover	5.734 (5.229)	8.664 (7.845) <sup>▲</sup>
AvgMaxMouseovers	4.486 (3.346)	5.943 (4.432) <sup>▲</sup>
<b>Scroll Features</b>		
TotalScrollDistance	105.532 (161.087)	182.154 (244.690) <sup>▲</sup>
AvgScrollDistance	55.118 (83.464)	64.382 (74.730) <sup>▲</sup>
MaxScrollPosition	39.067 (44.027)	53.610 (45.904) <sup>▲</sup>
AvgMaxScrollPosition	28.626 (36.012)	34.635 (35.586) <sup>▲</sup>
<b>Dwell-Time Features</b>		
TotalDwell	100.577 (112.695)	91.984 (105.488)
AvgDwell	42.998 (50.161)	29.351 (26.185) <sup>▼</sup>
<b>Duration</b>	193.596 (145.959)	223.964 (151.590) <sup>▲</sup>
<b>Interest</b>	2.838 (1.257)	2.635 (1.114)
<b>Prior Knowledge Features</b>		
PriorKnowledge	1.919 (0.937)	1.834 (0.845)
PriorSearch	1.437 (0.786)	1.378 (0.703)

# A Common Approach (with limitations)...

- More query abandonment, more difficulty...
- Searchers don't like issuing queries that do not yield relevant results



Whole-Session Analysis		
	easy	difficult
<b>Query Features</b>		
NumQueries	1.810 (1.462)	2.373 (1.641) <sup>▲</sup>
AvgQueryLength	5.298 (9.089)	5.770 (9.799)
<b>Click Features</b>		
NumClicks	3.263 (2.481)	4.618 (3.292) <sup>▲</sup>
AvgClicks	2.161 (1.739)	2.425 (2.033)
AvgClickRank	2.704 (1.737)	3.701 (3.517) <sup>▲</sup>
AvgTimeToFirstClick	8.613 (8.278)	8.351 (7.062)
NumViews	2.815 (2.055)	3.793 (2.623) <sup>▲</sup>
AvgViews	1.901 (1.507)	2.040 (1.703)
AvgViewRank	2.697 (1.795)	3.713 (3.499) <sup>▲</sup>
NumPageClicks	0.092 (0.450)	0.282 (0.937) <sup>▲</sup>
NumAbandon	0.294 (0.779)	0.378 (0.755) <sup>▲</sup>
PercentAbandon	0.078 (0.178)	0.106 (0.196) <sup>▲</sup>
NumViewsWithoutBook	0.479 (0.996)	1.071 (2.103) <sup>▲</sup>
PercentViewsWithoutBook	0.105 (0.193)	0.176 (0.253) <sup>▲</sup>
<b>Mouse Features</b>		
TotalMouseovers	23.039 (32.056)	42.602 (52.086) <sup>▲</sup>
AvgMouseovers	12.307 (13.160)	16.185 (15.026) <sup>▲</sup>
MaxMouseover	5.734 (5.229)	8.664 (7.845) <sup>▲</sup>
AvgMaxMouseovers	4.486 (3.346)	5.943 (4.432) <sup>▲</sup>
<b>Scroll Features</b>		
TotalScrollDistance	105.532 (161.087)	182.154 (244.690) <sup>▲</sup>
AvgScrollDistance	55.118 (83.464)	64.382 (74.730) <sup>▲</sup>
MaxScrollPosition	39.067 (44.027)	53.610 (45.904) <sup>▲</sup>
AvgMaxScrollPosition	28.626 (36.012)	34.635 (35.586) <sup>▲</sup>
<b>Dwell-Time Features</b>		
TotalDwell	100.577 (112.695)	91.984 (105.488)
AvgDwell	42.998 (50.161)	29.351 (26.185) <sup>▼</sup>
<b>Duration</b>	193.596 (145.959)	223.964 (151.590) <sup>▲</sup>
<b>Interest</b>	2.838 (1.257)	2.635 (1.114)
<b>Prior Knowledge Features</b>		
PriorKnowledge	1.919 (0.937)	1.834 (0.845)
PriorSearch	1.437 (0.786)	1.378 (0.703)

# A Common Approach (with limitations)...

- More queries, more difficulty...
- Searchers don't like issuing queries?



Whole-Session Analysis		
	easy	difficult
<b>Query Features</b>		
NumQueries	1.810 (1.462)	2.373 (1.641) <sup>▲</sup>
AvgQueryLength	5.398 (2.980)	5.779 (3.702)
NumQueryTerms	9.073 (8.251)	12.448 (10.333) <sup>▲</sup>
UniqueQueryTerms	6.504 (3.666)	8.091 (5.039) <sup>▲</sup>
TokenTypeRatio	1.315 (0.628)	1.471 (0.590) <sup>▲</sup>
AvgStopwords	0.201 (0.212)	0.204 (0.196)
AvgNonStopwords	0.799 (0.212)	0.796 (0.196)
NumAOLQueries	0.286 (0.705)	0.295 (0.731)
NumQuestionQueries	0.286 (0.573)	0.336 (0.625)
<b>Mouse Features</b>		
TotalMouseovers	23.039 (32.056)	42.602 (52.086) <sup>▲</sup>
AvgMouseovers	12.307 (13.160)	16.185 (15.026) <sup>▲</sup>
MaxMouseover	5.734 (5.229)	8.664 (7.845) <sup>▲</sup>
AvgMaxMouseovers	4.486 (3.346)	5.943 (4.432) <sup>▲</sup>
<b>Scroll Features</b>		
TotalScrollDistance	105.532 (161.087)	182.154 (244.690) <sup>▲</sup>
AvgScrollDistance	55.118 (83.464)	64.382 (74.730) <sup>▲</sup>
MaxScrollPosition	39.067 (44.027)	53.610 (45.904) <sup>▲</sup>
AvgMaxScrollPosition	28.626 (36.012)	34.635 (35.586) <sup>▲</sup>
<b>Dwell-Time Features</b>		
TotalDwell	100.577 (112.695)	91.984 (105.488)
AvgDwell	42.998 (50.161)	29.351 (26.185) <sup>▼</sup>
<b>Duration</b>	193.596 (145.959)	223.964 (151.590) <sup>▲</sup>
<b>Interest</b>	2.838 (1.257)	2.635 (1.114)
<b>Prior Knowledge Features</b>		
PriorKnowledge	1.919 (0.937)	1.834 (0.845)
PriorSearch	1.437 (0.786)	1.378 (0.703)



# A Common Approach (with limitations)...

- What underlying phenomena are queries capturing?
  - ▶ Abandonment?
  - ▶ More information?
  - ▶ Engagement?



Whole-Session Analysis		
	easy	difficult
<b>Query Features</b>		
NumQueries	1.810 (1.462)	2.373 (1.641) <sup>▲</sup>
AvgQueryLength	5.398 (2.980)	5.779 (3.702)
NumQueryTerms	9.073 (8.251)	12.448 (10.333) <sup>▲</sup>
UniqueQueryTerms	6.504 (3.666)	8.091 (5.039) <sup>▲</sup>
TokenTypeRatio	1.315 (0.628)	1.471 (0.590) <sup>▲</sup>
AvgStopwords	0.201 (0.212)	0.204 (0.196)
AvgNonStopwords	0.799 (0.212)	0.796 (0.196)
NumAOLQueries	0.286 (0.705)	0.295 (0.731)
NumQuestionQueries	0.286 (0.573)	0.336 (0.625)
<b>Mouse Features</b>		
TotalMouseovers	23.039 (32.056)	42.602 (52.086) <sup>▲</sup>
AvgMouseovers	12.307 (13.160)	16.185 (15.026) <sup>▲</sup>
MaxMouseover	5.734 (5.229)	8.664 (7.845) <sup>▲</sup>
AvgMaxMouseovers	4.486 (3.346)	5.943 (4.432) <sup>▲</sup>
<b>Scroll Features</b>		
TotalScrollDistance	105.532 (161.087)	182.154 (244.690) <sup>▲</sup>
AvgScrollDistance	55.118 (83.464)	64.382 (74.730) <sup>▲</sup>
MaxScrollPosition	39.067 (44.027)	53.610 (45.904) <sup>▲</sup>
AvgMaxScrollPosition	28.626 (36.012)	34.635 (35.586) <sup>▲</sup>
<b>Dwell-Time Features</b>		
TotalDwell	100.577 (112.695)	91.984 (105.488)
AvgDwell	42.998 (50.161)	29.351 (26.185) <sup>▼</sup>
<b>Duration</b>	193.596 (145.959)	223.964 (151.590) <sup>▲</sup>
<b>Interest</b>	2.838 (1.257)	2.635 (1.114)
<b>Prior Knowledge Features</b>		
PriorKnowledge	1.919 (0.937)	1.834 (0.845)
PriorSearch	1.437 (0.786)	1.378 (0.703)



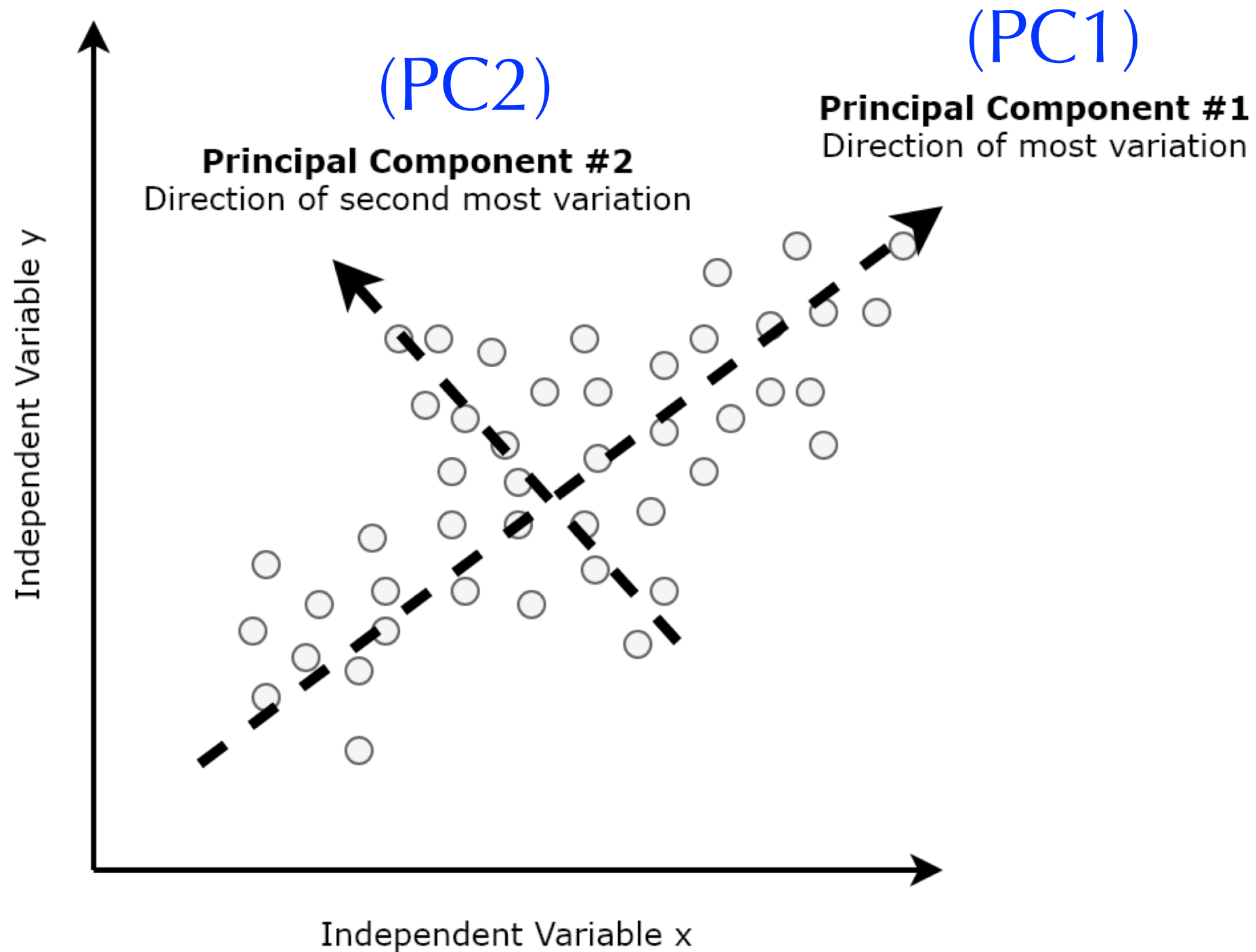
# This paper

- Using principal component analysis (PCA) to understand:
- The behavioral phenomena captured by behavioral measures
- The influence of behavioral phenomena on perceptions
  - ▶ workload
  - ▶ difficulty
  - ▶ time pressure
  - ▶ factors of user engagement
  - ▶ knowledge gains

# What is PCA?

- A common dimensionality reduction technique
- Clustering groups datapoints
- PCA groups measures/features used to “represent” datapoints
- The  $i^{\text{th}}$  principal component is a variance-maximizing direction that is orthogonal to all previous  $(i-1)$  components













# What is PCA?



- Dimensionality reduction achieved by representing the data using only **PC1**

# Secondary Analysis of Data From Three Studies (Studies 1-3)

- **Common objective:** bookmark pages relevant to the task

	System Manipulation	Task -Type Manipulation	Minimum # Bookmarks	Time Constraint
Study 1				
Study 2				
Study 3				

# Study 1

[Arguello et al., TOIS 2019]

- Laboratory study
- 32 participants each completed 4 tasks of the same type
- Two aggregated search layouts: blocked vs. blended
- Bookmark at least 10 pages in 15 minutes
- Post-task perception(s):
  - ▶ workload

# Layout Conditions: Blocked


Everything Web Images Videos News Shopping

**search blue** russia 2018 fifa world cup Search

WEB RESULTS: 1-10 of 47

**images**

**Image Results**



→ more for russia 2018 fifa world cup images

**News Results**

**Russia 2018 FIFA World Cup: Team previews for all 32 teams**  
Time is running out as the **2018 FIFA World Cup** kicks off on June 14 when host **Russia** takes on Saudi Arabia. And it's time for you to get caught up on every team, what there is to know, when they'll play and how you can watch. Below you will find links to ...  
[CBSSports.com](#) 19 hours ago

**World Cup 2018: Schedule, group matches, how to watch live**  
Group stage matches for the **2018 World Cup** in **Russia** are set ... Thirty-one soccer teams earned the right to play for the **FIFA World Cup** trophy and **Russia** was given an automatic bid into the competition to complete the field as is customary for the ...  
[Sporting News](#) 15 hours ago

**Kia Provides 424 Vehicles for Official Use at 2018 FIFA World Cup Russia**  
Kia Motors has provided 424 vehicles to assist operations at the much-anticipated **2018 FIFA World Cup Russia**. The handover of the provisional fleet took place May 31 at the Luzhniki Stadium in Moscow, where the **World Cup** Final will be held on July 15.  
[The Drive](#) seconds ago

→ more for russia 2018 fifa world cup news

**Video Results**

**12 DAYS TO GO! Goals galore at Switzerland 1954**  
Duration: 55 seconds

**Brazilian Player Alex Sandro full information (Best Brazilian player 2018 in Fifa World Cup Russia)**  
Duration: 3 minutes

**FIFA World Cup Russia 2018 ? Official Promo ??**  
Duration: 2 minutes

**Brazil vs Argentina Bangla Funny Dubbing | FIFA World Cup 2018 Russia | Binodon TV | Messi, Neymar**  
Duration: 2 minutes

**FIFA World Cup Russia 2018 ? Official Promo**  
Duration: 3 minutes

→ more for russia 2018 fifa world cup videos

**Shopping Results**

**PANINI FIFA World Cup Russia 2018. New. Pick Any 10 Stickers. Free Shipping.**  
\$5.0 Condition: Brand New

**2018 PANINI RUSSIA FIFA WORLD CUP SOCCER STICKERS - FACTORY SEALED 50 PACK BOX**  
\$37.95 Condition: Brand New

**Panini Fifa World Cup Russia 2018 - Individual Stickers Pick any 10**  
\$4.99 Condition: Brand New

**2018 Panini Russia FIFA World Cup Soccer Stickers - Factory Sealed 50 Pack Box**  
\$44.99 Condition: Brand New

**Panini FIFA World Cup Russia 2018 Stickers Pick 20 30 40 50 THOUSANDS AVAILABLE**  
\$9.9 Condition: Like New

→ more for russia 2018 fifa world cup shopping

Everything  
Web  
Images  
Videos  
News  
Shopping

web

news

video


shopping

1 2 3 4 5 next



# Layout Conditions: Blended (Truncated)

Everything
Web
Images
Videos
News
Shopping



WEB RESULTS: 1-10 of 47

**Everything**

Web






Images

Videos


News

Shopping


[Image Results](#)


[2018 FIFA World Cup Russia? - FIFA.com](#)  
Are you planning to visit **Russia** for the **2018 FIFA World Cup**?? Welcome2018.com is the tourist portal of the Russian government and provides fans travelling to **Russia** with useful information about the host nation of the next **FIFA World Cup**.  
<https://www.fifa.com/worldcup>

 [Russia 2018 FIFA World Cup: Team previews for all 32 teams](#)  
Time is running out as the **2018 FIFA World Cup** kicks off on June 14 when host **Russia** takes on Saudi Arabia. And it's time for you to get caught up on every team, what there is to know, when they'll play and how you can watch. Below you will find links to ...  
[CBSSports.com](#) 18 hours ago


[2018 FIFA World Cup - Wikipedia](#)  
A ceremony in Moscow launching the countdown from 1,000 days until the **2018 FIFA Football World Cup** begins in **Russia**.  
[https://en.wikipedia.org/wiki/2018\\_FIFA\\_World\\_Cup](https://en.wikipedia.org/wiki/2018_FIFA_World_Cup)

 [12 DAYS TO GO! Goals galore at Switzerland 1954](#)  
The 12 goals scored in Austria 7-5 Switzerland in the 1954 FIFA World Cup Switzerland? quarter-finals make it the highest-scoring match in the history of the global ...  
Duration: 55 seconds


[2018 FIFA World Cup Russia? Tourist Portal](#)  
**Russia** is the **world's** largest country. More than half of the population are football fans or at least follow the sport occasionally. **FIFA World Cup 2018** in **Russia** is a phenomenally powerful magnet.  
<welcome2018.com/en>

 [PANINI FIFA World Cup Russia 2018. New, Pick Any 10 Stickers. Free Shipping.](#)  
\$5.0  
Condition: Brand New

[Russia 2018 World Cup schedule: Complete fixture, dates ...](#)  
The **2018 FIFA World Cup** is right around the corner. In just about three weeks, the greatest soccer players on the planet will take center stage in **Russia**.  
<https://www.cbssports.com/soccer/world-cup/news/russia-2018-world...>

 [World Cup 2018: Schedule, group matches, how to watch live](#)  
Group stage matches for the **2018 World Cup** in **Russia** are set ... Thirty-one soccer teams earned the right to play for the **FIFA World Cup** trophy and **Russia** was given an automatic bid into the competition to complete the field as is customary for the ...  
[Sporting News](#) 15 hours ago

[Fan Guide: 2018 FIFA World Cup Tournament | FOX Sports](#)  
Find everything you need to know about the **2018 FIFA World Cup** in **Russia**, including key dates, qualifying & draw info, match schedules, stadiums & more at **FOX Sports**.  
<https://www.foxsports.com/soccer/fifa-world-cup/fan-guide>

 [Kia Provides 424 Vehicles for Official Use at 2018 FIFA World Cup Russia](#)  
Kia Motors has provided 424 vehicles to assist operations at the much-anticipated **2018 FIFA World Cup Russia**. The handover of the provisional fleet took place May 31 at the Luzhnik Stadium in Moscow, where the **World Cup** Final will be held on July 15.  
[The Drive](#) seconds ago

images

web

news

web

video

web

shopping

web

news

web

news

• • •

# Study 2

[Capra et al., SIGIR 2018]



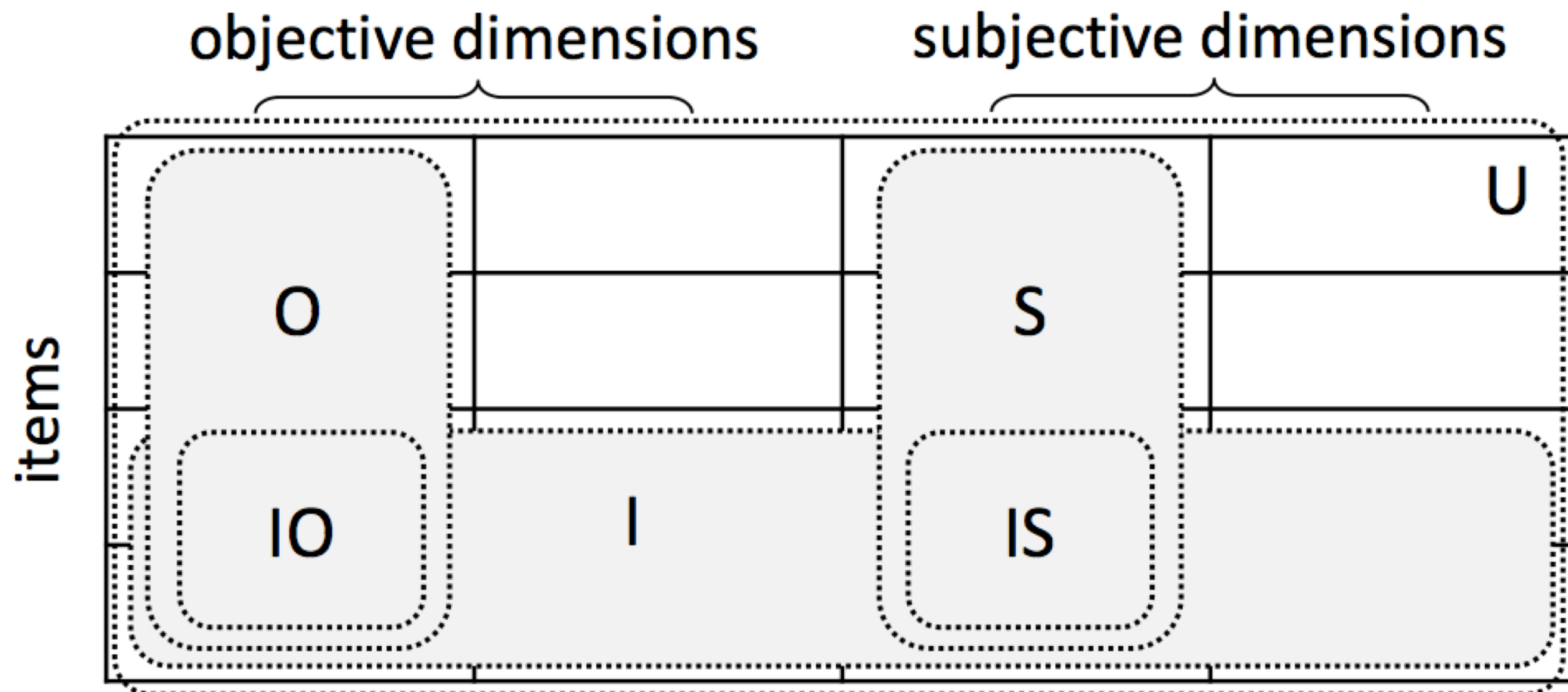
- 144 participants each completed 6 tasks of different types
- Same search interface
- No minimum bookmarks, no time limit
- Post-task perceptions of user engagement:
  - ▶ **Focused attention:** felt immersed, lost track of time
  - ▶ **Reward:** effort + success
  - ▶ **Aesthetic appeal**
  - ▶ **Perceived usability**



## Study 2

[Capra et al., SIGIR 2018]

- **Comparative tasks:** comparing items along dimensions
- Manipulated constraints specified in the task description
- Tasks with a specified dimension were more difficult



# Study 3

## [Arguello ECIR 2014]

- 20 tasks of varying levels of cognitive complexity
- Each task completed by 30 participants
- $20 \times 30 = 600$  search sessions
- Same search interface
- No minimum bookmarks, no time limit
- Post-task perception(s):
  - ▶ Difficulty
  - ▶ Time pressure
  - ▶ Knowledge gains

# Study 3

## [Arguello ECIR 2014]

- **Cognitive task complexity:** mental activities central to the task:
  - ▶ **Remember:** memorize, regurgitate
  - ▶ **Understand:** internalize, summarize, exemplify
  - ▶ **Analyze:** compare, contrast, differentiate
  - ▶ **Evaluate:** judge, critique, prioritize
  - ▶ **Create:** create something new

# Behavioral Measures

1. Queries
2. Avg. query length
3. Question queries
4. Queries w/o bookmarks
5. Queries w/o clicks
6. Queries w/o scrolls
7. Queries w/o mouseovers
8. Quick reformulations
9. Repeated intent queries
10. Bookmarks
11. Clicks
12. Views
13. Clicks w/o bookmarks
14. Views w/o bookmarks
15. Avg. click rank
16. Avg. view rank
17. Avg. bookmark rank
18. Paginations
19. Scroll distance
20. Mouseovers
21. Avg. mouseover rank
22. Mouseovers w/o clicks
23. Avg. query-to-click time
24. Time to 1st click
25. Time to 1st bookmark
26. Avg. landing pg. dwell time
27. Tot. landing pg. dwell time
28. Avg. time b/w events
29. Completion time
30. Unique queries
31. Unique query terms
32. Unique URLs clicked

# Behavioral Measures

1. Queries
2. Avg. query length
3. Question queries
4. Queries w/o bookmarks
5. Queries w/o clicks
6. Queries w/o scrolls
7. Queries w/o mouseovers
8. Quick reformulations
9. Repeated intent queries

10. Bookmarks
11. Clicks
12. Views
13. Clicks w/o bookmark
14. Views w/o bookmark
15. Avg. click rank
16. Avg. view rank
17. Avg. bookmark rank
18. Paginations

19. Scroll distance
20. Mouseovers
21. Avg. mouseover rank
22. Mouseovers w/o clicks
23. Avg. query-to-click time
24. Time to 1st click
25. Time to 1st bookmark
26. Avg. landing pg. dwell time
27. Tot. landing pg. dwell time
28. Avg. time b/w events

**Querying behavior**

# Behavioral Measures

1. Queries
2. Avg. query length
3. Question queries
4. Queries w/o bookmarks
5. Queries w/o clicks
6. Queries w/o scrolls
7. Queries w/o mouseovers
8. Quick reformulations
9. Repeated intent queries
10. Bookmarks
11. Clicks
12. Views
13. Clicks w/o bookmark
14. Views w/o bookmark
15. Avg. click rank
16. Avg. view rank
17. Avg. bookmark rank
18. Paginations
19. Scroll distance
20. Mouseovers
21. Avg. mouseover rank
22. Mouseovers w/o clicks
23. Avg. query-to-click time
24. Time to 1st click
25. Time to 1st bookmark
26. Avg. landing pg. dwell time
27. Tot. landing pg. dwell time
28. Avg. time b/w events

**Bookmarks**

# Behavioral Measures

## 1. Queries

### SERP-level interactions

9. Repeated intent queries

10. Bookmarks

11. Clicks

12. Views

13. Clicks w/o bookmarks

14. Views w/o bookmarks

15. Avg. click rank

16. Avg. view rank

17. Avg. bookmark rank

18. Paginations

19. Scroll distance

20. Mouseovers

21. Avg. mouseover rank

22. Mouseovers w/o clicks

23. Avg. query-to-click time

24. Time to 1st click

25. Time to 1st bookmark

26. Avg. landing pg. dwell time

27. Tot. landing pg. dwell time

28. Avg. time b/w events

29. Completion time

30. Unique queries

31. Unique query terms

32. Unique URLs clicked

# Behavioral Measures

## 1. Queries

### Temporal characteristics

- 9. Repeated intent queries
- 10. Bookmarks
- 11. Clicks
- 12. Views
- 13. Clicks w/o bookmarks
- 14. Views w/o bookmarks
- 15. Avg. click rank
- 16. Avg. view rank
- 17. Avg. bookmark rank
- 18. Paginations

- 19. Scroll distance
- 20. Mouseovers
- 21. Avg. mouseover rank
- 22. Mouseovers w/o clicks
- 23. Avg. query-to-click time
- 24. Time to 1st click
- 25. Time to 1st bookmark
- 26. Avg. landing pg. dwell time
- 27. Tot. landing pg. dwell time
- 28. Avg. time b/w events
- 29. Completion time
- 30. Unique queries
- 31. Unique query terms
- 32. Unique URLs clicked



# Behavioral Measures

## 1. Queries

### Deviation from others' search strategies

9. Repeated intent queries
10. Bookmarks
11. Clicks
12. Views
13. Clicks w/o bookmarks
14. Views w/o bookmarks
15. Avg. click rank
16. Avg. view rank
17. Avg. bookmark rank
18. Paginations

19. Scroll distance
20. Mouseovers
21. Avg. mouseover rank
22. Mouseovers w/o clicks
23. Avg. query-to-click time
24. Time to 1st click
25. Time to 1st bookmark
26. Avg. landing pg. dwell time
27. Tot. landing pg. dwell time
28. Avg. time b/w events
29. Completion time
30. Unique queries
31. Unique query terms
32. Unique URLs clicked

# PCA Decisions

- Eigen-decomposition of covariance vs. correlation matrix
- Choosing the number of components
- Choosing a rotation technique
- Combining measures into components

# Results

- Using principal component analysis (PCA) to understand:
- The behavioral phenomena captured by behavioral measures
- The influence of behavioral phenomena on perceptions
  - ▶ workload
  - ▶ difficulty
  - ▶ time pressure
  - ▶ factors of user engagement
  - ▶ knowledge gains

Results: PCA output

# PCA Output: Study 1

- Component loadings are correlation values in the range [-1,+1]

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# PCA Output: Study 1

- 6-component solution explained 76% of the variance

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# PCA Output: Study 1

- AbandQs: query abandonment

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# PCA Output: Study 1

- AbandCs: click abandonment

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70



# PCA Output: Study 1

- DeepSERP: SERP exploration at low ranks

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# PCA Output: Study 1

- **Pace:** pace of interaction

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# PCA Output: Study 1

- NLQs: natural language queries

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# PCA Output: Study 1

- SlowCs: time to click

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# PCA Output: Study 1

- Measure ambiguity: completion time

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# PCA Output: Study 2

- 5-component solution explained 70% of the variance

	PC1 (AbandQs)	PC2 (Effort)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)
QueriesWOClicks	0.94	0.10	0.07	0.08	0.03
QuickReforms	0.88	0.14	-0.01	-0.01	-0.03
QueriesWOBooks	0.86	0.38	0.06	0.13	0.05
Queries	0.82	0.52	-0.02	0.09	0.06
QueriesWOMouse	0.81	-0.11	0.13	0.04	-0.04
RepeatedIntentQs	0.79	0.46	0.04	0.11	0.06
QueriesWOScrolls	0.74	0.01	-0.14	-0.02	0.05
UniqueQueries	0.70	0.59	-0.06	0.05	0.04
Clicks	0.23	0.83	0.31	0.10	0.00
UniqueURLs	0.15	0.79	0.23	0.03	0.10
ClicksWOBooks	0.24	0.74	0.36	0.10	0.00
Mouseovers	0.28	0.71	0.34	0.13	-0.04
MouseWOClicks	0.39	0.68	0.43	0.10	0.04
ScrollDistance	0.22	0.68	0.45	0.17	0.06
Bookmarks	-0.02	0.55	-0.01	-0.03	-0.02
Paginations	0.13	0.53	0.64	0.08	0.08
AvgMouseRank	-0.06	0.26	0.86	0.03	0.00
AvgClickRank	-0.05	0.23	0.86	0.02	-0.05
AvgBookRank	-0.06	0.18	0.81	-0.01	-0.05
AvgTimeBWEvents	-0.10	-0.04	-0.02	0.89	0.06
CompletionTime	0.15	0.40	0.09	0.85	0.03
TimeToFirstBook	0.20	-0.01	0.08	0.82	-0.01
QuestionQueries	0.14	0.02	0.11	-0.04	0.84
AvgQueryLength	-0.13	-0.20	0.16	-0.01	0.78
Avg1stClickTime	-0.05	0.13	-0.16	0.20	0.26
TimeToFirstClick	0.01	0.01	-0.03	0.01	0.10
UniqueQueryTerms	0.44	0.35	-0.11	-0.06	0.36



# PCA Output: Study 2

- **Effort:** amount of information required by the task

	PC1 (AbandQs)	PC2 (Effort)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)
QueriesWOClicks	0.94	0.10	0.07	0.08	0.03
QuickReforms	0.88	0.14	-0.01	-0.01	-0.03
QueriesWOBooks	0.86	0.38	0.06	0.13	0.05
Queries	0.82	0.52	-0.02	0.09	0.06
QueriesWOMouse	0.81	-0.11	0.13	0.04	-0.04
RepeatedIntentQs	0.79	0.46	0.04	0.11	0.06
QueriesWOScrolls	0.74	0.01	-0.14	-0.02	0.05
UniqueQueries	0.70	0.59	-0.06	0.05	0.04
Clicks	0.23	0.83	0.31	0.10	0.00
UniqueURLs	0.15	0.79	0.23	0.03	0.10
ClicksWOBooks	0.24	0.74	0.36	0.10	0.00
Mouseovers	0.28	0.71	0.34	0.13	-0.04
MouseWOClicks	0.39	0.68	0.43	0.10	0.04
ScrollDistance	0.22	0.68	0.45	0.17	0.06
Bookmarks	-0.02	0.55	-0.01	-0.03	-0.02
Paginations	0.13	0.53	0.64	0.08	0.08
AvgMouseRank	-0.06	0.26	0.86	0.03	0.00
AvgClickRank	-0.05	0.23	0.86	0.02	-0.05
AvgBookRank	-0.06	0.18	0.81	-0.01	-0.05
AvgTimeBWEvents	-0.10	-0.04	-0.02	0.89	0.06
CompletionTime	0.15	0.40	0.09	0.85	0.03
TimeToFirstBook	0.20	-0.01	0.08	0.82	-0.01
QuestionQueries	0.14	0.02	0.11	-0.04	0.84
AvgQueryLength	-0.13	-0.20	0.16	-0.01	0.78
Avg1stClickTime	-0.05	0.13	-0.16	0.20	0.26
TimeToFirstClick	0.01	0.01	-0.03	0.01	0.10
UniqueQueryTerms	0.44	0.35	-0.11	-0.06	0.36



# PCA Output: Study 2

- Measure ambiguity: number of queries

	PC1 (AbandQs)	PC2 (Effort)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)
QueriesWOClicks	0.94	0.10	0.07	0.08	0.03
QuickReforms	0.88	0.14	-0.01	-0.01	-0.03
QueriesWOBooks	0.86	0.38	0.06	0.13	0.05
Queries	0.82	0.52	-0.02	0.09	0.06
QueriesWOMouse	0.81	-0.11	0.13	0.04	-0.04
RepeatedIntentQs	0.79	0.46	0.04	0.11	0.06
QueriesWOScrolls	0.74	0.01	-0.14	-0.02	0.05
UniqueQueries	0.70	0.59	-0.06	0.05	0.04
Clicks	0.23	0.83	0.31	0.10	0.00
UniqueURLs	0.15	0.79	0.23	0.03	0.10
ClicksWOBooks	0.24	0.74	0.36	0.10	0.00
Mouseovers	0.28	0.71	0.34	0.13	-0.04
MouseWOClicks	0.39	0.68	0.43	0.10	0.04
ScrollDistance	0.22	0.68	0.45	0.17	0.06
Bookmarks	-0.02	0.55	-0.01	-0.03	-0.02
Paginations	0.13	0.53	0.64	0.08	0.08
AvgMouseRank	-0.06	0.26	0.86	0.03	0.00
AvgClickRank	-0.05	0.23	0.86	0.02	-0.05
AvgBookRank	-0.06	0.18	0.81	-0.01	-0.05
AvgTimeBWEvents	-0.10	-0.04	-0.02	0.89	0.06
CompletionTime	0.15	0.40	0.09	0.85	0.03
TimeToFirstBook	0.20	-0.01	0.08	0.82	-0.01
QuestionQueries	0.14	0.02	0.11	-0.04	0.84
AvgQueryLength	-0.13	-0.20	0.16	-0.01	0.78
Avg1stClickTime	-0.05	0.13	-0.16	0.20	0.26
TimeToFirstClick	0.01	0.01	-0.03	0.01	0.10
UniqueQueryTerms	0.44	0.35	-0.11	-0.06	0.36

# PCA Output: Study 3

- 8-component solution explained 76% of the variance

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (Effort)	PC6 (NLQs)	PC7 (SlowCs)
QueriesWOClicks	0.91	-0.05	0.12	0.01	0.03	0.06	0.08
QueriesWOBooks	0.90	0.27	0.03	0.05	0.05	0.00	0.03
QuickReforms	0.83	0.11	0.30	-0.02	0.07	-0.07	0.08
Queries	0.82	0.28	0.01	0.03	0.41	-0.02	0.01
QueriesWOScrolls	0.79	0.09	-0.13	-0.07	0.27	0.02	0.02
QueriesWOMouse	0.78	-0.02	0.05	0.03	-0.09	0.05	0.15
ClicksWOBooks	0.22	0.87	0.17	0.12	-0.01	-0.05	-0.08
ViewsWOBooks	0.10	0.87	0.14	0.09	-0.02	-0.04	-0.04
Clicks	0.18	0.75	0.33	0.06	0.41	-0.06	-0.10
Views	0.09	0.69	0.34	0.02	0.50	-0.07	-0.08
AvgMouseRank	0.05	0.12	0.94	0.06	0.09	-0.04	0.11
AvgViewRank	-0.03	0.08	0.93	0.05	0.03	0.09	-0.06
AvgClickRank	0.00	0.07	0.93	0.06	0.04	0.07	0.09
AvgBookRank	-0.03	0.10	0.92	0.05	0.02	0.10	-0.05
Paginations	0.13	0.14	0.76	0.00	0.06	-0.15	0.19
Mouseovers	0.18	0.40	0.71	0.05	0.32	-0.13	0.03
MouseWOClicks	0.33	0.26	0.63	0.03	0.32	-0.12	0.06
AvgTimeBWEvents	-0.05	-0.14	0.02	0.89	-0.02	-0.03	0.13
AvgDwellTime	-0.06	-0.10	-0.04	0.88	-0.05	-0.01	-0.04
TotalDwellTime	0.00	0.32	0.18	0.74	0.25	-0.04	-0.07
TimeToFirstBook	0.15	0.31	0.04	0.74	-0.13	0.03	0.25
CompletionTime	0.17	0.42	0.31	0.60	0.45	-0.03	0.08
Bookmarks	0.03	0.17	0.37	-0.05	0.78	-0.06	-0.08
UniqueQueries	0.33	0.10	-0.07	0.06	0.65	-0.04	0.08
UniqueURLs	0.09	0.47	0.38	0.06	0.50	0.08	-0.01
QuestionQueries	0.18	-0.07	-0.02	-0.02	0.11	0.81	-0.04
AvgQueryLength	-0.11	0.06	0.00	-0.03	-0.19	0.80	0.04
TimeToFirstClick	0.28	-0.10	0.08	0.09	0.00	-0.01	0.85
Avg1stClickTime	-0.05	0.04	0.06	0.19	0.03	0.09	0.76
ScrollDistance	0.10	0.00	0.06	-0.04	-0.03	-0.08	0.71
UniqueQueryTerms	-0.01	0.29	0.01	-0.03	0.11	0.12	0.14

# PCA Output: Study 3

- Measure ambiguity: number of page views

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (Effort)	PC6 (NLQs)	PC7 (SlowCs)
QueriesWOClicks	0.91	-0.05	0.12	0.01	0.03	0.06	0.08
QueriesWOBooks	0.90	0.27	0.03	0.05	0.05	0.00	0.03
QuickReforms	0.83	0.11	0.30	-0.02	0.07	-0.07	0.08
Queries	0.82	0.28	0.01	0.03	0.41	-0.02	0.01
QueriesWOScrolls	0.79	0.09	-0.13	-0.07	0.27	0.02	0.02
QueriesWOMouse	0.78	-0.02	0.05	0.03	-0.09	0.05	0.15
ClicksWOBooks	0.22	0.87	0.17	0.12	-0.01	-0.05	-0.08
ViewsWOBooks	0.10	0.87	0.14	0.09	-0.02	-0.04	-0.04
Clicks	0.18	0.75	0.33	0.06	0.41	-0.06	-0.10
Views	0.09	0.69	0.34	0.02	0.50	-0.07	-0.08
AvgMouseRank	0.05	0.12	0.94	0.06	0.09	-0.04	0.11
AvgViewRank	-0.03	0.08	0.93	0.05	0.03	0.09	-0.06
AvgClickRank	0.00	0.07	0.93	0.06	0.04	0.07	0.09
AvgBookRank	-0.03	0.10	0.92	0.05	0.02	0.10	-0.05
Paginations	0.13	0.14	0.76	0.00	0.06	-0.15	0.19
Mouseovers	0.18	0.40	0.71	0.05	0.32	-0.13	0.03
MouseWOClicks	0.33	0.26	0.63	0.03	0.32	-0.12	0.06
AvgTimeBWEvents	-0.05	-0.14	0.02	0.89	-0.02	-0.03	0.13
AvgDwellTime	-0.06	-0.10	-0.04	0.88	-0.05	-0.01	-0.04
TotalDwellTime	0.00	0.32	0.18	0.74	0.25	-0.04	-0.07
TimeToFirstBook	0.15	0.31	0.04	0.74	-0.13	0.03	0.25
CompletionTime	0.17	0.42	0.31	0.60	0.45	-0.03	0.08
Bookmarks	0.03	0.17	0.37	-0.05	0.78	-0.06	-0.08
UniqueQueries	0.33	0.10	-0.07	0.06	0.65	-0.04	0.08
UniqueURLs	0.09	0.47	0.38	0.06	0.50	0.08	-0.01
QuestionQueries	0.18	-0.07	-0.02	-0.02	0.11	0.81	-0.04
AvgQueryLength	-0.11	0.06	0.00	-0.03	-0.19	0.80	0.04
TimeToFirstClick	0.28	-0.10	0.08	0.09	0.00	-0.01	0.85
Avg1stClickTime	-0.05	0.04	0.06	0.19	0.03	0.09	0.76
ScrollDistance	0.10	0.00	0.06	-0.04	-0.03	-0.08	0.71
UniqueQueryTerms	-0.01	0.29	0.01	-0.03	0.11	0.12	0.14

# Important Trends

- Search sessions are characterized by similar phenomena
  - ▶ Effort, AbandQs, AbandCs, DeepSERP, Pace, ...
- Some measures are more ambiguous than others
  - ▶ **Study 1**: completion time → AbandCs, Pace
  - ▶ **Study 2**: queries → AbandQs, Effort
  - ▶ **Study 3**: page views → AbandCs, Effort
- Experimental design can influence a measure's meaning
  - ▶ **Study 1**: bookmarks → Pace (negative)
  - ▶ **Studies 3**: bookmarks → Effort

# Results

- Using principal component analysis (PCA) to understand:
- The behavioral phenomena captured by behavioral measures
- The influence of behavioral phenomena on perceptions
  - ▶ workload
  - ▶ difficulty
  - ▶ time pressure
  - ▶ factors of user engagement
  - ▶ knowledge gains

# Combining Measures into Components

- Linear combination (after normalization)

	PC1 (AbandQs)	PC2 (AbandCs)	PC3 (DeepSERP)	PC4 (Pace)	PC5 (NLQs)	PC6 (SlowCs)
QueriesWOBooks	0.87	0.09	0.27	0.04	0.06	-0.03
QueriesWOClicks	0.86	-0.03	0.21	0.21	-0.04	0.02
Queries	0.85	0.16	0.10	-0.29	0.15	0.04
RepeatedIntentQs	0.85	0.14	0.24	-0.02	0.02	0.12
QuickReforms	0.83	0.00	0.09	-0.18	-0.20	-0.05
UniqueQueries	0.82	0.15	0.09	-0.16	0.37	0.00
QueriesWOMouse	0.71	0.02	-0.08	0.15	-0.16	-0.09
UniqueQueryTerms	0.62	0.04	-0.05	0.04	0.36	0.30
QueriesWOScrolls	0.60	0.16	-0.44	-0.33	-0.03	-0.04
Clicks	0.16	0.83	0.22	-0.37	0.07	-0.18
ClicksWOBooks	0.18	0.79	0.25	0.13	0.02	-0.24
UniqueURLs	0.04	0.76	0.08	-0.05	0.12	0.09
CompletionTime	0.05	0.57	0.08	0.53	0.08	0.34
MouseWOClicks	0.34	0.18	0.85	-0.11	0.14	0.06
Mouseovers	0.16	0.40	0.82	-0.09	0.07	-0.02
Paginations	-0.09	0.18	0.78	0.01	-0.09	0.10
ScrollDistance	0.36	-0.06	0.77	-0.03	0.22	0.12
AvgTimeBWEvents	-0.26	-0.05	-0.07	0.83	-0.03	0.31
Bookmarks	-0.02	0.27	-0.07	-0.78	-0.01	0.09
TimeToFirstBook	0.06	0.19	-0.17	0.70	-0.12	0.22
QuestionQueries	0.02	0.08	0.13	0.01	0.83	-0.08
AvgQueryLength	0.01	0.10	0.03	-0.09	0.82	0.10
TimeToFirstClick	0.04	-0.02	0.02	0.12	0.06	0.82
Avg1stClickTime	0.05	-0.13	0.34	0.30	-0.04	0.70

# Effects on Perceptions: Study 1-3

Study/dependent variable	AbandQ	AbandC	DeepSERP	Pace	NLQs	SlowC	Effort
<b>Study 1</b>							
workload	.32***	.27**	.26**	.52***	.19*	.24**	–
<b>Study 2</b>							
focused attention	<i>ns</i>	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
reward	-.08**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	-.11***
aesthetic appeal	-.05**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
perceived usability	-.18***	–	-.12***	<i>ns</i>	<i>ns</i>	–	-.21***
<b>Study 3</b>							
difficulty	.13***	.23***	.15***	<i>ns</i>	<i>ns</i>	<i>ns</i>	.10*
time pressure	.10**	.23***	<i>ns</i>	.13**	<i>ns</i>	<i>ns</i>	.14***
knowledge increase	<i>ns</i>	<i>ns</i>	<i>ns</i>	.13 ( <i>p</i> = .052)	<i>ns</i>	<i>ns</i>	<i>ns</i>

- Multi-level modeling (think: regression)
- Results consistent with prior work



# Effects on Perceptions: Study 1

Study/dependent variable	AbandQ	AbandC	DeepSERP	Pace	NLQs	SlowC	Effort
<b>Study 1</b>							
workload	.32***	.27**	.26**	.52***	.19*	.24**	–
<b>Study 2</b>							
focused attention	<i>ns</i>	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
reward	-.08**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	-.11***
aesthetic appeal	-.05**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
perceived usability	-.18***	–	-.12***	<i>ns</i>	<i>ns</i>	–	-.21***
<b>Study 3</b>							
difficulty	.13***	.23***	.15***	<i>ns</i>	<i>ns</i>	<i>ns</i>	.10*
time pressure	.10**	.23***	<i>ns</i>	.13**	<i>ns</i>	<i>ns</i>	.14***
knowledge increase	<i>ns</i>	<i>ns</i>	<i>ns</i>	.13 ( <i>p</i> = .052)	<i>ns</i>	<i>ns</i>	<i>ns</i>

- Every component had a sig. positive effect on workload

# Effects on Perceptions: Study 2

Study/dependent variable	AbandQ	AbandC	DeepSERP	Pace	NLQs	SlowC	Effort
<b>Study 1</b>							
workload	.32***	.27**	.26**	.52***	.19*	.24**	–
<b>Study 2</b>							
focused attention	<i>ns</i>	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
reward	-.08**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	-.11***
aesthetic appeal	-.05**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
perceived usability	-.18***	–	-.12***	<i>ns</i>	<i>ns</i>	–	-.21***
<b>Study 3</b>							
difficulty	.13***	.23***	.15***	<i>ns</i>	<i>ns</i>	<i>ns</i>	.10*
time pressure	.10**	.23***	<i>ns</i>	.13**	<i>ns</i>	<i>ns</i>	.14***
knowledge increase	<i>ns</i>	<i>ns</i>	<i>ns</i>	.13 ( <i>p</i> = .052)	<i>ns</i>	<i>ns</i>	<i>ns</i>

- Several components had a sig. negative effect on reward, aesthetic appeal, and perceived usability

# Effects on Perceptions: Study 3

Study/dependent variable	AbandQ	AbandC	DeepSERP	Pace	NLQs	SlowC	Effort
<b>Study 1</b>							
workload	.32***	.27**	.26**	.52***	.19*	.24**	–
<b>Study 2</b>							
focused attention	<i>ns</i>	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
reward	-.08**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	-.11***
aesthetic appeal	-.05**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
perceived usability	-.18***	–	-.12***	<i>ns</i>	<i>ns</i>	–	-.21***
<b>Study 3</b>							
difficulty	.13***	.23***	.15***	<i>ns</i>	<i>ns</i>	<i>ns</i>	.10*
time pressure	.10**	.23***	<i>ns</i>	.13**	<i>ns</i>	<i>ns</i>	.14***
knowledge increase	<i>ns</i>	<i>ns</i>	<i>ns</i>	.13 ( <i>p</i> = .052)	<i>ns</i>	<i>ns</i>	<i>ns</i>

- Several components had a sig. positive effect on difficulty and time pressure

# Effects on Perceptions: Study 3

Study/dependent variable	AbandQ	AbandC	DeepSERP	Pace	NLQs	SlowC	Effort
<b>Study 1</b>							
workload	.32***	.27**	.26**	.52***	.19*	.24**	–
<b>Study 2</b>							
focused attention	<i>ns</i>	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
reward	-.08**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	-.11***
aesthetic appeal	-.05**	–	<i>ns</i>	<i>ns</i>	<i>ns</i>	–	<i>ns</i>
perceived usability	-.18***	–	-.12***	<i>ns</i>	<i>ns</i>	–	-.21***
<b>Study 3</b>							
difficulty	.13***	.23***	.15***	<i>ns</i>	<i>ns</i>	<i>ns</i>	.10*
time pressure	.10**	.23***	<i>ns</i>	.13**	<i>ns</i>	<i>ns</i>	.14***
knowledge increase	<i>ns</i>	<i>ns</i>	<i>ns</i>	.13 ( <i>p</i> = .052)	<i>ns</i>	<i>ns</i>	<i>ns</i>

- **Pace** had a marginally sig. positive effect on knowledge increase
- Participants perceived greater knowledge gains when they searched slower (spent more time reading pages)

# Conclusions

- PCA can help us:
  - interpret behavioral measures
  - study behavioral phenomena and perceptions
- Behavioral measures can be ambiguous (suggest multiple phenomena)
- Using PCA, ambiguous measures can be interpreted based on their correlations with other (more interpretable) measures
- Behavioral measures can “mean” different things depending on the experimental design



Thank you!  
Questions?