

Preference, Consensus, and Choice in Crowdsourced Relevance



Paul Bennett, Microsoft Research (CLUES)

Joint work with Ben Carterette, Max Chickering, Susan Dumais, Eric Horvitz, Edith Law, and Anton Mityagin.

Road Map

➔ Why Preferences?

- Learning Consensus from Preferences at Scale
- Beyond Consensus
- Discussion

bing

Web

12v car battery charger



Web Shopping More▼

RELATED SEARCHES

12 Volt Car Battery Charger
Portable Car Battery Charger
Car Battery Corrosion
Car Battery Booster
Car Battery Tester
Car Battery Guide
8 Volt Car Battery
Motorcycle Batteries USA

SEARCH HISTORY

Search more to see your history

See all

Clear all · Turn off

ALL RESULTS

1-10 of 3,980,000 results · [Advanced](#)

[12 Volt Battery Charger](#)

From \$29.99. 12v Battery Chargers for Cars & Trucks. Very Low Prices.

www.harborfreight.com/chargers

[Duralast® at AutoZone](#)

Check Store Pricing & Availability. Visit Duralast Parts Official Site

www.AutoZone.com

[battery car charger](#)

Charge Devices In Your Car W/ iGo! Order Today And Get Free Shipping!

www.igo.com/carchargers

[Car Battery Charger](#)

Best Car Battery Charger, Home Car Battery Charger, Eonon, Sunforce, Schumacher, Black & Decker, Jensen, Sunsei, Solar, Ac , 12v Car Battery Charger, Portable Car ...

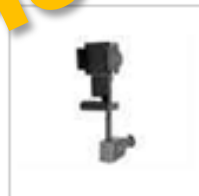
carbatterycharger.us

[12 Volt Smart Battery Charger & 12v Gel Cell & Solar Batteries](#)

12 volt battery chargers, solar, come with a money back guarantee and will not void or affect manufacturer's warranty. For cars, autos, motorcycles, and ...

www.batterystuff.com/battery-chargers/12-volt

[Shop 12v car battery charger](#)



bing.com/shopping

Brand	Category	Price
Aftermarket Gr...	Battery Chargers	below \$2
CTA Digital	Chargers & Jumpers	\$2-\$23
Macally Mace	Batteries	above \$23
Mizoo		

[12v Car Battery Charger - Compare Prices, Reviews and Buy at ...](#)

12v Car Battery Charger - 1,176 results like the Energizer 12V Alkaline Photo Cell Battery - 2 Pack, HP VV122AA Auto Adapter - 40 W For Notebook - 12 V DC, Garmin 010 ...

www.nextag.com/12v-car-battery-charger

[Use 12v Car Battery Charger To Keep Your Batteries Alive!](#)

If you drive a car and you don't want to deal with the headaches of having a dead car battery, you

Sponsored sites

Sponsored sites

[Auto Battery Chargers](#)

Find New Battery Chargers Today.

Enjoy Big Savings at JC Whitney!

www.JCWhitney.com/BatteryChargers

[Duracell® Official Site](#)

Find The Right Duracell Product For Your Electronic Needs.

www.Duracell.com

[Car Battery](#)

Save up to 40% on Car Batteries. We Compare Best Prices!

CarBatteries.ShopCompare.net

[12v Car Battery Charger](#)

Explore 100+ Batteries. Save on 12v Car Battery Charger!

Batteries.BizRate.com

[12v Charger at Amazon](#)

Find Car Parts, Accessories, Tools. Garage Equipment, Car Care & More

www.Amazon.com/automotive

[car battery at Sears®](#)

Buy Online, and Pick Up at Your Closest Sears Store. Shop Auto Now!

www.Sears.com/Auto

[Find Car Battery Chargers](#)

Your Source For Auto Parts. Find Car Battery Chargers Quickly.

AutomotiveInDepth.com

[Largest Battery Site](#)

Over 1,000,000 Batteries In Stock ready to ship today.

www.AtBatt.com

[See your message here](#)

Why relevance judgments?

- Used widely in search and advertising to:
 - Train ranking algorithms
 - Measure progress
 - Which system to deploy
 - Is the system better than it previously was?
 - Is it better than alternative models?
 - Is that difference significant?
 - Assess performance against competitors
 - Identify which components of the system need improvement

How relevant is this page when you search for:

12v car battery charger

tfdemo2 | End Session | Logout

Fair

Good

Excellent

Perfect

Bad

http://www.amazon.com/exec/obidos/redirect?link_code=ur2&...

amazon.com Your Amazon.com Electronics See all 42 Product Categories Your Account | Cart | Your Lists | Help |

Browse Brands | Top Sellers | Camera & Photo | Computers & PC Hardware | Audio & Video | Cell Phones & Service | Office Electronics | Today

Search Electronics battery charger 60 Gift Certificates

Category
Any Category
Electronics
Camera & Photo (8,941)
Accessories & Supplies (37,757)
Home Automation & Security (24)
Car Electronics (19)
Audio & Video (254)
Office Electronics (65)
Computers & Add-Ons (174)
GPS & Navigation (4)

Brand
Any Brand
Canon (115)
Rayovac (11)
Sony (63)
La Crosse Technology (1)
CTA Digital (54)
Vector (8)
Bargaincell (134)
Hitech (479)
Olympus (47)
Sakar (32)
Nikon (60)
Schumacher (34)
See more...

Seller
Any Seller
BoxWave Corporation (10,274)
Gomadic Corp. (9,068)
Amazon.com (4,886)
Battery Heads (4,570)
Batteries (4,519)
Accessory Genie (1,487)
Batteryware (802)
1800eMart (775)
K & B Electronics (587)
Eforcity (481)
ANTOnline (478)
Hitech (419)
See more...


Price


Electronics > "battery charger"


Related Searches: [aa battery charger](#), [battery charger aa](#), [battery charger aaa](#).


Showing 1 - 24 of 40,121 Results < Previous | Page: 1 2 3 ... | Next > Sort by


Over 100 Battery Chargers products in Electronics


- 

Sony BCG-34HE4 Super-Quick Worldwide Battery Charger with 4 AA NiMH Batteries
Buy new: ~~\$20.00~~ **\$19.99**
25 Used & new from \$14.80
Get it by **Thursday, Oct 11** if you order in the next **20 hours** and choose one-day shipping.
Eligible for **FREE Super Saver Shipping**.
★★★★★
- 

Rayovac Universal Battery Charger for Rechargeable AA, AAA, C, D, and 9-Volt Batteries
Buy new: ~~\$24.99~~ **\$19.27**
2 Used & new from \$18.95
Get it by **Thursday, Oct 11** if you order in the next **22 hours** and choose one-day shipping.
Eligible for **FREE Super Saver Shipping**.
★★★★★
- 

Sony Power C AA Batteries
Buy new: ~~\$10.00~~ **\$9.99**
25 Used & new from \$8.99
Get it by **Thursday, Oct 11** if you order in the next **22 hours** and choose one-day shipping.
Eligible for **FREE Super Saver Shipping**.
★★★★★
- 

La Crosse Technology BC-900 AlphaPower Battery Charger
Buy new: ~~\$60.00~~ **\$49.99**
4 Used & new from \$48.95
Get it by **Thursday, Oct 11** if you order in the next **22 hours** and choose one-day shipping.
Eligible for **FREE Super Saver Shipping**.
★★★★★
- 

Black and Decker VEC10868BD 6/4/2 AMP BATTERY CHARGER
Buy new: ~~\$90.00~~ **\$32.88**
6 Used & new from \$32.88
Get it by **Thursday, Oct 11** if you order in the next **22 hours** and choose one-day shipping.
Eligible for **FREE Super Saver Shipping**.
★★★★★
- 

Schumacher Speedcharge Charger
Buy new: ~~\$49.99~~ **\$49.99**
4 Used & new from \$49.99
In Stock
★★★★★

From Judgments to Performance

Rank for two systems

1	A	D
2	F	B
3	B	E
4	H	H
5	D	A
6	E	F
7	C	G
8	G	C

Get five-point rel.

1	5	4
2	4	5
3	5	3
4	2	2
5	4	5
6	3	4
7	1	1
8	1	1

Convert to "Gain"

1	31	15
2	15	31
3	31	7
4	3	3
5	15	31
6	7	15
7	0	0
8	0	0

Weight by rank

1.00	31	15
0.63	15	31
0.50	31	7
0.43	3	3
0.39	15	31
0.36	7	15
0.33	0	0
0.32	0	0

Multiply and Add to get DCG
System 1: 65.55
System 2: 56.69

Divide by ideal (68.30) to get NDCG/100
System 1: 96.0
System 2: 83.0

What's wrong with absolute?

- High Variability in Judged Relevance of Page
 - Can lead to incorrect conclusions of system improvement (and deployment of wrong version).
 - Can cause incorrect assessment of performance relative to competitors.
 - Introduces noise in training data for ranker.
- Coarser-grained distinctions
 - May mask real discernible differences in page quality.

Which page is more relevant when you search for:

12v car battery charger

Here

Duplicate
Pages

Here

Bad <http://www.compactappliance.com/xq/JSP,itemType.CATE...>

CA COMPACT APPLIANCE
SMALL SOLUTIONS - SUPERIOR SERVICE

Call Our EXPERTS: 800-297-6076 Sales: 24 Hours A Day Service: M-F 9am-5pm CST

COLUMBUS DAY SAVINGS EVENT Save 5% OFF ALL ORDERS THIS MONDAY & TUESDAY ONLY! Plus Get Free Shipping On 100% Of Products. Click Here For Details

Clearance Incredible Prices On Refurbished Products View Complete List

Home > Garage & Utility > Power > **Battery Chargers** View All 1 | 2 Next

Sort by: **Best Sellers** | Lowest Price | Highest Price

Departments

- Heaters
- Home Air Quality
- Compact Refrigerators
- Wine Coolers
- Beer & Beverage
- Freezers
- Laundry
- Dishwashers
- Cooking
- Portable Coolers
- Patio Furniture
- Garage & Utility
- Tankless Water Heaters
- Small Home Appliances

Specialty Areas

- Refurbished Items
- Ice Makers
- 12 Volt / Travel Coolers
- Washer Dryer Combo
- Portable Dishwashers
- Portable Air Conditioner
- Trash Compactors

Shopping Services

- Satisfaction Guarantee
- Returns & Exchanges
- Contact Us
- Help Topics
- Corporate Sales
- Email Newsletter

Vector 2/4/6 Amp Continuous Charge Rate - 12 Volt Battery Charger
List Price: \$69.99
Our Price: \$49.00
Qualifies for FREE SHIPPING

Black & Decker 6/4/2 Amp Smart Battery Charger
List Price: \$69.99
Our Price: \$45.00
Qualifies for FREE SHIPPING

Vector 2/6/10 Amp Charge Rate - 12 Volt Battery Charger
List Price: \$69.99
Our Price: \$59.00
Qualifies for FREE SHIPPING

Black & Decker 2/6/10 Amp Smart Battery Charger
List Price: \$79.99
Our Price: \$69.00
Qualifies for FREE SHIPPING

Black & Decker 2/10/15 Amp Smart Battery Charger/ Starter
List Price: \$99.99
Our Price: \$79.00
Qualifies for FREE SHIPPING

Vector 2/10/15 Amp Charge Rate - 12 Volt Battery Charger
List Price: \$99.99
Our Price: \$79.00
Qualifies for FREE SHIPPING

Black & Decker 2/5/10/2 AMP Smart Battery Charger/ Starter
List Price: \$129.99
Our Price: \$99.00
Qualifies for FREE SHIPPING

Vector 2/10/25 Amp Charge Rate - 12 Volt Battery Charger with 75 Amp Engine Start
List Price: \$149.99
Our Price: \$99.00
Qualifies for FREE SHIPPING

Vector 2/10/30 Amp Charge Rate - 12 Volt Battery Charger with 80 Amp Engine Start
List Price: \$169.99
Our Price: \$99.00
Qualifies for FREE SHIPPING

Bad http://www.amazon.com/exec/obidos/redirect?link_code=ur2&...

amazon.com Your Amazon.com

Electronics > **battery charger**

Related Searches: **aa battery charger**, **battery charger aa**, **battery charger aaa**.

Showing 1 - 24 of 40,121 Results

Over 100 **Battery Chargers** products in Electronics

Sony BCG-34HE4 Super-Quick Worldwide Battery Charger with 4 AA NiMH Batteries
Buy new: \$29.99 \$19.99
25 Used & new from \$14.80
Get it by Thursday, Oct 11 if you order in the next 20 hours and choose one-day shipping.
Eligible for FREE Super Saver Shipping.
★★★★★

Rayovac Universal Battery Charger for Rechargeable AA, AAA, C, D, and 9-Volt Batteries
Buy new: \$24.99 \$19.27
2 Used & new from \$18.95
Get it by Thursday, Oct 11 if you order in the next 22 hours and choose one-day shipping.
Eligible for FREE Super Saver Shipping.
★★★★★

Sony Power Charger with 4 AA Batteries
Buy new: \$49.99 \$16.25
25 Used & new from \$16.25
Get it by Thursday, Oct 11 if you order in the next 22 hours and choose shipping.
Eligible for FREE Super Saver S
★★★★★

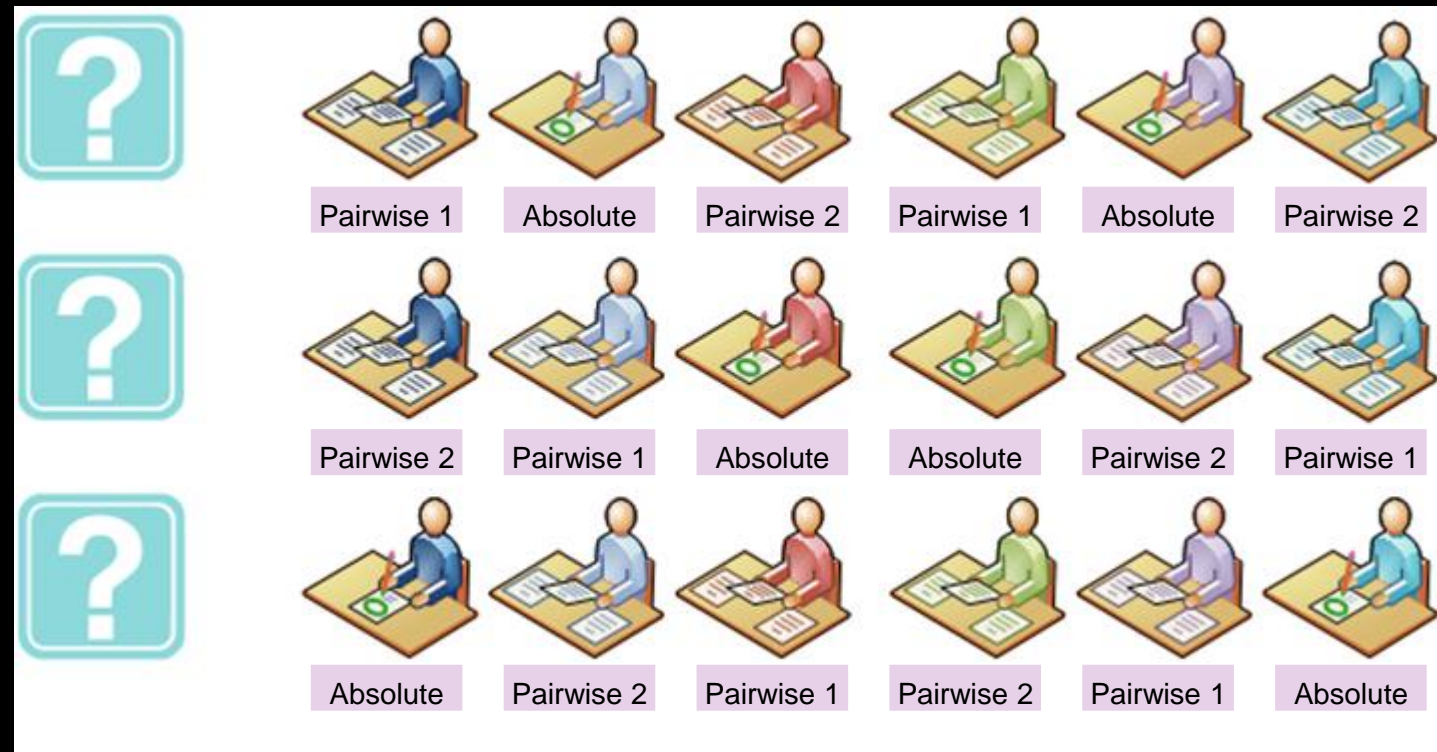
La Crosse Technology BC-900 AlphaPower Battery Charger
Buy new: \$69.99 \$49.99
4 Used & new from \$49.95
Get it by Thursday, Oct 11 if you order in the next 22 hours and choose one-day shipping.
Eligible for FREE Super Saver Shipping.
★★★★★

Black and Decker VEC1086BBD 6/4/2 AMP BATTERY CHARGER
Buy new: \$99.99 \$32.88
6 Used & new from \$32.88
Get it by Thursday, Oct 11 if you order in the next 22 hours and choose one-day shipping.
Eligible for FREE Super Saver Shipping.

Schumacher SC-6500A Speedcharge Automatic B
Buy new: \$149.99 \$90.43
4 Used & new from \$79.99
In Stock
★★★★★

Comparing Judgment Types

- Judges use absolute vs. relative interfaces
- Two problem domains
 - Search
 - Ads
- Measured
 - Agreement
 - Time
 - Also, studied advanced techniques to reduce number of judgments, calibrate to actual scale, etc.



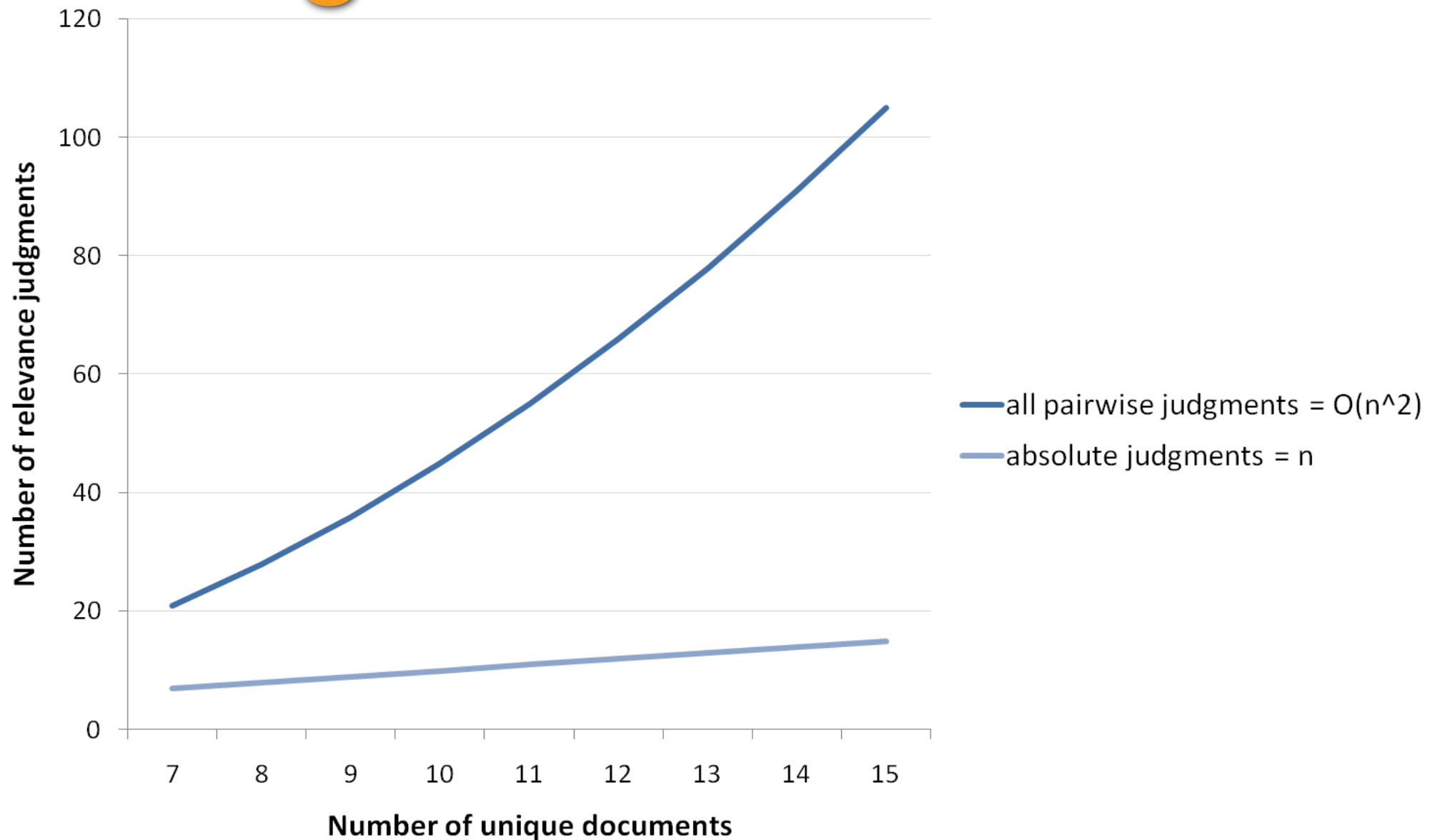
Summary Findings

Search Inferred Preferences from Absolute					Search Relative Preferences				
	A < B	A,B Bad	A > B	Total		A < B	A,B Bad	A > B	Total
A < B	0.657					0.752	0.033	0.215	2580
A,B Bad	0.297	0.380					0.567	0.225	413
A > B	0.278	0.053	0.669					0.765	2757

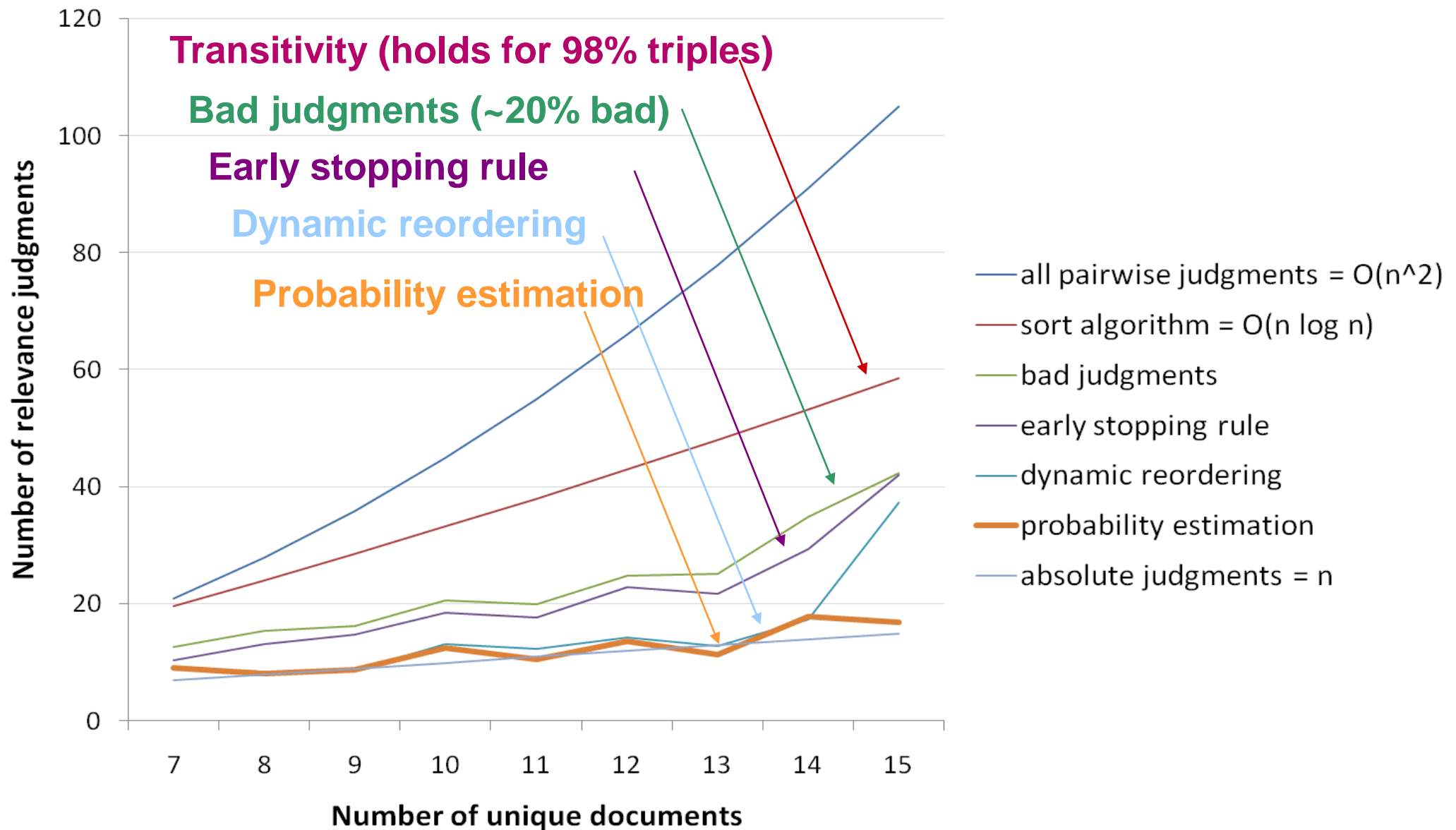
Ad Inferred Preferences from Absolute					Ad Relative Preferences				
	A < B	A,B Bad	A > B	Total		A < B	A,B Bad or Dup	A > B	Total
A < B	0.635					0.700	0.022	0.277	2691
A,B Bad	0.263	0.436					0.500	0.279	272
A > B	0.377	0.009	0.614					0.724	2985

- Relative Preferences have higher interjudge agreement than absolute.
- Faster per judgment (Two to Three times quicker)
- Other Observations
 - Finer-grained
 - Judges like it better!
- So what are the problems?

The Big O Problem

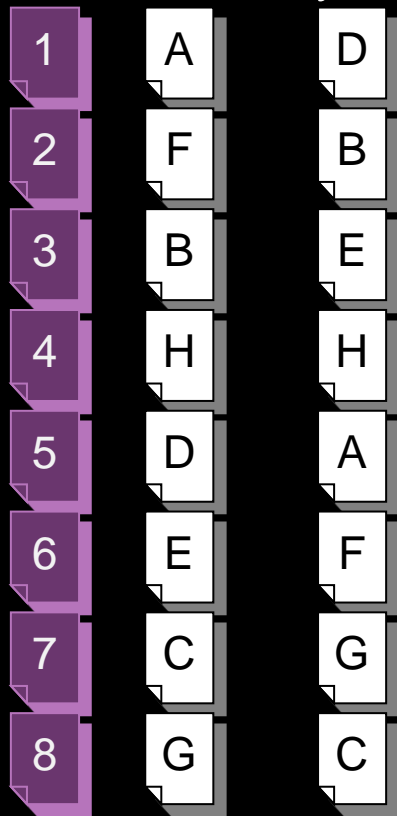


The Road from n^2 to n



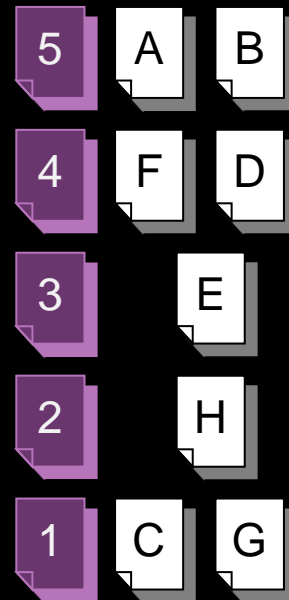
Computing Average Preference Precision

Rank for two systems



Out-Of-Order Count
System 1: 0
System 2: 0

Get Partial Order

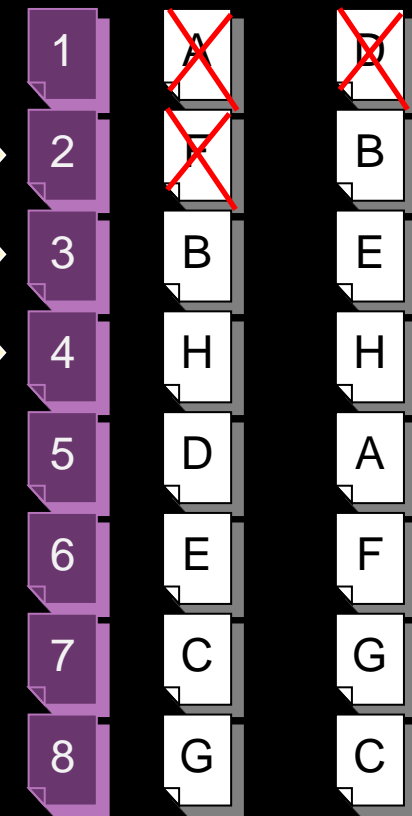


$1 - \text{Out-Of-Order} / [n(n-1)/2]$
System 1: 0.857
System 2: 0.714

Or Full Order



Count Out-of-Order Pairs Above Each Rank



Summarizing Preferences for Relevance

- Absolute judgments are noisy and effect system evaluation.
 - Relative judgments have much higher agreement.
 - Relative judgments are faster per judgment.
 - Can reduce total number of relative judgments needed.
 - Relative judgments situate the assessor in a higher degree of context.
- Both learning algorithms and evaluation measures for preference-based judgments are available.

Road Map

- Why Preferences?

➔ Learning Consensus from Preferences at Scale

- Beyond Consensus
- Discussion

Ranking for Search

- Start with a dataset of ground truth
 - An editorial judgment of relevance?
 - A click?
- Apply your favorite ranking learning algorithm
- Problem solved!

Editorial Judgments as Truth

- Pros
 - Control full process
 - Can (somewhat) calibrate judges to a consensus standard
- Cons
 - Ownership of query – what is the user's need?
 - Is relevance *topical* only?
 - Does the quality, authoritativeness, readability matter?
 - How about focus, composition, or artistry for images?

Clicks as Truth

- Pros
 - Whatever properties are rolled into relevance for the user, it all gets wrapped into that click.
- Cons
 - Positional Biases
 - What does “no click” mean?
 - An item that isn’t displayed can’t be clicked.

Consensus Opinion

- Desired as ground truth
 - A ranking that is as close to consensus opinion as possible.
- Get rankings from a large number of users
 - Sample the query stream
 - Have a large number of users rank items for many different queries.

Picture This

Live Search Club: Games: All Games: Picture This - Windows Internet Explorer

http://club.live.com/Pages/Games/GameList.aspx?game=Picture_This&tab=All%20Games


File Edit View Favorites Tools Help

Live Search Club: Games: All Games: Picture This

Live Search Set Live Search as Default Bookmark This Page Help

Microsoft®
Live Search Club Sign In

Home Games Prizes My Account Ways to Earn Community About



Play Now

Match search wits with another player to pick the right search images and win.

Good teamwork wins. Searches return certain images. Select the same picture your teammate does and score points. The two-minute game flies by, so choose quickly. Hint: earn more time when you both flag the same image as bad.

[Rules](#) [Add to My Games](#)

Every Game is Powered by **Microsoft Live Search**

Join Now

Sign In To Get Prizes

350 Tickets 1100 Tickets

Most Popular Games

Flexicon Chicktionary

EARN TICKETS WITH EVERY SEARCH

Top Games **My Games** **All Games** **Genre**

Chicktionary

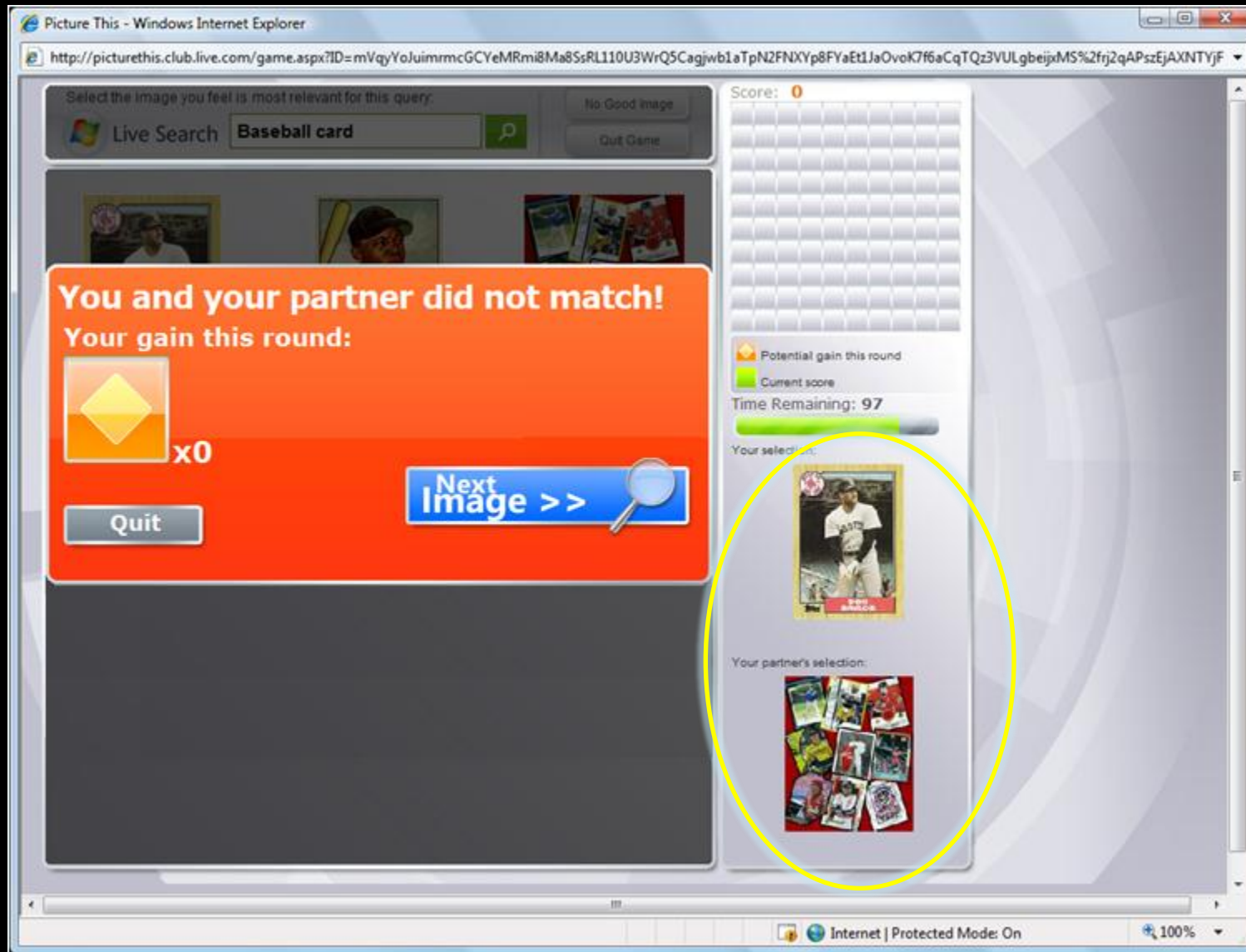
Can you defeat the wordsmith hens and scratch out every word before flying the coop? Channel your inner dictionary and extract all the words from the jumble. Reshuffle letters for a fresh look.

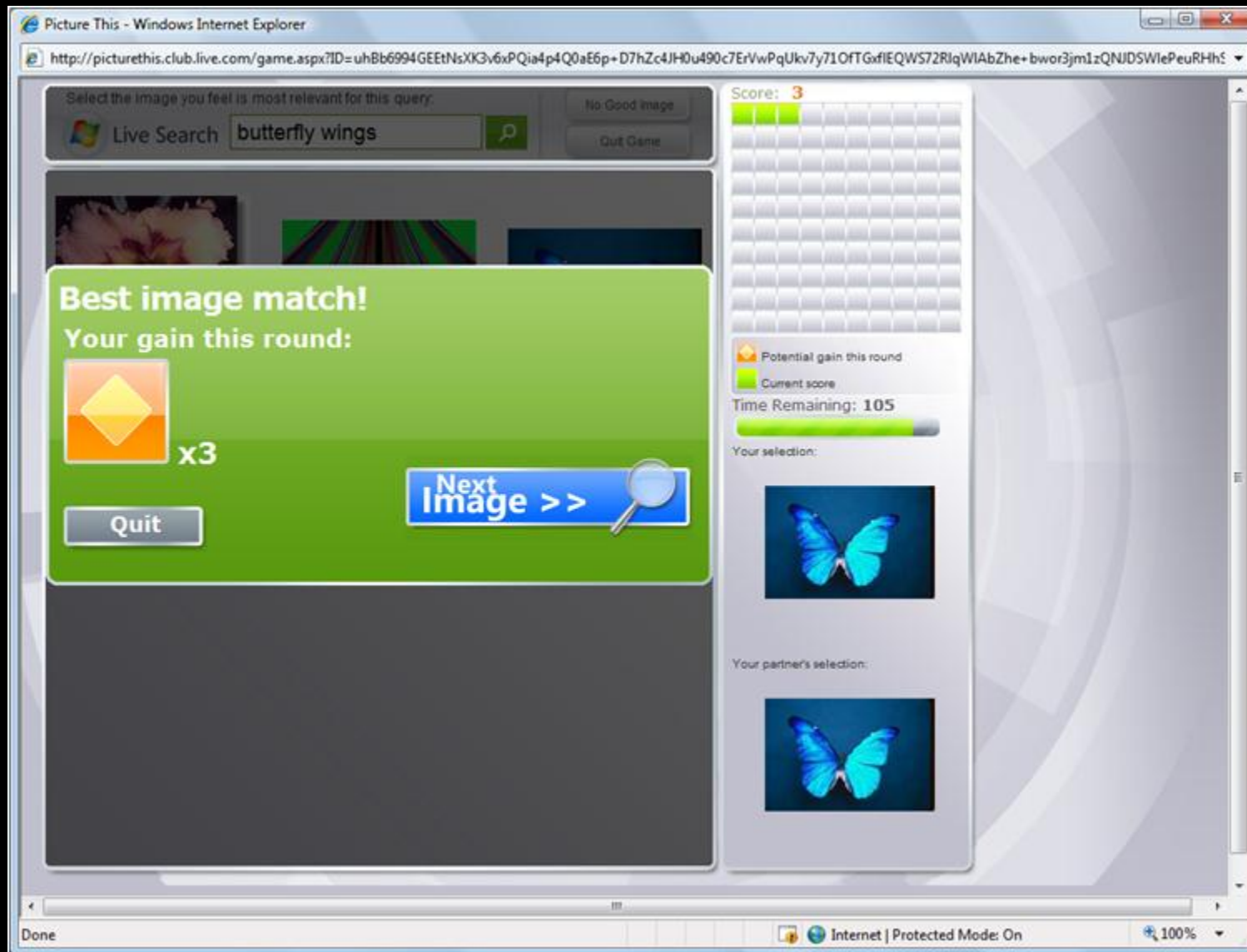
PLAY NOW [Add to My Games](#)

Clink

Done Internet | Protected Mode: On 100%







A Different Kind of Click

- Positional Biases
 - Order of choices is randomized.
- What does “no click” mean?
 - Clicks are required to play the game.
 - User can choose “No Good Image” or flag an image as “Bad” (irrelevant or detrimental)
- An item that isn’t displayed can’t be clicked.
 - Can experiment with potentially irrelevant items without risking reputation of search.

Adaptive Game Elements

- Number of images to choose from varies from 2 to 10 (based on performance).
- Comparison set varies from “easy” to “hard”
 - Easy: Images estimated relevance spread across spectrum.
 - Hard: Images close to the same level of relevance.
- Adaptivity
 - Take advantage of more “discerning” partners by increasing game difficulty.
 - Make the game more entertaining.

Partner Robots

- Sometimes no human is available to partner with (e.g. odd man out, low system usage)
- Robot partner is used to play the game.
- Robot chooses non-deterministically using basic model of relevance from preferences collected so far.

Fraud Mitigation

- Discussion boards dedicated to how to cheat on Club Bing.
- Random pairing of partners
- No fixed strategy
 - randomization of choices, no score for “no good image”
- Weeding out Bots
 - Players occasionally challenged with a CAPTCHA
 - Game site uses bot detection algorithm
 - Honeypot queries (queries with a highly agreed upon right answer)

Are Responses Random?

- Users permanently routed by ID to a “training” or “testing” server
- Training Set:
 - 49% agreement on human-human rounds.
 - 3.53 actions available on average.
 - Random clicking would yield 28% agreement.
- Testing Set:
 - 50% agreement on human-human rounds.
 - Random clicking would yield 28% agreement.

Game Data

- Simple Experiment
 - Split users into train and test.
 - Use preferences of “train” users to predict preferences of “test” users *when partners agree*.
 - Default hypothesis is user preferences are explained by Image Search’s ranking
- Goal is to predict preferences where two users agree in the test set.
- Agreement only and Raw (Full) used in training.
- 34 days, 427 queries, 95 images/query, 18M (effective) pairwise preferences.

Preprocessing

- “Raw” form
 - Our robots preferences removed.
- “Agreement” form
 - Removed all preferences where partners disagreed.
 - Removed all preferences where human was paired with a robot.
 - Kept only one preference for every pair of preference agreements.

Data Overview

- 34 days, 427 queries
- 95 images per query (94 + “neutral”)
 - More than 50 because of churn

	Games	Rounds	Human-Human Rounds
Training	154,060	1,491,206	1,144,409
Testing	155,322	1,522,375	1,159,570

	Preferences	Effective Pairwise Preferences
Raw Training	2,599,531	8,860,418
Raw Testing	2,645,984	9,267,890

	Preferences	Effective Pairwise	No Good Image
Agreement Training	537,651	1,731,317	1,800
Agreement Testing	612,736	1,838,987	1,447

Consensus Ranking

- Goal Reminder: Learn a ground truth ranking.
- Input: Features are “Was image ID X displayed?”
- Label: which ID was chosen as best.
- Model: Gives us consensus ranking.
 - Intuition: a good model predicts what people think is best.

Preference Models

- Progression from most naïve (simplest) to fewest assumptions (most complex)
 - Frequency Model
 - Global win probability of an image
 - Pairwise Probability Model
 - Models pairwise interactions but not comparison set
 - Go Model
 - Interactions in comparison set by conditioning on current set and response when learning and predicting.

Data Representation

- Introduce virtual “Neutral Image” to handle
 - “No Good Image”: Preference for neutral image over every image displayed
 - “Flag as Bad”: Preference for neutral image over image flagged.
- Pairwise preferences:
 - Image *A* preferred to *B*, *C*, and *D* can be represented as three pairwise preferences: $A > B$, $A > C$, $A > D$.

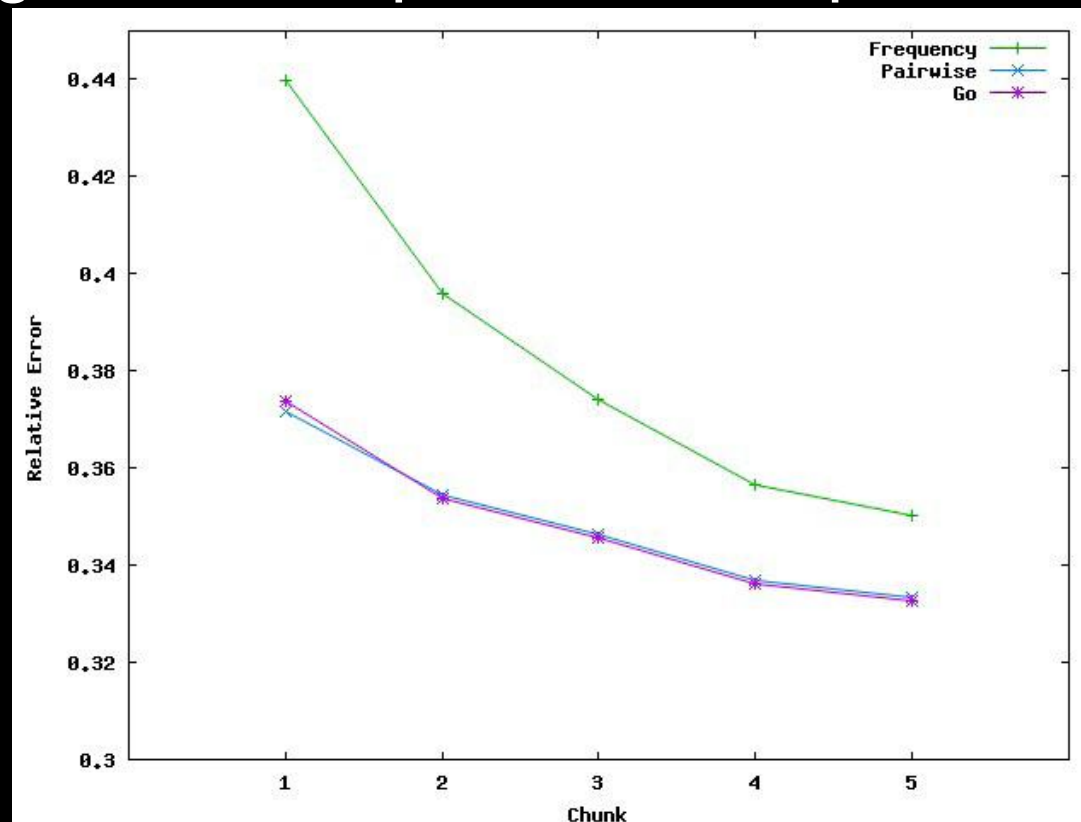
Error Relative to Image Search

	Agreement Training	Raw Training
Frequency	0.3713	0.3504
Pairwise	0.3451	0.3335
Go	0.3408	0.3325

- All methods perform at least twice and up to three times better than search baseline.
- Given a large enough training sample, simple method works well.
- More structured, complex models can use the (potentially) noisier raw training versions better.

Learning Curve

- Break raw training data into 5 ~7-day chunks
- Each training set is a superset of the previous.



Learning Curve Lessons

- Frequency model performs increasingly worse with less data
 - Higher sparsity relative to number of pairs.
- Worth using complex models if we want to get the minimal number of preferences per query and cover more queries.

Versus Editorial Judgments

- Kendall's Tau:
(agreements – disagreements) / num_differing_pairs
- Use set of already collected editorial judgments for 60 of 427 queries.
- Examined correlation of Pairwise model (trained on Agree)

	Kendall's Tau	Differing Pairs	Agreements	Disagreements	Raw Pairs
Pairwise	0.6742	27299	22852	4447	67638

Related Work

- Social Labeling: ESP Game (von Ahn & Dabbish, 2004), Peekaboom (von Ahn *et al.*, 2006), TagATune (Law *et al.*, 2007)
 - ESP does not tell you about relevance and tends to least common denominator descriptions.
- Preference Learning
 - Similar in spirit to Joachims (2002) with different source of data. No experimenting with engine itself as in Radlinski & Joachims (2005) or Radlinski *et al.* (2008)

Our Contributions

- Versus other social labeling
 - We use queries drawn from search logs → language distribution seen in practice drives labeling.
- First study of large-scale preference data as a signal to augment other sources of relevance information.
- Unique approach yields click-like relevance data
 - No positional biases, no risk of frustrating users when given non-relevant results, no danger to product reputation.
- Systematic study of a number of preference models
 - For plentiful data, many models are fine.
 - For sparse data, pairwise and Go both work well.
- April 2008 – Jan 2009
 - 2.5M visitors, 60M searches (~90M clicks)
 - Conceivable to collect judgments in academic setting.

Road Map

- Why Preferences?
- Learning Consensus from Preferences at Scale
- ➔ Beyond Consensus
- Discussion

The need for expertise

amazon.com Hello. Sign in to get personalized recommendations. New customer? Start here. FREE Two-Day Shipping for College Sponsored by Post-it & Scotch

Your Amazon.com Today's Deals Gifts & Wish Lists Gift Cards Your Account Help

Shop All Departments Search Grocery & Gourmet Food GO Cart Wish List

Grocery & Gourmet Food Bestsellers Breakfast Foods Coffee & Tea International Foods Snacks & Candy Gifts & Seasonal Natural & Organic Special Offers Subscribe & Save



Clam Base
by [Better Than Bouillon](#)
★★★★☆ (2 customer reviews)

Price: **\$5.95**

In Stock.
Ships from and sold by [Superior Quality Foods](#).

[See larger image](#)
[Share your own customer images](#)

\$5.95 + \$5.90 shipping
In Stock. Sold by **Superior Quality Foods**
Quantity:
[Add to Cart](#)
or
[Sign in](#) to turn on 1-Click ordering.
[Add to Wish List](#)
[Add to Shopping List](#)

Express Checkout with PayPhrase
Enter PayPhrase [What's this?](#) [Create PayPhrase](#)

More Buying Choices
Have one to sell? [Sell yours here](#)

[Share](#)   

Product Features

- No Added MSG
- Made primarily from clams
- Highly Concentrated

The Ideal Assessor

- Interest: Assessor for a query is someone who would likely *issue* a query.
- Expertise: Assessor can *distinguish* relevant, satisfying results from irrelevant ones.
- Confidence: Assessors judgments *agree* with measured utility (high signal to noise).

Choice

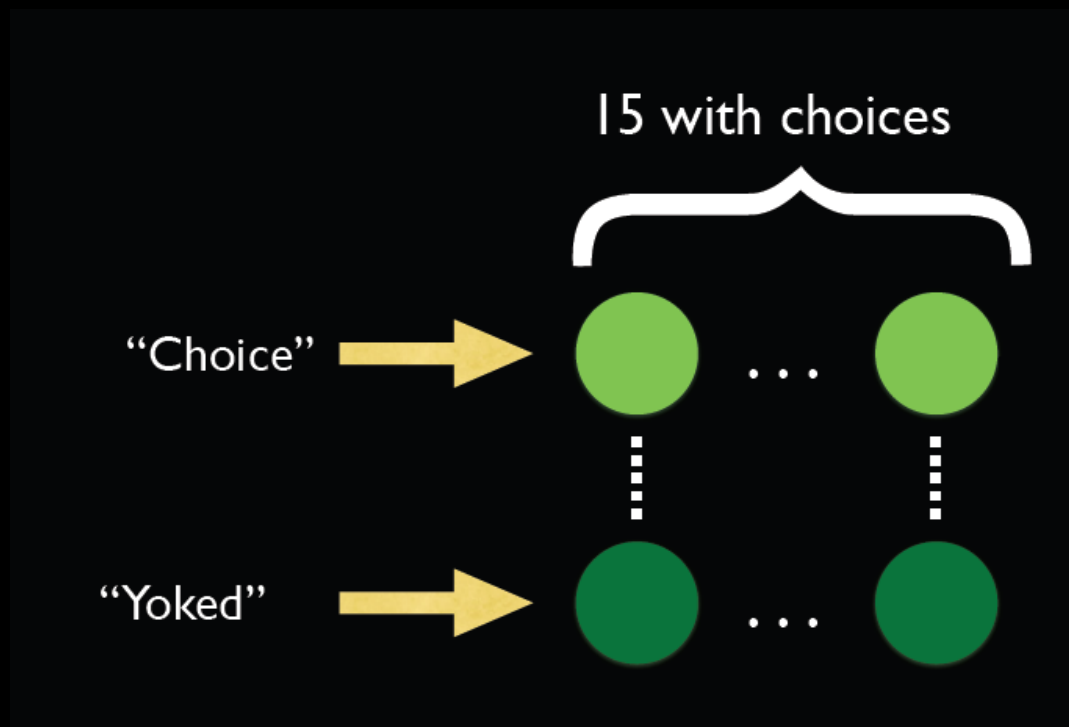
Choose a query to judge:

- ☐ wtvv television
- ☐ clam base
- ☐ cellular south
- ☐ cosa soccer
- ☐ eb games

Next

- For relevance judgments, can we leverage choice as a signal of expertise, interests and confidence?

Experiment Design



- Balance choices by length, ambiguity, and topic as a proxy for ease of choice selection, difficulty of task, and differentiability of candidates.
- For each query, pre- and post-judgment survey.

Survey

clam base

Current Search Query:

[Task 4 of 25] Step 1 of 7

Before you judge the relevance of the webpages for this search query, please answer the following questions:

(1) Specify up to 3 intents (i.e., what the user was looking for) behind the search query and mark the one that is the most likely.

Most Likely

☐

☐

☐

(2) How confident are you that the most likely intent you specified is what the user is looking for?

Very unsure Very confident

☐ ☐ ☐ ☐

1 2 3 4

(3) How knowledgeable are you in the topic of the search query:

Zero knowledge Expert

☐ ☐ ☐ ☐

1 2 3 4

(4) How interested are you in the topic of the search query:

Zero interests Very interested

☐ ☐ ☐ ☐

1 2 3 4

Next

Influence of Choice



- Intent modification: yoked subjects make non-superficial changes to the intentions more often (13.84% of all, and 5.35% of most likely intention).

Summary & Challenges

- Demonstrated
 - Simple mechanism of choice increased self-assessed interest, expertise, and confidence.
 - Also reduces the impact of exposure to the task as measured by intent modification – implying expertise.
- Challenges
 - Objective measurements of expertise beyond intent modification.
 - Given choices, predict interests/expertise and narrow set of choices around them.
 - Balance available experts with demand for tasks.

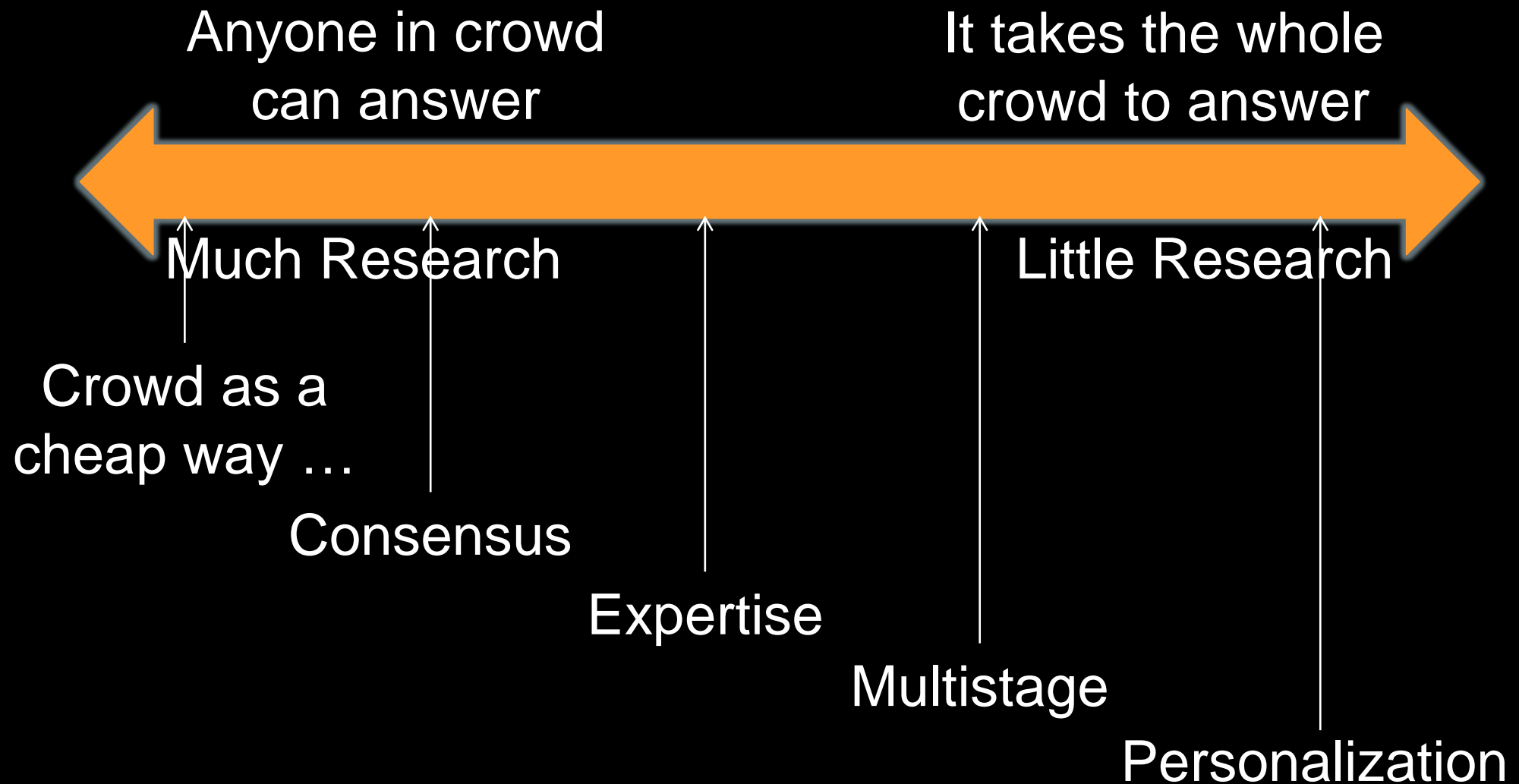
Road Map

- Why Preferences?
 - Learning Consensus from Preferences at Scale
 - Beyond Consensus
- ➔ Discussion

Specific Challenges

- Modeling player/judgment/worker performance
 - Personalized rankings, discrimination skill, identifying expertise
- Active Learning for Query & Result Selection
 - Which query, which results, and how many opinions?
- Using crowdsourcing to identify noisy *editorial* judgments or difficult to interpret click data.
- Identifying, measuring, and exploiting sources of expertise.

A Spectrum of Crowdsourcing



Microsoft®

Your potential. Our passion.™