香港浸會大學
HONG KONG BAPTIST UNIVERSITY

DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

User-Centric Design and Evaluation of Explanation for Recommendation

Dr. Li Chen

lichen@comp.hkbu.edu.hk

July 30, 2020

Invited talk for the 3rd International Workshop on Explainable Recommendation and Search (EARS 2020), in conjunction with SIGIR'20

Recommendation is almost everywhere

- What products you could buy …
- What movies you could watch…
- What music you could listen to…
- Who you could date…
- What restaurants/hotels you could visit …
- What images you could view…
- Etc.

Explanation for Recommender Systems

- *What is explanation?*
  - "*making clear by giving a detailed description*" (Tintarev and Masthoff, 2012)
- In recommender system, it has been mainly used to
  - Increase the system's transparency
    - explain the recommendation process (i.e., the logic of underlying algorithm)
  - Persuade users to try
    - justify why the recommendation might be good for a user

Nava Tintarev and Judith Masthoff. 2012. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction* 22, 4–5 (Oct. 2012), 399–439.

Different kinds of explanation

Explanation Style	We recommend *U2* because:
(I) User-based	User *Aren* with whom you share similar tastes in artists, listens to U2.
(II) Item-based	(a) People who listen to your profile item *AC/DC* also listen to U2. (b) Last.fm's data indicates that U2 is similar to *Coldplay* that is in your profile.
(III) Content	(a) U2 has similar tags as *Beatles* that is in your profile. (b) U2 is tagged with *rock* that is in your profile.
(IV) Social	Your friend *Cindy* likes U2.
(V) Item popularity	U2 is a very popular in the last.fm database with 3.5 million listeners and 94 million playcounts.

Courtesy image from Kouki et al. (2019)

Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2019. Personalized Explanations for Hybrid Recommender Systems. In Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19). ACM, New York, NY, USA, 379–390.

Collaborative (social) style

Courtesy image from Herlocker et al. (2000)

The histogram with grouping interface that **performed best** in the study of Herlocker et al. (2000).

Courtesy image from Gedikli et al. (2014)

IMDb's popular over all average rating interface that **performed worst** in the study of Herlocker et al. (2000).

Only "persuasiveness" was considered as the explanation purpose

Such explanation can cause users to overestimate item quality (Bilgic and Mooney, 2005)

Jonathan L. Herlocker, Joseph A. Konstan, and John Riedl. 2000. Explaining collaborative filtering recommendations. In *Proceedings of ACM Conference on Computer Supported Cooperative Work (CSCW'00)*. ACM, NY, 241–250.

Mustafa Bilgic and Raymond J. Mooney. 2005. Explaining recommendations: Satisfaction vs. Promotion. In *Proceedings of the Workshop Beyond Personalization, in Conjunction with IUI'05*. ACM, San Diego, California, 13–18.

Content-based explanation

Courtesy image from Vig et al. (2009)

"We recommend the movie Fargo because it is tagged with 'quirky' and you have enjoyed other movies tagged with 'quirky'"

Slot	Word	Count	Strength	Explain
DESCRIPTION	HEART	2	94.14	Explain
DESCRIPTION	BEAUTIFUL	1	17.07	Explain
DESCRIPTION	MOTHER	3	11.55	Explain
DESCRIPTION	READ	14	10.63	Explain
DESCRIPTION	STORY	16	9.12	Explain

Courtesy image from Bilgic and Mooney (2005)

Keyword style explanation

Courtesy image from Gedikli et al. (2014)

Personalized tag cloud

Jesse Vig, Shilad Sen, and John Riedl. 2009. Tagsplanations: Explaining recommendations using tags. In *Proceedings of the 14th International Conference on Intelligent User Interfaces (IUI'09)*. ACM, NY, 47–56.

Fatih Gedikli, Dietmar Jannach, and Mouzhi Ge. 2014. How should I explain? A comparison of different explanation types for recommender systems. *Int. J. Hum. Comput. Stud.* 72, 4 (April 2014), 367–382.

Explanation generation algorithm

You might be interested in [feature], on which this product performs well.
You might be interested in [feature], on which this product performs poorly.

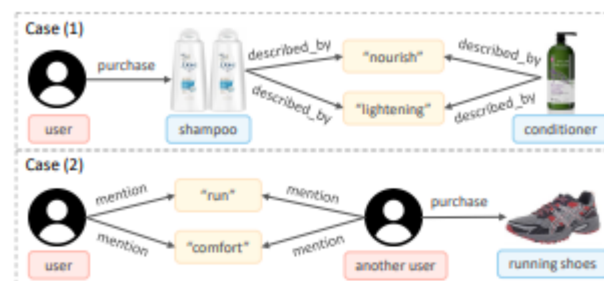
Explicit factor models for explainable recommendation (courtesy image from Zhang et al. 2014)

Yelp (user), L-Attn-only model: local attention
They carry some rare things that you can't find anywhere else. The staff is pretty damn cool too best in Arizona . I prefer ma-and-pa. They treat you the best and they value your business extreme . They are good people great atmosphere and music. I definitely believe that Lux has the best coffee I've ever had at this point. Screw all my previous reviews. This place has coffee down , they make damn good toast too .
Yelp (user), D-Attn model: local attention
They carry some rare things that you can't find anywhere else. The staff is pretty damn cool too best in Arizona. I prefer ma-and-pa. They treat you the best and they value your business extreme . They are good people great atmosphere and music. I definitely believe that Lux has the best coffee I've ever had at this point. Screw all my previous reviews. This place has coffee down, they make damn good toast too .

Explainable deep models based on attention mechanism (courtesy image from Seo et al. 2017)

Rating	Tips
4.64	***This is a great product for a great price.***
5	Great product at a great price.
4.87	***I purchased this as a replacement and it is a perfect fit and the sound is excellent.***
5	Amazing sound.
4.69	***I have been using these for a couple of months.***
4	Plenty of wire gets signals and power to my amp just fine quality wise.
4.87	***One of my favorite movies.***
5	This is a movie that is not to be missed.
4.07	***Why do people hate this film.***
4	Universal why didnt your company release this edition in 1999.
2.25	***Not as good as i expected.***
5	Jack of all trades master of none.
1.46	***What a waste of time and money.***
1	The coen brothers are two sick bastards.
4.34	***Not bad for the price.***
3	Ended up altering it to get rid of ripples.

Automatic text generation (courtesy image from Li et al., 2017)



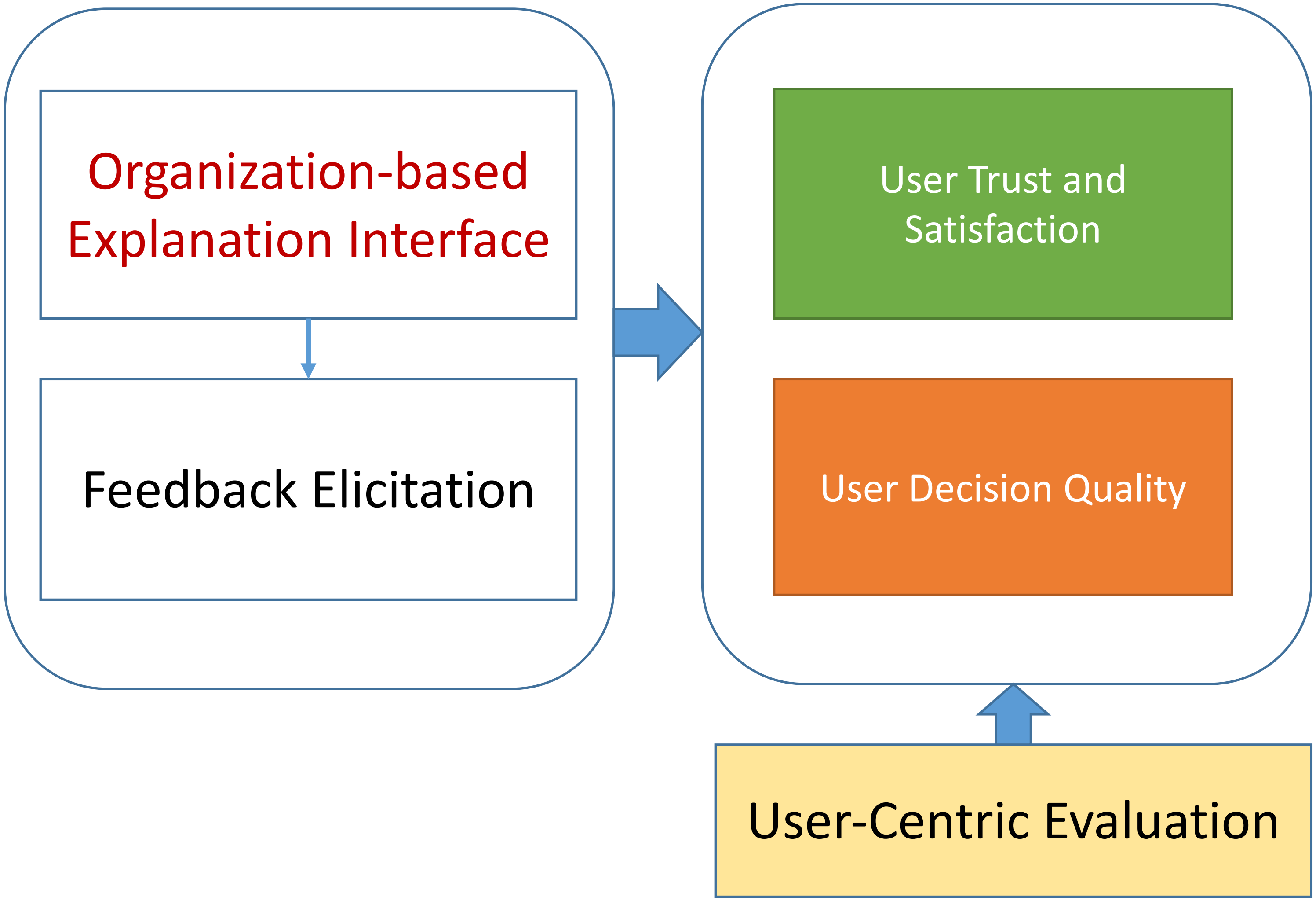
Knowledge graph reasoning (courtesy image from Xian et al., 2019)

Yongfeng Zhang and Xu Chen. Explainable Recommendation: A Survey and New Perspectives. *Foundations and Trends in Information Retrieval*: Vol. 14, No. 1, pp 1-101. Now Publishers.

- Limitations of related work
  - Specific to a single item
  - Low-risk product domains (with users' historical data)
  - Primarily emphasize on **transparency** and **persuasiveness**

- Less *from users' perspective* to design and evaluate the explanation for recommendation
  - *User trust?*
  - *User's decision quality?*
  - *Feedback elicitation from (new) users through explanation?*

Our Focus (1)

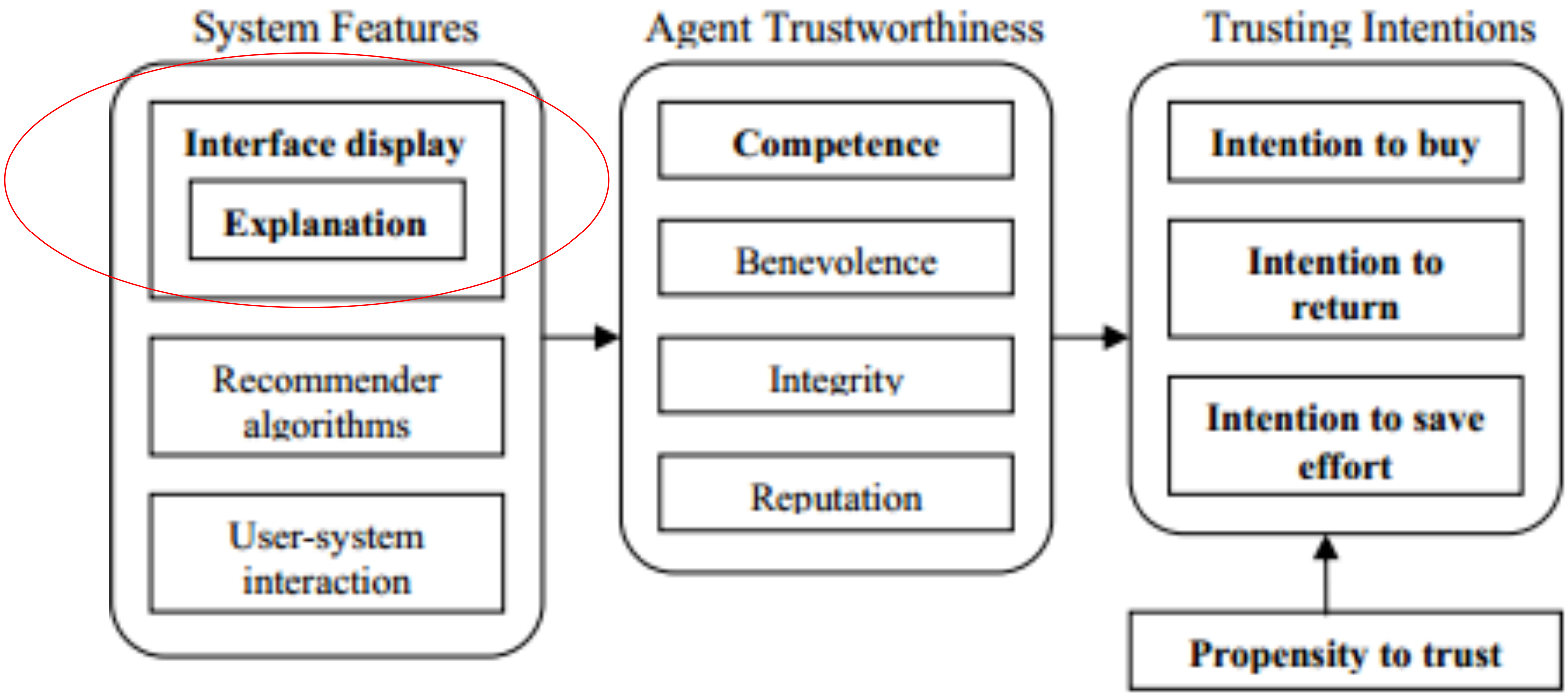

Organization-based Explanation Interface
Feedback Elicitation
User Trust and Satisfaction
User Decision Quality
User-Centric Evaluation

Motivation

Replaced by computer

- Trust is difficult to build and easy to lose in the online environment
- Low trust will stop customers from performing particular actions (e.g., transacting, purchasing, returning)
- Key factor to the success of e-commerce (Gefen, 2003)

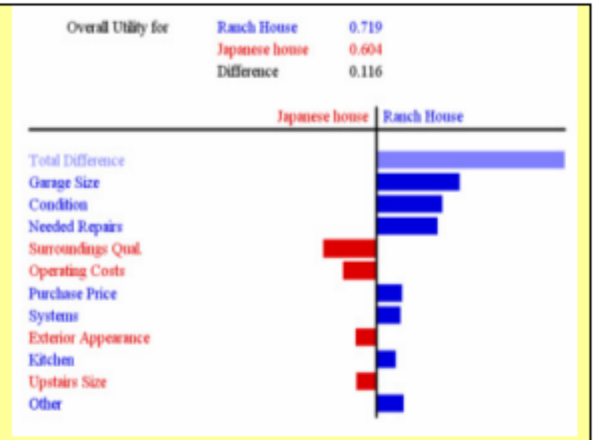
Trust Model for Recommender Systems



Li Chen and Pearl Pu. Trust Building in Recommender Agents. In *Proceedings of the Workshop on Web Personalization, Recommender Systems and Intelligent User Interfaces at the 2nd International Conference on E-Business and Telecommunication Networks (ICETE'05)*, pages 135-145, Reading, UK, October 3-7, 2005.

A user survey on 53 subjects

The ranch house seems better than Japanese house according to your preferences, since it has advantages on garage size, condition, needed repairs, purchase price, systems, kitchen and other features. However, the Japanese house still has some benefits on surroundings quality, operating costs, exterior appearance and upstairs size.



Explanation realized in text vs. graphics

House 18 is an interesting house. In fact, it has a convenient location in the Ecublens neighborhood. House 18 is close to work (1.7 miles).	House 18 is an interesting house. In fact, it has a convenient location in the Ecublens neighborhood. Even though house 18 is somewhat far from the park (3 miles), it is close to work (1.7 miles) and a rapid transportation stop (1 mile). House 18 offers a beautiful view, and it has a wonderful exterior.

Short and concise explanation sentences vs. long and detailed ones

User preference depends on product domain and background knowledge

Search Results

Ranking	ID	Type	Price	Area	Bathroom	Kitchen	Distance	
1 why?	27	room in a house	500	15	private	private	5	Basket
2 why?	30	room in a house	500	22	private	not available	10	Basket
3 why?	71	room in a house	490	18	private	not available	10	Basket
4 why?	77	shared apartment	550	20	private	not available	10	Basket
5 why?	69	shared apartment	470	15	shared	shared	10	Basket
6 why?	34	room in a house	550	25	shared	private	10	Basket
7 why?	72	room in a house	500	12	shared	private	15	Basket
								More

The recommendations with simple "why" explanation component

Search Results

There are three apartments satisfying your preferences on price, bathroom and distance

ID	Type	Price (Fs)	Area (m2)	Bathroom	Kitchen	Distance (mins)	
27	room in a house	500	15	private	private	5	Basket
30	room in a house	500	22	private	not available	10	Basket
71	room in a house	490	18	private	not available	10	Basket

Although these apartments are slightly expensive, they offer superior benefits on some of the other attributes

ID	Type	Price (Fs)	Area (m2)	Bathroom	Kitchen	Distance (mins)	
77	shared apartment	550	20	private	not available	10	Basket
34	room in a house	550	25	shared	private	10	Basket
							More

These apartments satisfy your price need, but not on all other preferences

ID	Type	Price (Fs)	Area (m2)	Bathroom	Kitchen	Distance (mins)	
69	shared apartment	470	15	shared	shared	10	Basket
72	room in a house	500	12	shared	private	15	Basket
							More

Organization-based explanation interface, where the category title replaces the "why" component

Explanation can be an effective means to inspire user trust in the recommender system;
Organization-based interface can be more effective than the simple "why" interface

Design principles

- **Principle 1**: *Categorize remaining recommendations according to their similar tradeoff properties relative to the top candidate*
- **Principle 2**: *Propose improvements and compromises in the category title using conversational language; keep the number of tradeoff attributes under five to avoid information overload*
- **Principle 3**: *Eliminate dominated categories, and diversify the categories in terms of their titles and contained recommendations*
- **Principle 4**: *Include actual products in a recommended category*
- **Principle 5**: *Rank recommendations within each category by exchange rate rather than similarity measure*

Pearl Pu and Li Chen. Trust Building with Explanation Interfaces. In *Proceedings of International Conference on Intelligent User Interfaces (IUI'06)*, pages 93-100, Sydney, Australia, January 29-February 1, 2006.

Organization interface vs. Ranked list

The most popular product

	Manufacturer	Price	Processor speed	Battery life	Installed memory	Hard drive capacity	Display size	Weight
◉	—	$2'095.00	1.67 GHz	4.5 hour(s)	512 MB	80 GB	38.6 cm	2.54 kg

We also recommend the following products because

they are cheaper and lighter, but have lower processor speed

	Manufacturer	Price	Processor speed	Battery life	Installed memory	Hard drive capacity	Display size	Weight
○	—	$1'499.00	1.5 GHz	5 hour(s)	512 MB	80 GB	33.8 cm	1.91 kg
○	—	$1'739.99	1.5 GHz	4.5 hour(s)	512 MB	80 GB	38.6 cm	2.49 kg
○	—	$1'625.99	1.5 GHz	5 hour(s)	512 MB	80 GB	30.7 cm	2.09 kg
○	—	$1'426.99	1.5 GHz	5 hour(s)	512 MB	60 GB	30.7 cm	2.09 kg
○	—	$1'929.00	1.2 GHz	4 hour(s)	512 MB	60 GB	26.9 cm	1.41 kg
○	—	$1'595.00	1 GHz	5.5 hour(s)	512 MB	40 GB	26.9 cm	1.41 kg

they have higher processor speed and bigger hard drive capacity, but are heavier

	Manufacturer	Price	Processor speed	Battery life	Installed memory	Hard drive capacity	Display size	Weight
○	—	$1'220.49	1.8 GHz	5 hour(s)	1 GB	100 GB	38.1 cm	2.95 kg
○	—	$2'148.99	2 GHz	4 hour(s)	1 GB	100 GB	39.1 cm	2.9 kg
○	—	$1'379.00	3.3 GHz	2 hour(s)	512 MB	100 GB	43.2 cm	4.31 kg
○	—	$2'235.00	1.8 GHz	2.5 hour(s)	1 GB	100 GB	43.2 cm	3.99 kg
○	—	$2'319.00	1.7 GHz	4.5 hour(s)	512 MB	100 GB	43.2 cm	3.13 kg
○	—	$2'075.00	1.8 GHz	1.67 hour(s)	512 MB	100 GB	43.2 cm	4.4 kg

they have longer battery life and lighter weight, but smaller display size

	Manufacturer	Price	Processor speed	Battery life	Installed memory	Hard drive capacity	Display size	Weight
○	—	$1'529.00	1.7 GHz	6.5 hour(s)	512 MB	80 GB	33.8 cm	1.77 kg
○	—	$1'599.00	1.7 GHz	6.5 hour(s)	512 MB	80 GB	33.8 cm	1.91 kg
○	—	$1'125.00	1.5 GHz	6 hour(s)	512 MB	80 GB	30.7 cm	2 kg
○	—	$2'099.99	1.2 GHz	9 hour(s)	512 MB	60 GB	26.9 cm	1.41 kg
○	—	$1'649.00	1.1 GHz	8.5 hour(s)	512 MB	40 GB	26.9 cm	1.36 kg
○	—	$969.00	1.2 GHz	6 hour(s)	256 MB	39 GB	30.7 cm	2.22 kg

they are cheaper, but heavier

	Manufacturer	Price	Processor speed	Battery life	Installed memory	Hard drive capacity	Display size	Weight
○	—	$1'179.00	3.2 GHz	2 hour(s)	512 MB	80 GB	39.1 cm	3.62 kg
○	—	$1'425.00	1.6 GHz	5.5 hour(s)	512 MB	80 GB	39.1 cm	2.86 kg
○	—	$1'190.00	3.2 GHz	1 hour(s)	512 MB	80 GB	39.1 cm	3.72 kg
○	—	$1'629.00	1.8 GHz	5.8 hour(s)	512 MB	60 GB	38.1 cm	2.81 kg
○	—	$627.10	1.6 GHz	1.5 hour(s)	256 MB	40 GB	38.1 cm	2.81 kg
○	—	$520.00	1.13 GHz	3.5 hour(s)	128 MB	30 GB	35.8 cm	2.59 kg

Organization interface

The most popular product

		Manufacturer	Price	Processor speed	Battery life	Installed memory	Hard drive capacity	Display size	Weight
◉		—	$2'095.00	1.67 GHz	4.5 hours	512 MB	80 GB	38.6 cm	2.54 kg

We also recommend the following products

		Manufacturer	Price	Processor speed	Battery life	Installed memory	Hard drive capacity	Display size	Weight
○	Why?	—	$1'220.49	1.8 GHz	5 hours	1 GB	100 GB	38.1 cm	2.95 kg
○	Why?	—	$2'148.99	2.0 GHz	4 hours	1 GB	100 GB	39.1 cm	2.90 kg
○	Why?	—	$1'379.00	3.3 GHz	2 hours	512 MB	100 GB	43.2 cm	4.31 kg
○	Why?	—	$1'179.00	3.2 GHz	2 hours	512 MB	80 GB	39.1 cm	3.62 kg
○	Why?	—	$1'529.00	1.7 GHz	6.5 hours	512 MB	80 GB	33.8 cm	1.77 kg
○	Why?	—	$1'599.00	1.7 GHz	6.5 hours	512 MB	80 GB	33.8 cm	1.91 kg
○	Why?	—	$1'425.00	1.6 GHz	5.5 hours	512 MB	80 GB	39.1 cm	2.86 kg
○	Why?	—	$2'235.00	1.8 GHz	2.5 hours	1 GB	100 GB	43.2 cm	3.99 kg
	This product has higher processor speed and bigger hard drive capacity but is heavier		$1'190.00	3.2 GHz	1 hours	512 MB	80 GB	39.1 cm	3.72 kg
			$1'125.00	1.5 GHz	6 hours	512 MB	80 GB	30.7 cm	2 kg
○	Why?	—	$2'319.00	1.67 GHz	4.5 hours	512 MB	100 GB	43.2 cm	3.13 kg
○	Why?	—	$1'499.00	1.5 GHz	5 hours	512 MB	80 GB	33.8 cm	1.91 kg
○	Why?	—	$1'739.99	1.5 GHz	4.5 hours	512 MB	80 GB	38.6 cm	2.49 kg
○	Why?	—	$1'629.00	1.8 GHz	5.8 hours	512 MB	60 GB	38.1 cm	2.81 kg
○	Why?	—	$1'625.99	1.5 GHz	5 hours	512 MB	80 GB	30.7 cm	2.09 kg
○	Why?	—	$1'426.99	1.5 GHz	5 hours	512 MB	60 GB	30.7 cm	2.09 kg
○	Why?	—	$2'099.99	1.2 GHz	9 hours	512 MB	60 GB	26.9 cm	1.41 kg
○	Why?	—	$2'075.00	1.8 GHz	1.67 hours	512 MB	100 GB	43.2 cm	4.4 kg
○	Why?	—	$1'649.00	1.1 GHz	8.5 hours	512 MB	40 GB	26.9 cm	1.36 kg
○	Why?	—	$627.10	1.6 GHz	1.5 hours	256 MB	40 GB	38.1 cm	2.81 kg
○	Why?	—	$969.00	1.2 GHz	6 hours	256 MB	39 GB	30.7 cm	2.22 kg
○	Why?	—	$520.00	1.13 GHz	3.5 hours	128 MB	30 GB	35.8 cm	2.59 kg
○	Why?	—	$1'929.00	1.2 GHz	4 hours	512 MB	60 GB	26.9 cm	1.41 kg
○	Why?	—	$1'595.00	1.0 GHz	5.5 hours	512 MB	40 GB	26.9 cm	1.41 kg

Ranked list

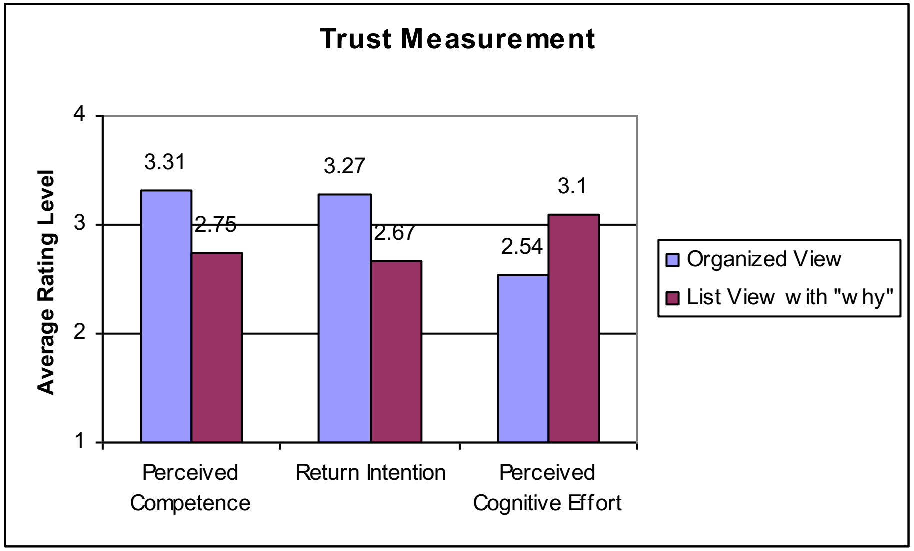
Participants: 72; Material: online product finder (digital cameras and notebooks); Procedure: within-subjects

Results

Items in the **Perceived Competence** construct	Mean	
	Organized view	List view with "why"
I felt comfortable using the interface;	3.24	2.78
This interface enabled me to compare different products very efficiently.	3.38	2.72
	Cronbach's alpha = 0.84	

Items in the **Intention to Return** construct	Mean	
	Organized view	List view with "why"
If I had to buy a product online in the future and an interface such as this was available, I would be very likely to use it;	3.11	2.56
I don't like this interface, so I would not use it again (*reverse scale*).	3.40	2.79
	Cronbach's alpha = 0.91	

Items in the **Cognitive Effort** construct	Mean	
	Organized view	List view with "why"
I easily found the information I was looking for (*reverse scale*);	2.47	3.07
Selecting a product using this interface required too much effort.	2.61	3.14
	Cronbach's alpha = 0.73	



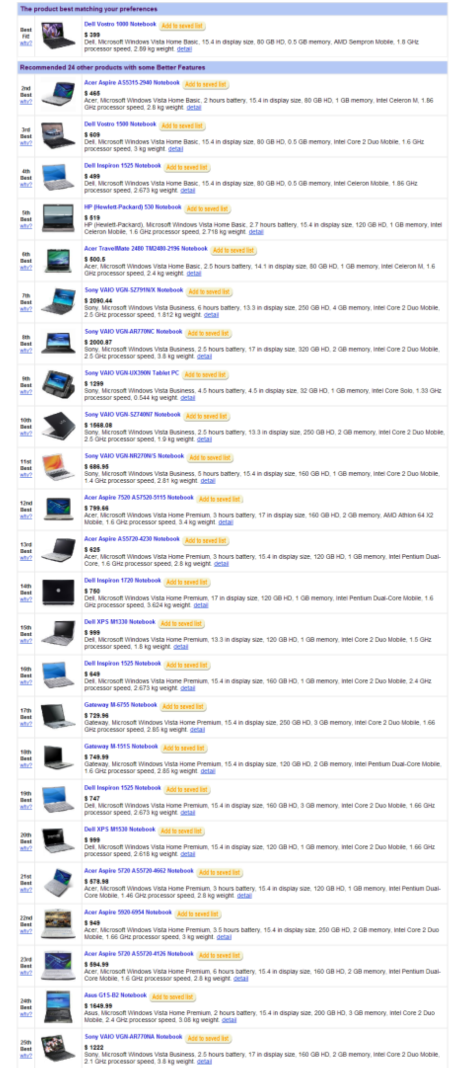
	Perceived Competence	Intention to Return	Cognitive Effort	Completion Time
Perceived Competence	1	.778** (.000)	-.826 ** (.000)	-.018 (.830)
Intention to Return	.778** (.000)	1	-.675** (.000)	-.042 (.619)
Cognitive Effort	-.826 ** (.000)	-.675** (.000)	1	.069 (.414)
Completion Time	-.018 (.830)	-.042 (.619)	.069 (.414)	1

** Correlation is significant at the 0.01 level (2-tailed).

Eye-tracking experiment (21 participants)

ORG-Vertical layout

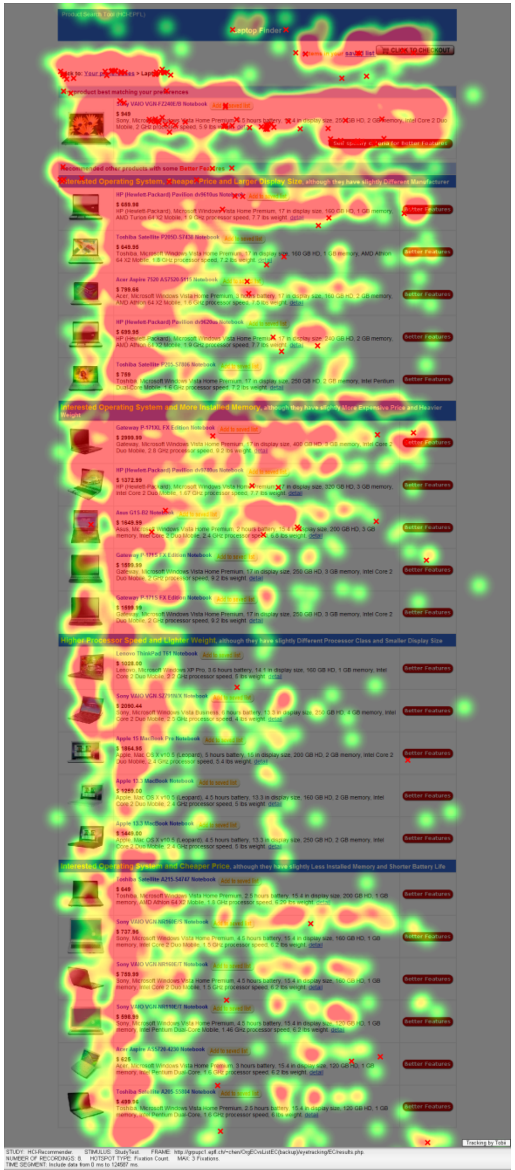
ORG-Quadrant layout



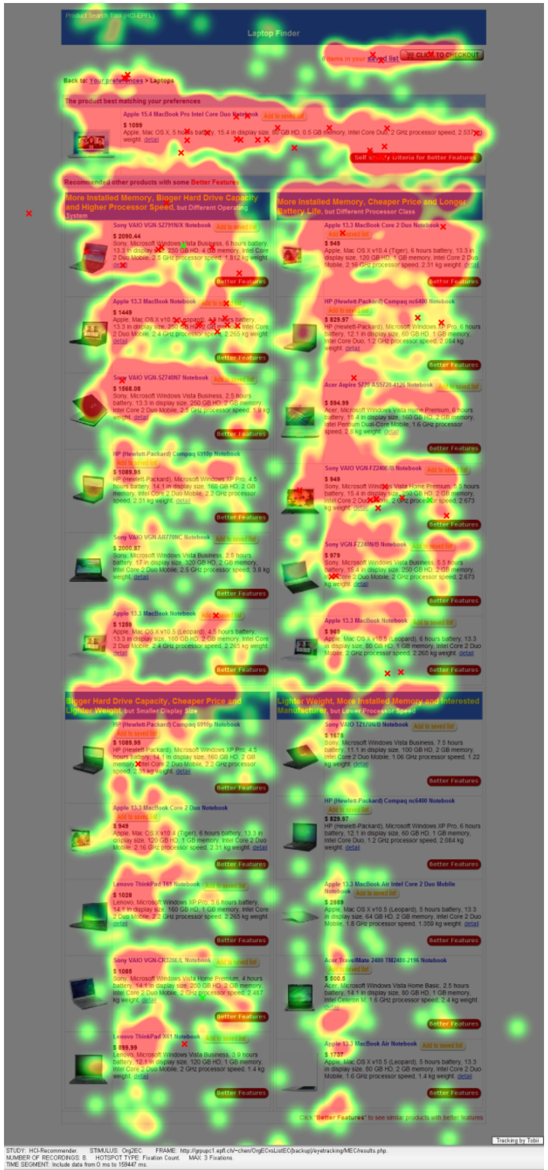
LIST

Li Chen and Pearl Pu. Users' Eye Gaze Pattern in Organization-based Recommender Interfaces. In *Proceedings of ACM International Conference on Intelligent User Interfaces (IUI'11)*, pages 311-314, Palo Alto, California, USA, February 13-16, 2011.

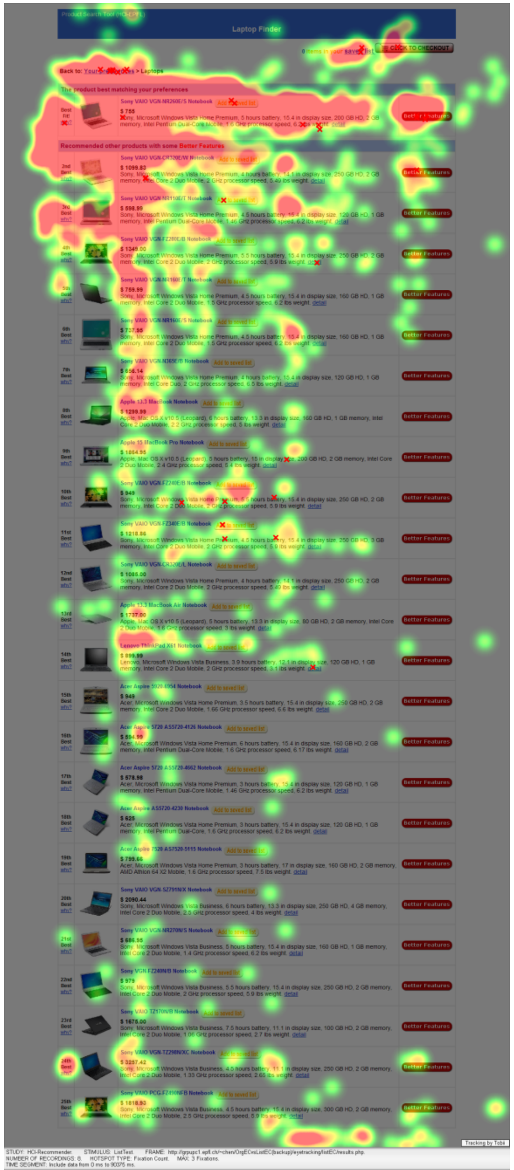
Hotspot plot



ORG-Vertical layout



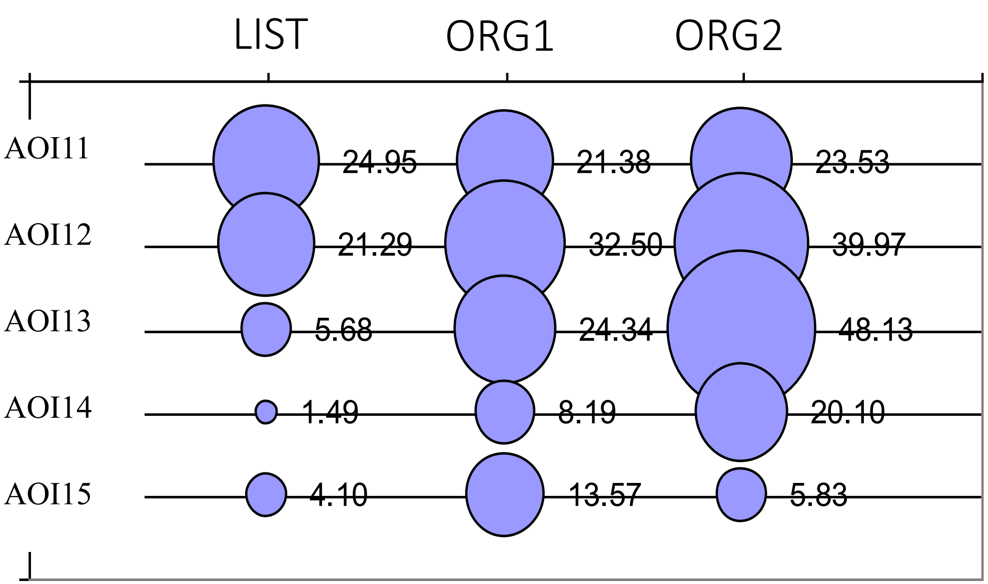
ORG-Quadrant layout



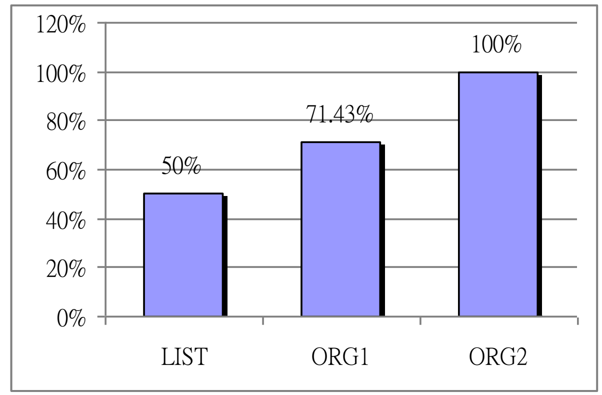
LIST

AOI analysis & user choices

Fixation duration on each AOI



Percent of users who have finally made product choice



Distribution of users' choices among AOIs

	Average selections	Top item (AOI1)	AOI2	AOI3	AOI4	AOI5
LIST	1.33	25%	75%			
ORG1	1.86	23%	31%	15%	8%	23%
ORG2	3.2	12.5%	37.5%	37.5%	12.5%	

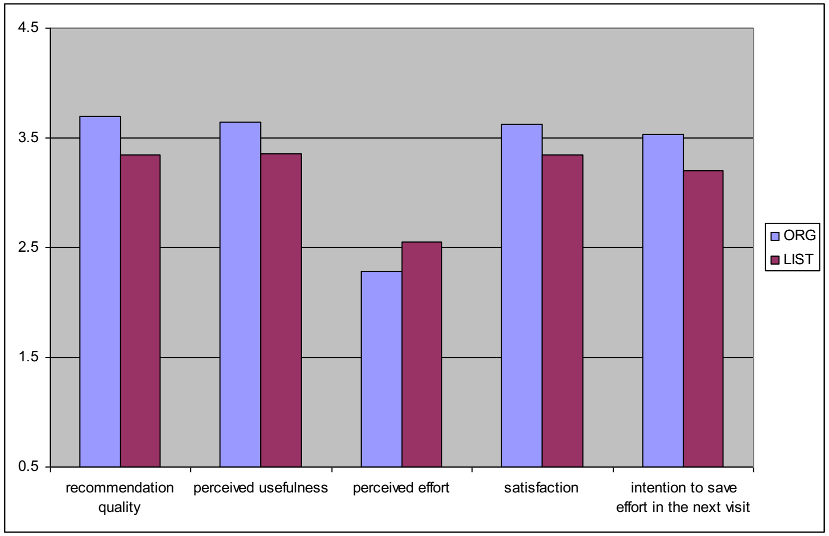
Cross-cultural user evaluation (120 participants)

	Oriental Culture (60)	*Western Culture (60)*
Nation	China (60)	Switzerland (41); Other European countries (19)
Gender	Female (23); Male (37)	Female (15); Male (45)
Average age	21~30 (57); >30 (3)	<21 (14); 21~30 (44); >30 (2)
Major/ job domain	Computer, mathematics, environment, electronics, architecture, etc.	Computer, education, mechanics, electronics,, architecture, etc.
Computer knowledge	4.34 (advanced)	4.08 (advanced)
Internet usage	4.83 (almost daily)	4.98 (almost daily)
e-commerce site visits	3.69 (1-3 times a month)	3.36 (a few times every 3 months)
e-shopping experiences	3.25 (a few times every 3 months)	2.92 (a few times every 3 months)

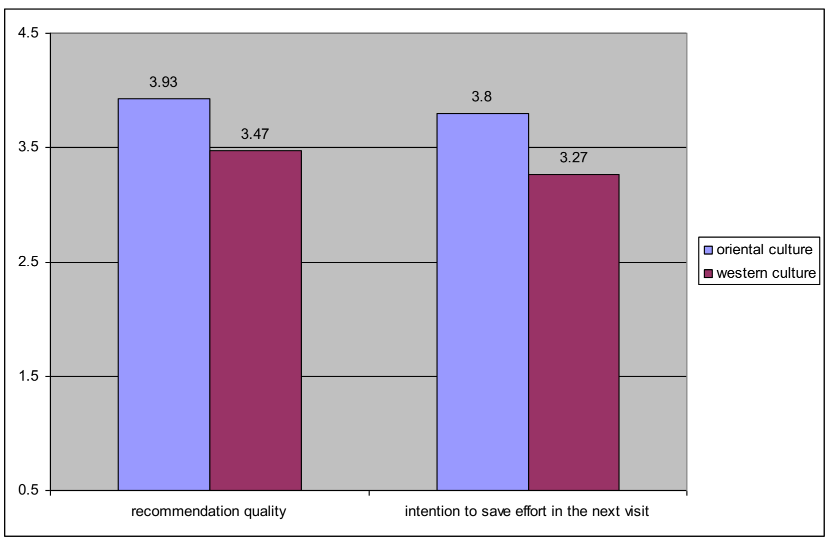
The **Oriental culture** focuses on holistic thought, continuity, and interrelationships of objects.

The **Western culture** puts greater emphasis on analytical thought, detachment, and attributes of objects.

Li Chen and Pearl Pu. A Cross-Cultural User Evaluation of Product Recommender Interfaces. In *Proceedings of ACM Conference on Recommender Systems (RecSys'08)*, pages 75-82, Lausanne, Switzerland, October 23-25, 2008.



Overall comparison



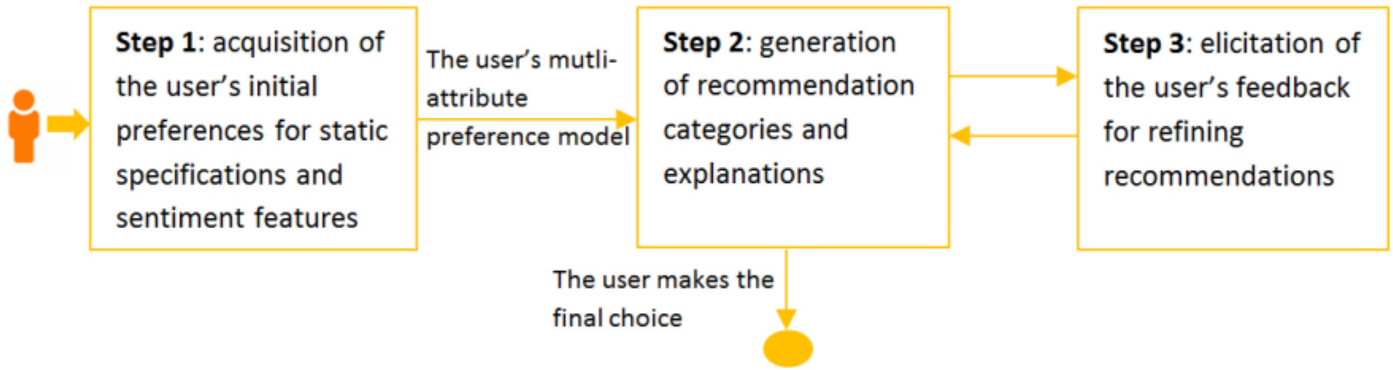
ORG interface

Major findings

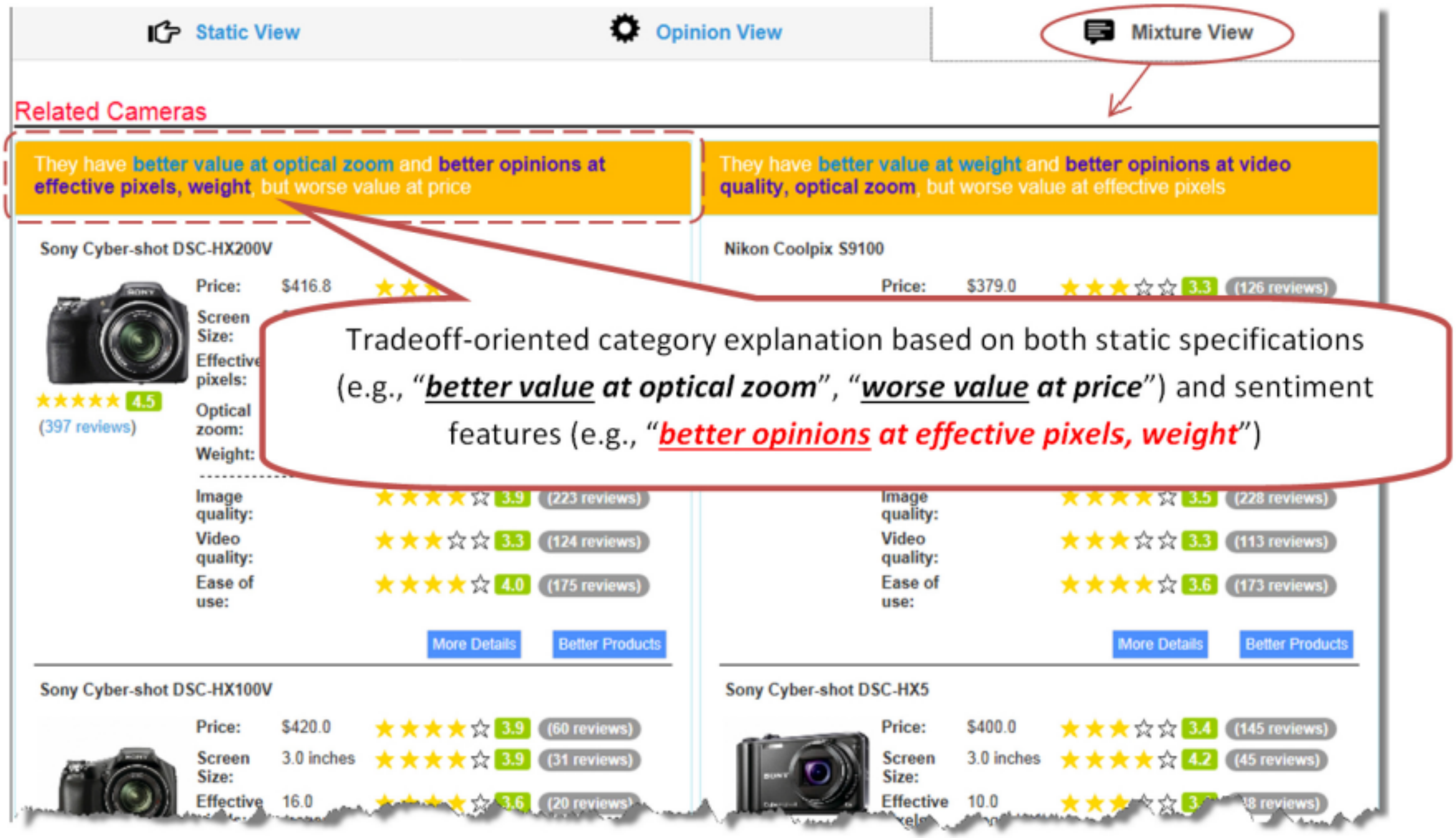
- People from different cultural backgrounds basically performed similar regarding both objective performance and subjective perceptions
- Significant favor of ORG against LIST
- Stronger among Chinese participants

Review-enhanced ORG interface

- *Motivation*: Buyers' decision certainty and purchase likelihood will be likely increased after obtaining advice from other customers' opinions (Askalidis and Malthouse, 2016)
- *Our idea*: Integrate sentiment features as extracted from product reviews to enhance the interface's explanatory power
  - to educate users about product knowledge
  - to help users to construct stable preferences
  - to enable users to make more informed and confident decisions



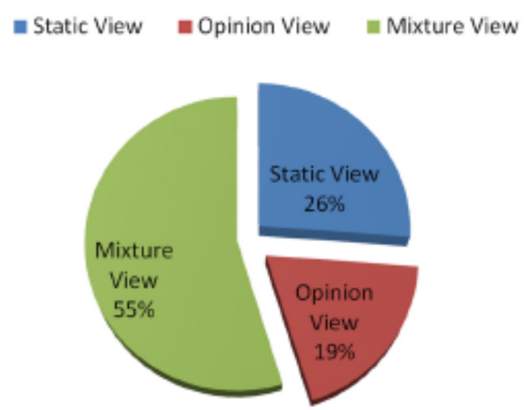
Interface design and comparison



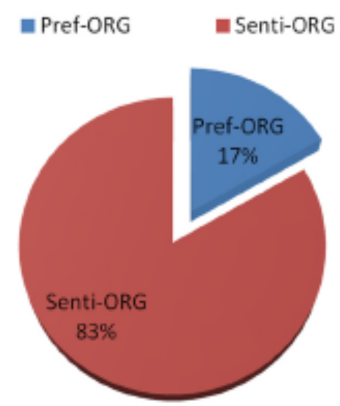
Li Chen, Dongning Yan, and Feng Wang. User Evaluations on Sentiment-based Recommendation Explanations. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 9(4), Article 20, 2019.

User studies (94 participants)

- Two within-subjects experimental setups (before-after and counter-balancing)
- Major finding:
  - Incorporation of sentiment features can significantly increase users' **product knowledge**, **preference certainty**, **perceived information usefulness**, and **purchase intention**
  - Decision efficiency is not necessarily correlated with users' decision effectiveness and system perceptions

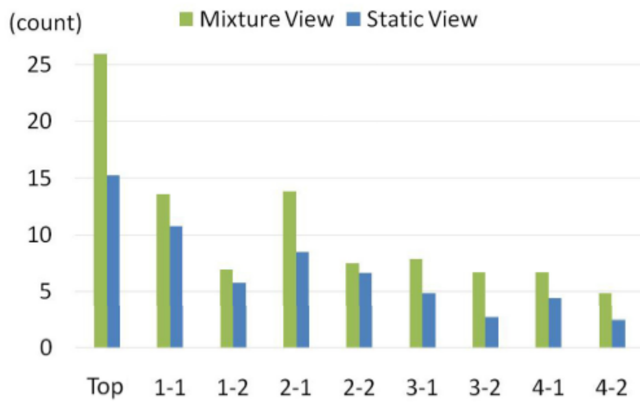


(a) Users' self-reported preferences for the three explanatory views.

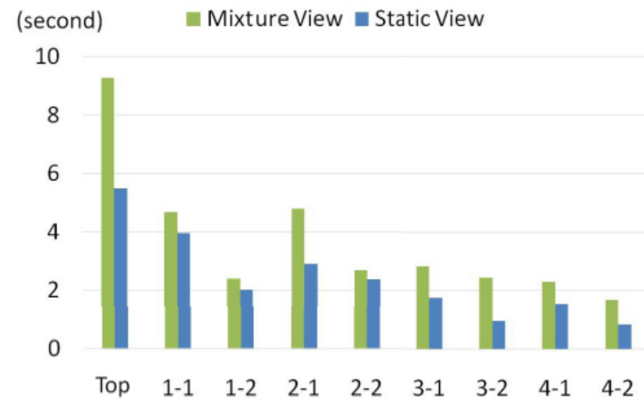


(b) Users' self-reported preferences for the two systems: Pref-ORG and Senti-ORG.

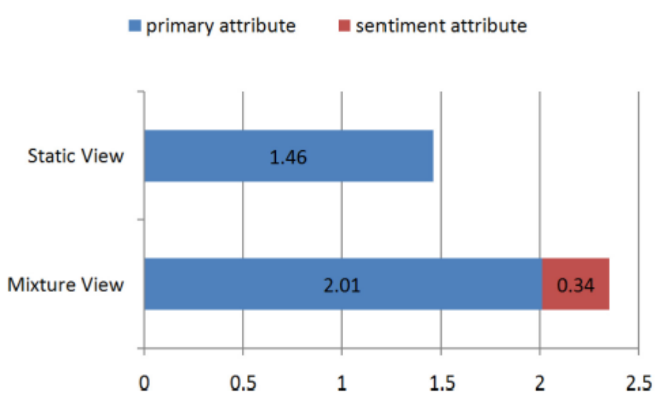
Eye-tracking study (37 participants)



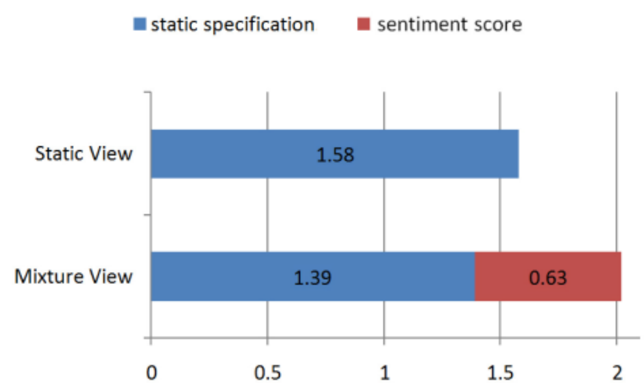
(a) Average fixation count at the individual product position.



(b) Average fixation duration at the individual product position.

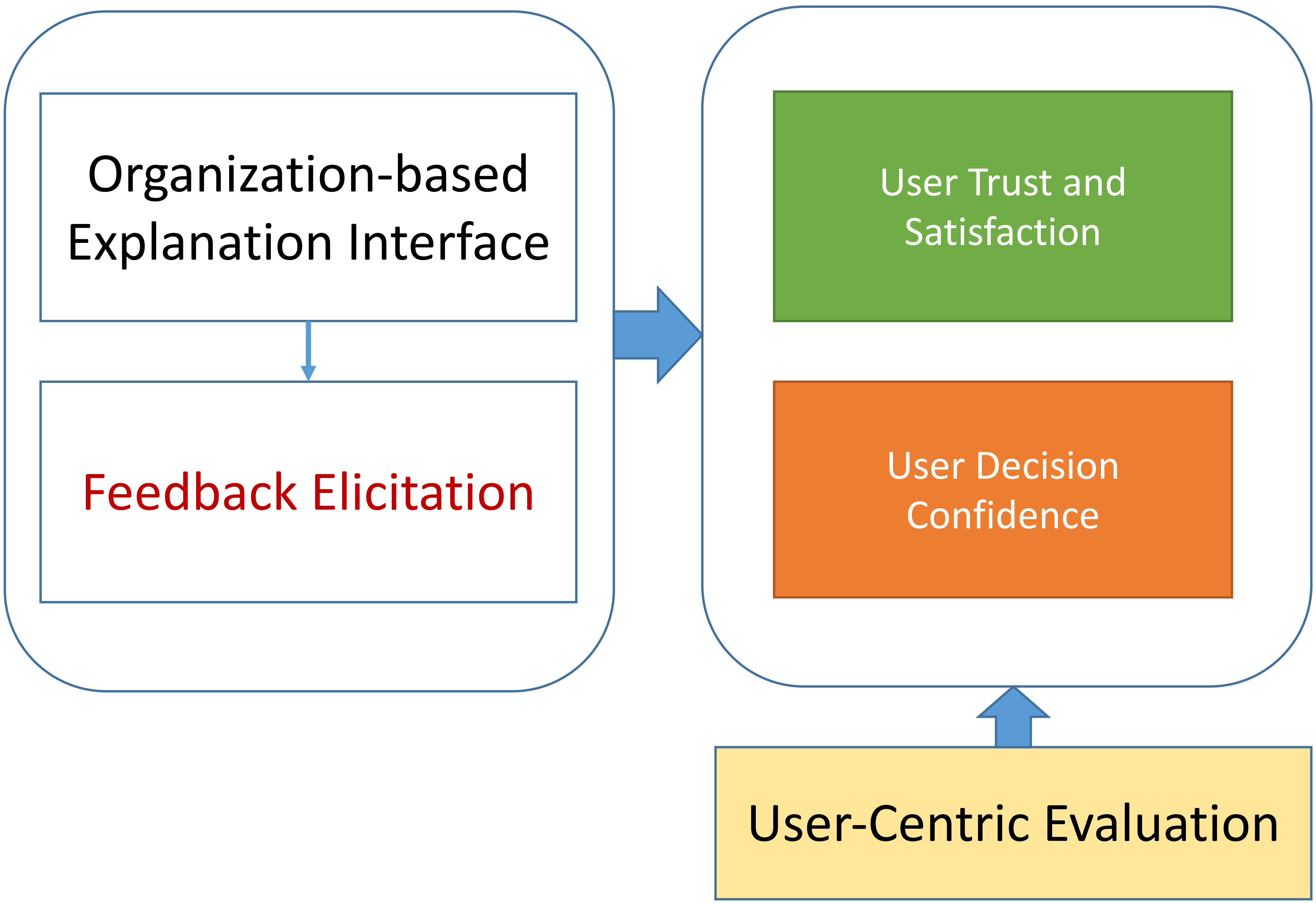


(a) The average number of attributes examined in each product evaluation.



(b) The average fixation count on static specifications (and sentiment scores in Mixture View) of primary attributes.

Our Focus (2)



Organization-based Explanation Interface
Feedback Elicitation
User Trust and Satisfaction
User Decision Confidence
User-Centric Evaluation

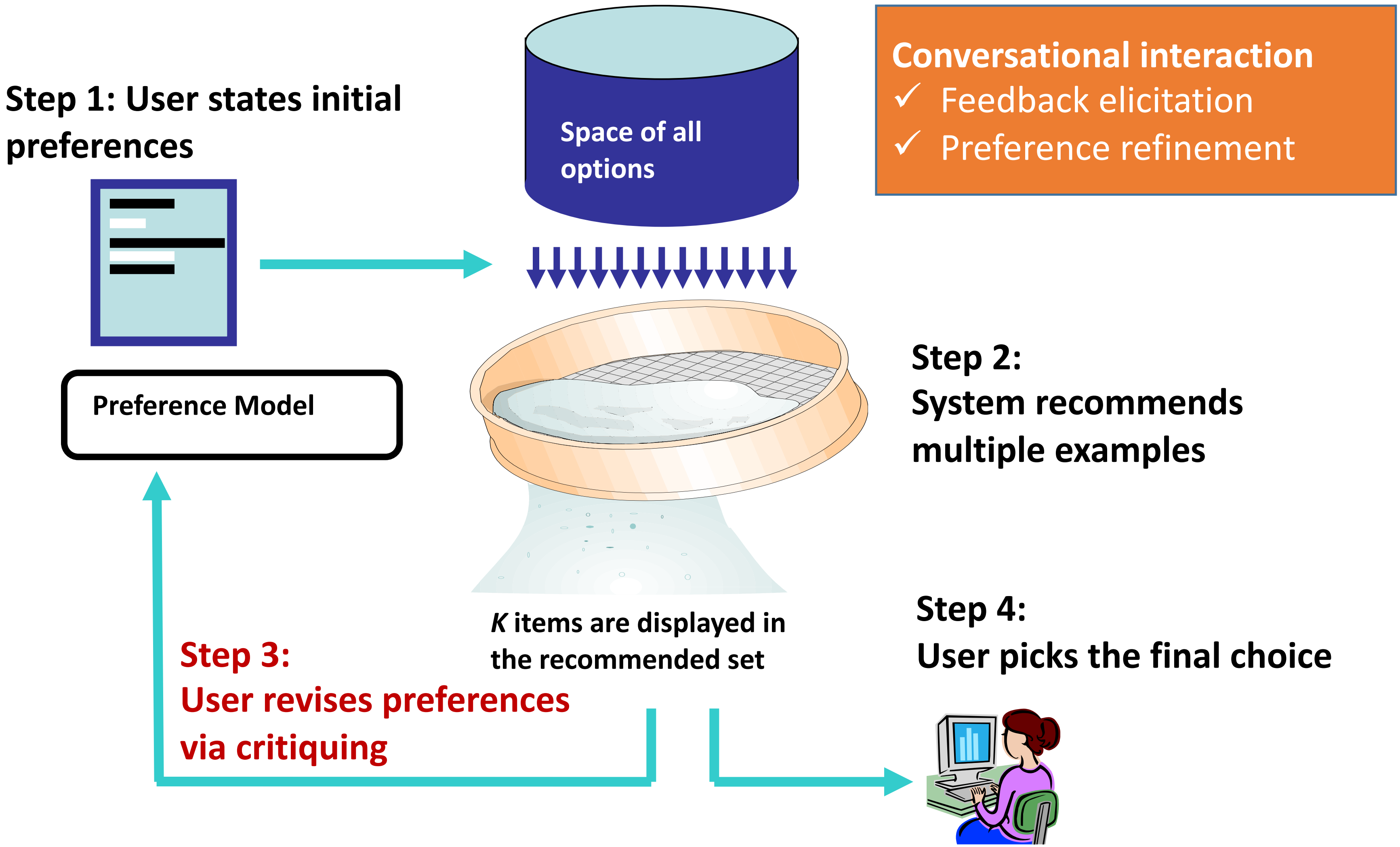
Motivation

Unfamiliar product domain

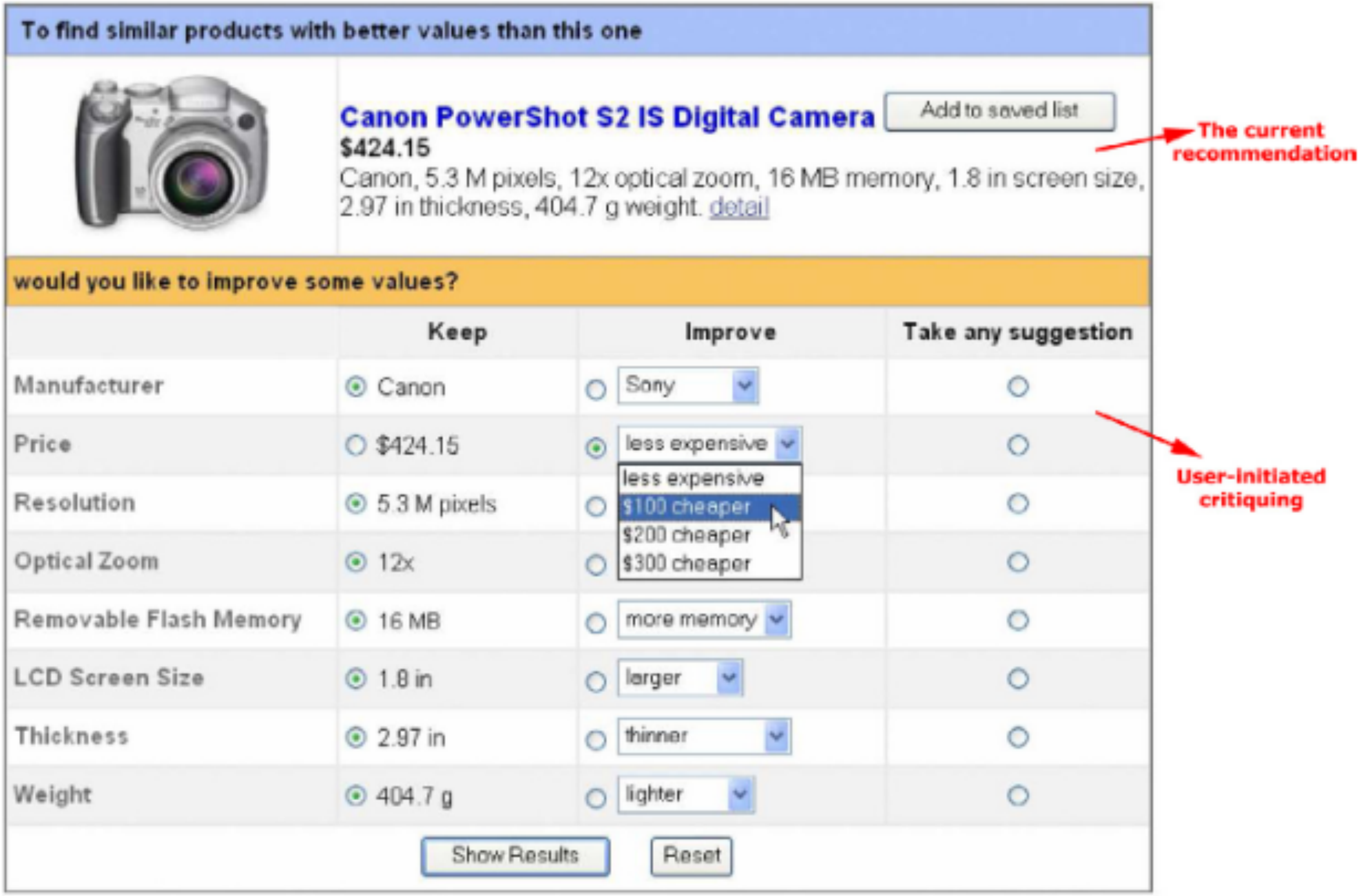
Adaptive Decision Making

- Users are likely to construct their preferences in a context-dependent and adaptive fashion during the decision process (Tversky and Simonson 1993; Payne et al. 1993, 1999; Carenini and Poole 2000).
- Users become aware of their latent preferences only when proposed solutions violate them (Pu and Faltings 2000, 2002).
- Compensatory decision strategy (i.e., tradeoff making) normally leads to rational and high-quality decision (Frisch and Clemen, 1994)

Critiquing-based Recommender Systems



Li Chen and Pearl Pu. Critiquing-based Recommenders: Survey and Emerging Trends. *User Modeling and User-Adapted Interaction Journal (UMUAI)*, vol. 22(1), pages 125-150, 2012.



Example Critiquing interface (Chen and Pu, 2006)

Li Chen and Pearl Pu. Evaluating Critiquing-based Recommender Agents. In *Proceedings of Twenty-first National Conference on Artificial Intelligence (AAAI'06)*, pages 157-162, Boston, USA, July 16-20, 2006.

Hybrid critiquing interface

System-suggested critiques

User-initiated critiquing

Li Chen, Dongning Yan, and Feng Wang. User Perception of Sentiment-Integrated Critiquing in Recommender Systems. *International Journal of Human-Computer Studies (IJHCS)*, vol. 121, pages 4-20, 2019.

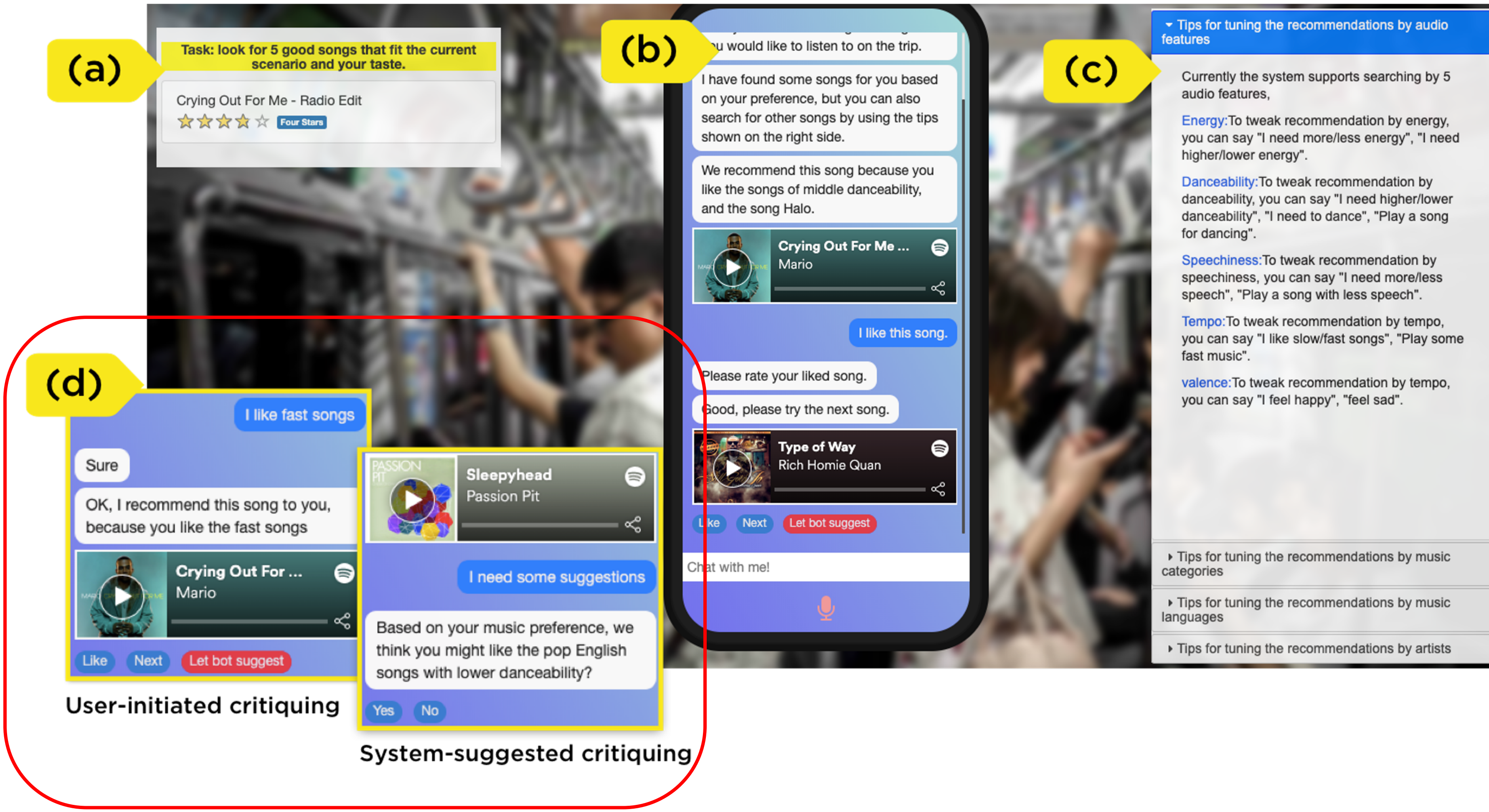
- User-initiated critiquing type
  - Similarity-based (e.g., "*Find some item similar to this one*")
  - Quality-based (e.g., "*Find a similar product, but cheaper*")
  - Quantity-based (e.g., "*Find something similar to this one, but at least $100 cheaper*")

- System-suggested critique (tradeoff-oriented explanation)
  - (revisit) Design **Principle 2** for ORG interface: *Propose improvements and compromises in the category title using conversational language*
  - Favor critique candidate with high tradeoff utility (more gains relative to losses)
  - Diversify multiple critique suggestions

User evaluation results

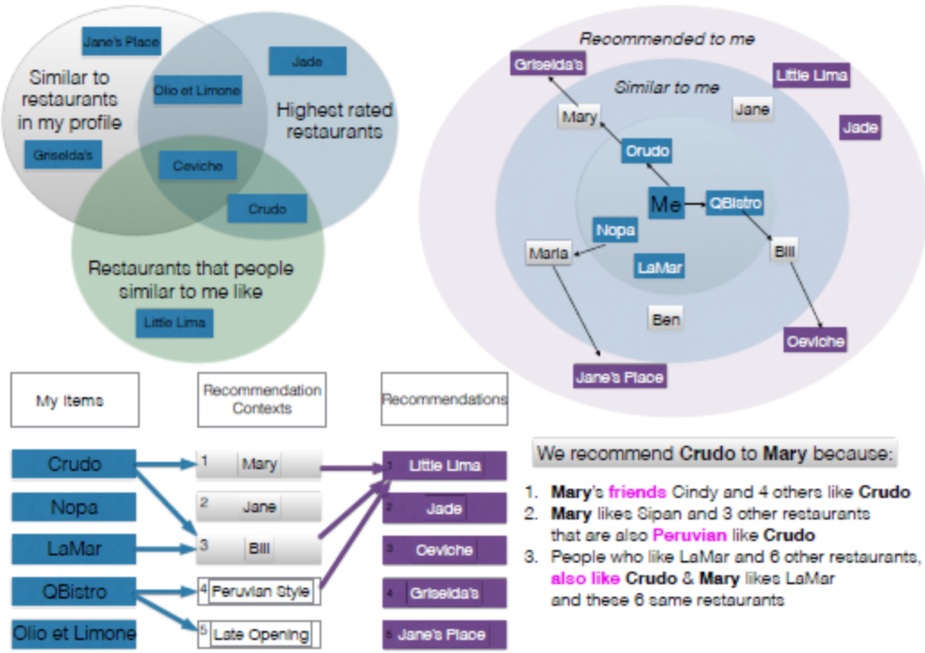
- Critiquing-based system can significantly improve users' **decision accuracy** by **up to 57%**, against non-critiquing based
- Hybrid critiquing (combining both user-initiated critiquing and system-suggested critiques) can achieve the **desired user control** and effectively save users' interaction effort
- Incorporation of <u>sentiment features</u> into the critiquing interface can further improve users' **decision quality**

Critiquing-based conversation in Chatbot



Yucheng Jin, Wanling Cai, Li Chen, Nyi Nyi Htun, and Katrien Verbert. MusicBot: Evaluating Critiquing-based Music Recommenders with Conversational Interaction. In *Proceedings of 28th ACM International Conference on Information and Knowledge Management (CIKM'19)*, pages 951–960, Beijing, China, November 3-7, 2019.

Recent trend – Hybrid explanation



Courtesy image from Kouki et al. (2017)

Style	Sample
Baseline	Recommend based on your visit logs
Demographic	Recommend for "*women in 30s*"
Content	Recommend for who often visit "*Italian restaurant*"
Context	Recommend for use "*with husband/wife*"
Demographic + Context	Recommend for use "*in business entertaining*" of "*men in 50s*"
Content + Context	Recommend for use "*in solitude*" for who often visit "*noodle*"
Demographic + Content + Context	Recommend for use "*with close friends (with drink)*" of "*women in 20s*" who often visit "*cafe*"

Courtesy image from Sato et al. (2018)

Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2017. User preferences for hybrid explanations. In *Proceedings of the 11th ACM Conference on Recommender Systems (RecSys'17)*. ACM, New York, NY, 84–88.

Masahiro Sato, Budrul Ahsan, Koki Nagatani, Takashi Sonoda, Qian Zhang, and Tomoko Ohkuma. 2018. Explaining Recommendations Using Contexts. In *23rd International Conference on Intelligent User Interfaces (IUI '18)*. ACM, New York, NY, USA, 659–664.

Recent trend – Effect of personal characteristics

Personality	"Calm participants (low neuroticism) preferred popularity-based explanations, while anxious participants (high neuroticism) preferred item-based CF explanations" (Kouki et al., 2019).
Need for cognition	"Users with a low need for cognition tend to benefit more from explanations" (Millecamp et al., 2019)
Domain knowledge	"High musical sophistication users felt more supported in making a decision if the RS provides explanations" (Millecamp et al., 2020).

Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2019. Personalized Explanations for Hybrid Recommender Systems. In *IUI '19*. ACM, New York, NY, USA, 379–390.

Martijn Millecamp, Nyi Nyi Htun, Cristina Conati, and Katrien Verbert. 2019. To Explain or Not to Explain: The Effects of Personal Characteristics when Explaining Music Recommendations. In *IUI '19*. ACM, New York, NY, USA, 397–407.

Martijn Millecamp, Nyi Nyi Htun, Cristina Conati, and Katrien Verbert. 2020. What's in a User? Towards Personalising Transparency for Music Recommender Interfaces. In *UMAP '20*, ACM, New York, NY, USA, 173–182.

Conclusion

User-centric design

Algorithm (explanation generation)

+

Interface (explanation presentation)

Transparency | Persuasiveness | Trust | Decision quality

Improved user experiences and satisfaction with the recommender system

Thanks!

Contact info:

Dr. Li CHEN

lichen@comp.hkbu.edu.hk

Homepage:

http://www.comp.hkbu.edu.hk/~lichen/