

Argument Search with Voice Assistants

Master's Thesis by Kevin Lang

Referees:

Prof. Dr. Benno Stein

Prof. Dr. Ing. Eva Hornecker

Advisor:

Johannes Kiesel

- Motivation
- Study 1: Online Survey
- Study 2: Wizard of Oz Experiment
- Summary
- Conclusion
- Future Work

Why...

- adopting a pet?
- buying a car?
- voting for this candidate?



Where can I find arguments?

- Sources?
- Trustworthy?
- Convincing?
- Counter-arguments?



The screenshot shows the args search engine interface. At the top, there is a logo with three colored segments (red, green, blue) followed by the word "args". A search bar contains the query "Should I vote for Trump?". Below the search bar, it says "Page 1 of 521 arguments, 263 pro, 258 con (retrieved in 1715.5ms)" and provides links to "Pro vs. Con View" and "Topic Space View".

Pro

- It is without question that this Presidential...**
<http://www.debate.org/debates/Vote-for-trump/1/>
 Trump's comments in the media have drawn a lot of attention. People saying he is a loose canon ... on the American people. Donald Trump has ideas as I will continue to explain. He wants ... ▾ score
- In this debate my opponent must argue that...**
<http://www.debate.org/debates/Republicans-should-try-to-derail-Trump/1...>
 In this debate my opponent must argue that republicans **should** keep and support Trump as their nominee while I argue that they **should** deny him the nomination or **vote for** ... ▾ score
- I'd like to start out my closing statements by...**
<http://www.debate.org/debates/Hunting-should-NOT-be-illegal./1/>
 was over whether or not hunting **should** remain legal, not the justification of killing. I said yes ... ado, it's time for my closing statements. I have used many examples to prove ... ▾ score

Con

- I'm just curious, who actually supports trump? And...**
<http://www.debate.org/debates/Vote-for-trump/1/>
 I'm just curious, who actually supports trump? And also, what is their reasoning for it. This is why I made this debate. I have no rules or structure for this debate except that round ... ▾ score
- Alright. I'm going to begin by countering my...**
<http://www.debate.org/debates/Hunting-should-NOT-be-illegal./1/>
 claimed I'm telling voters how to **vote** ; I'm explaining why a point **SHOULD** be ignored. My opponent asserts ... to argue against the value ; biopower flows through. Next, I agree. ... ▾ score
- I accept this debate and will argue that the...**
<http://www.debate.org/debates/Resolved-The-Father-Should-Not-Have-Say...>
 I accept this debate and will argue that the father of an unborn child **should** have 50% equal say ... to accept the debate, I will simply use this round as acceptance and allow the ... ▾ score

- Argument Search Engine for the Web (Wachsmuth et al., 2017)



I'm afraid I can't do that, Dave.

2001: A Space Odyssey, 1968



Hello, computer?

Star Trek IV: The Voyage Home, 1986

Conversational Voice Assistant

- convenient to use and hands-free
 - used for many small tasks
- Search for arguments
- Future goals:
 - Voice assistant as discussion partner
 - Decision making



Core Questions in this Thesis

Why people want to use a voice assistant for argument search?

How does the user interact with the novel system?

Which responses do they expect from it?

Online Survey

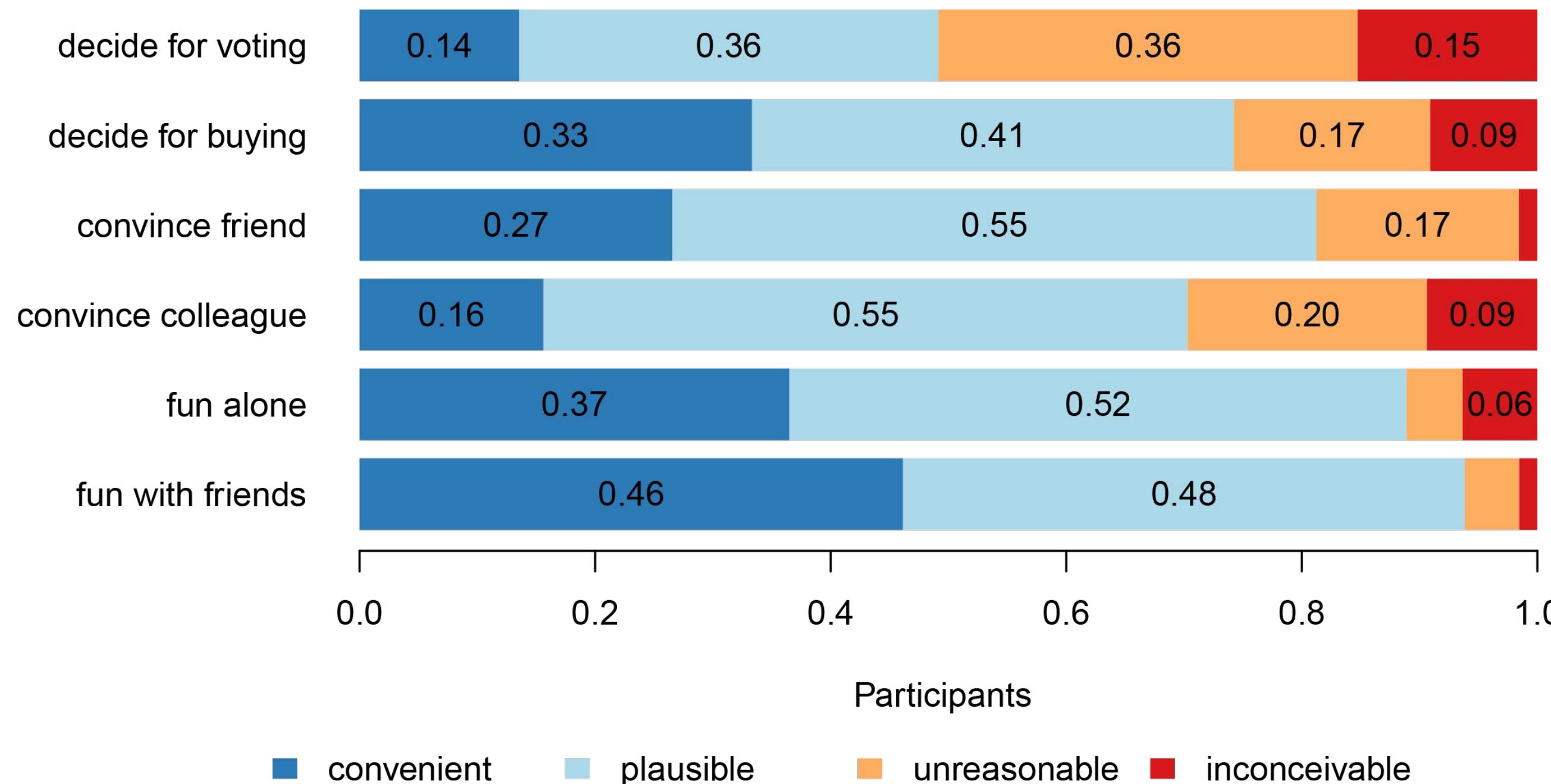
- Asking about the acceptance of:
 - Motivations
 - Situations (Locations, Audiences)
 - Possible Features



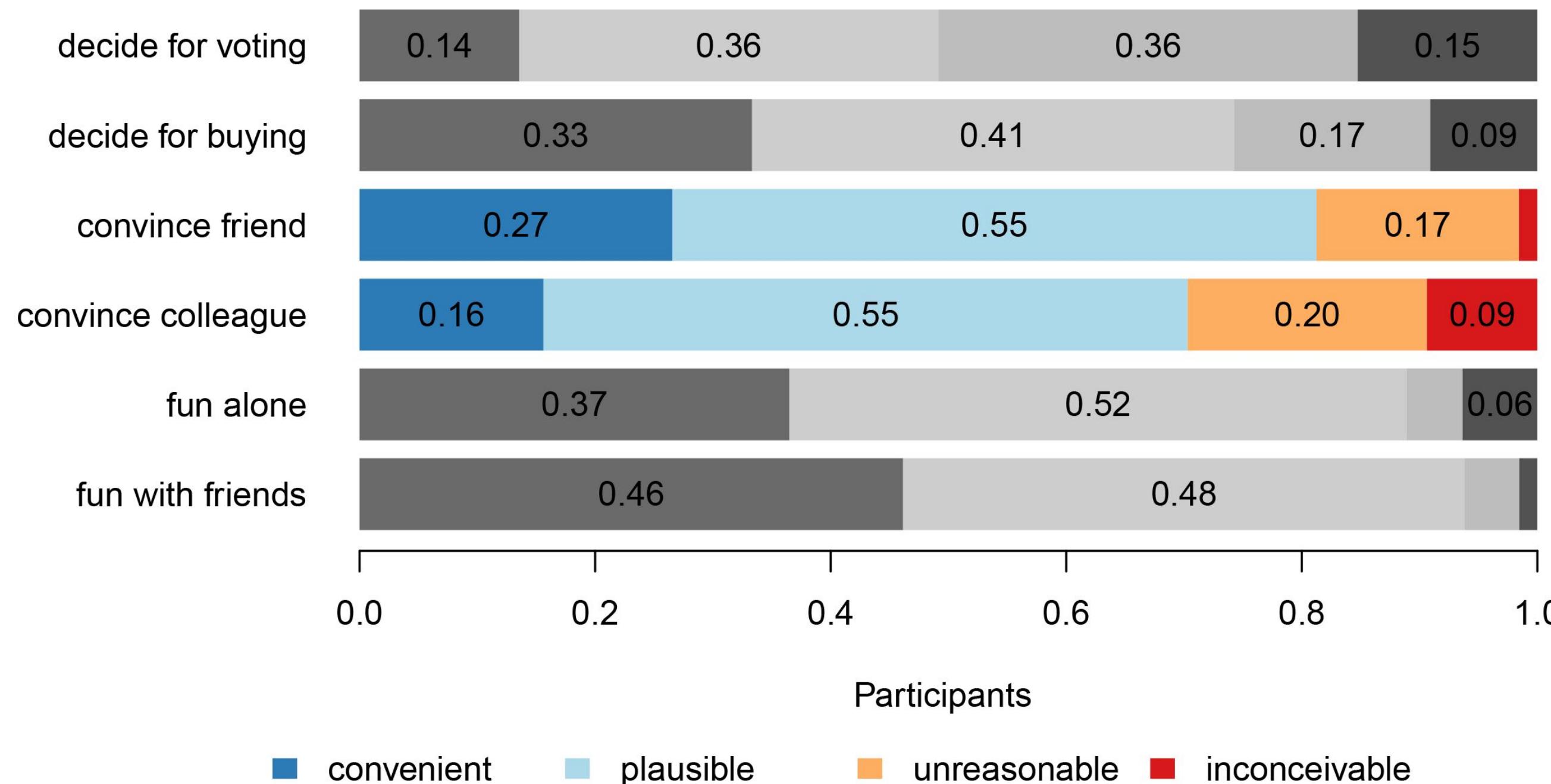
Process

- 67 participants
- 39 English, 28 German
- 18~30 years(49), 31~49(11), 50~64(5), 65+ (1)

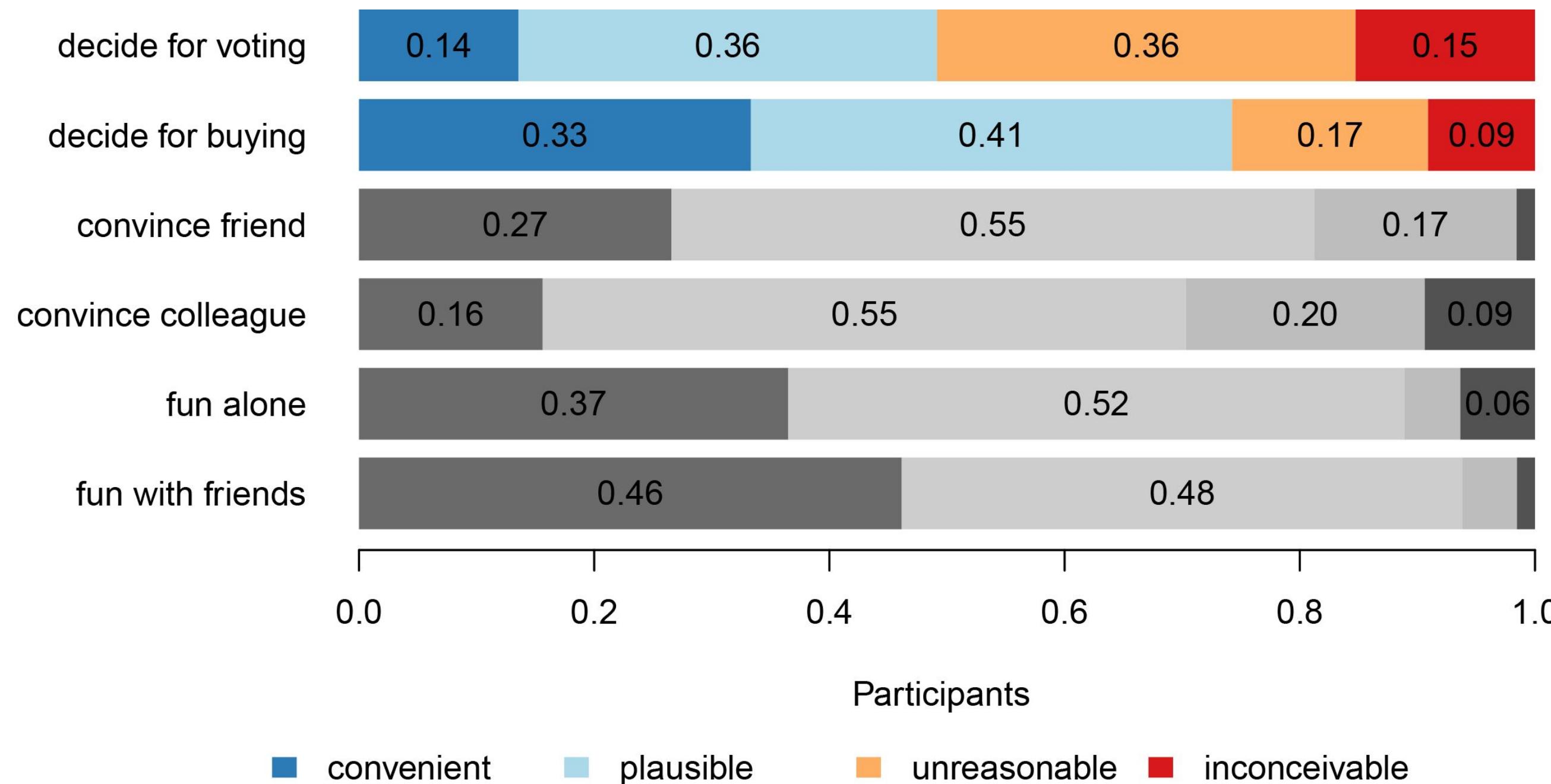
Motivations



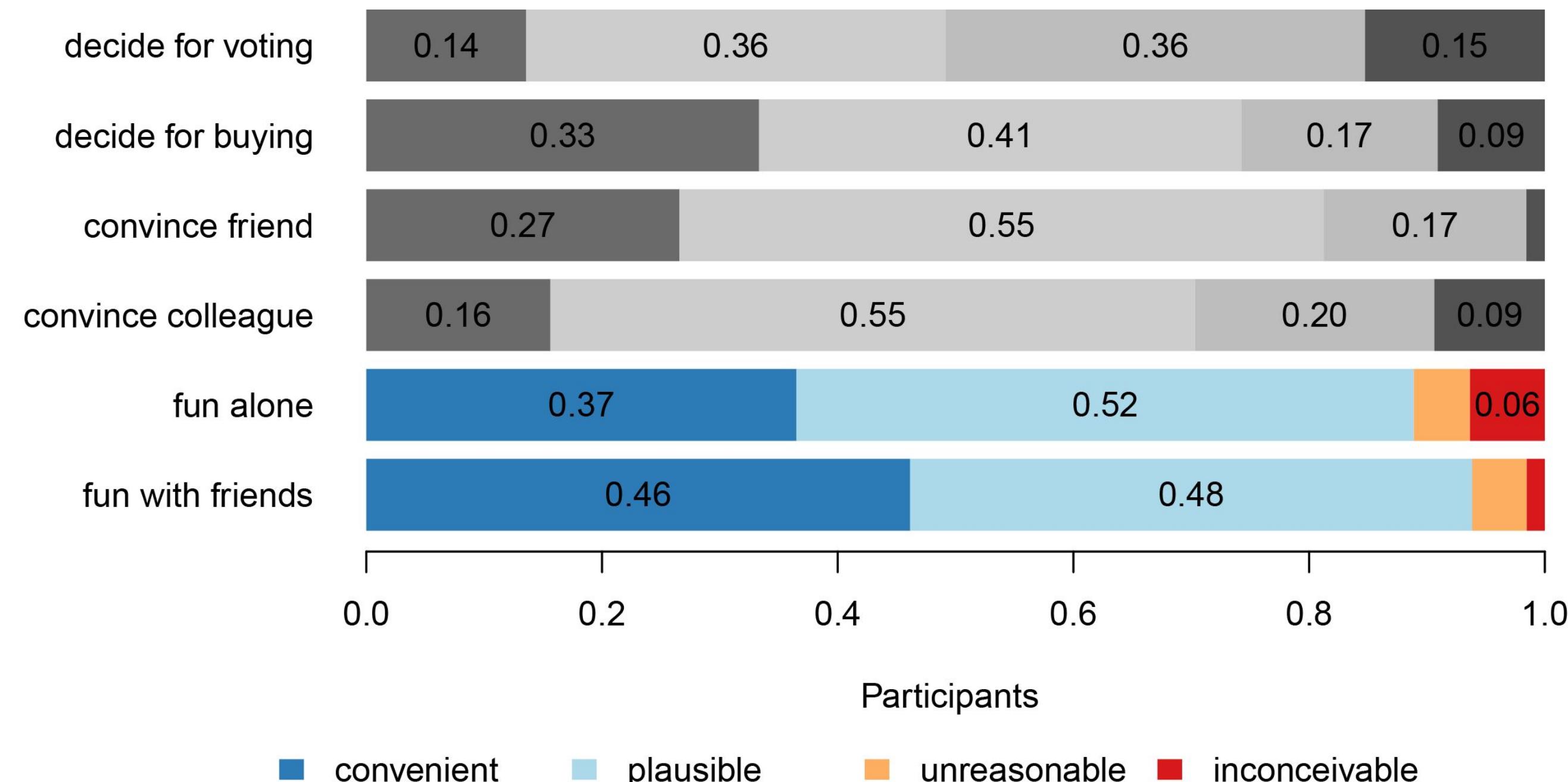
Motivations



Motivations

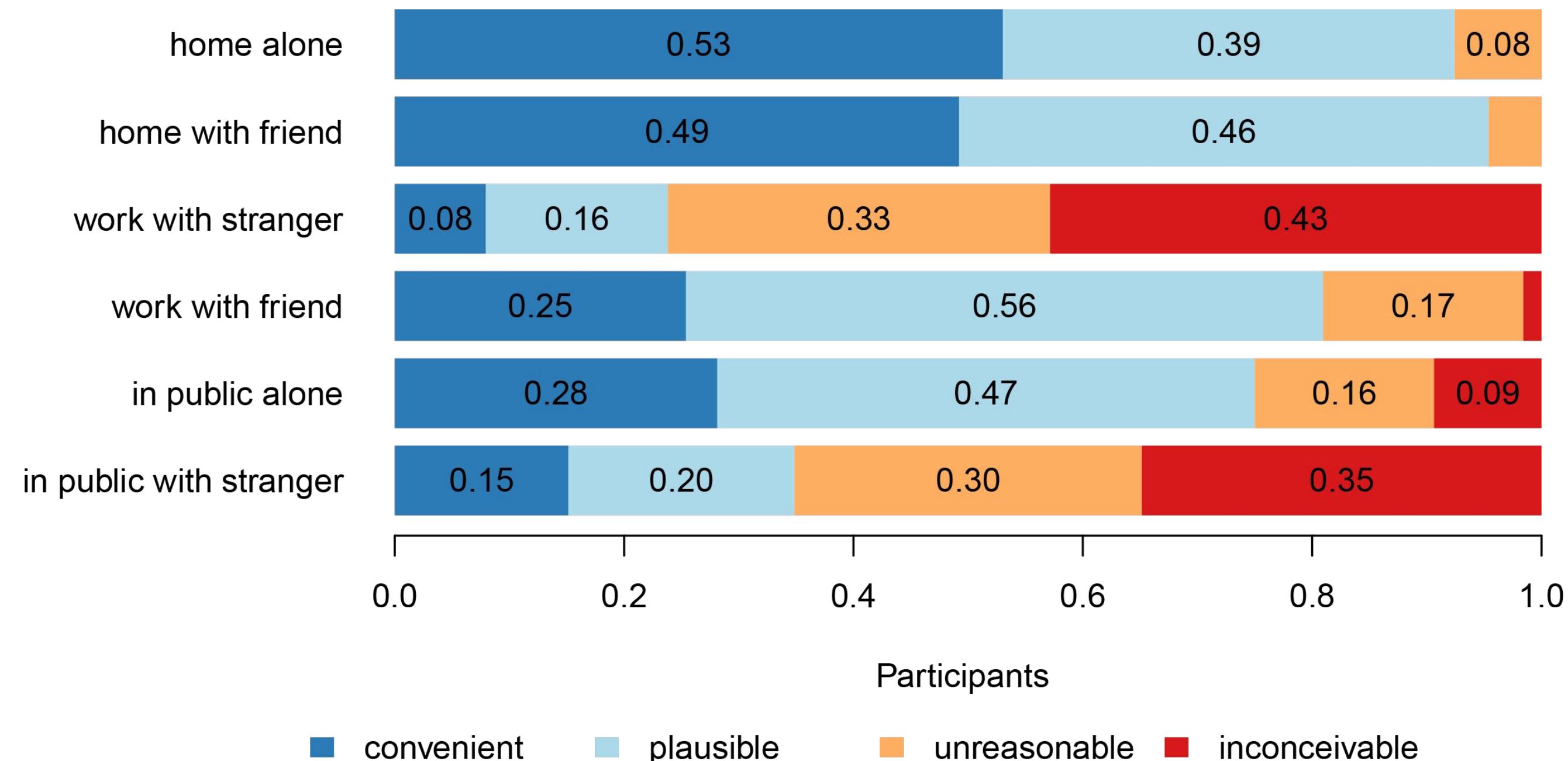


Motivations



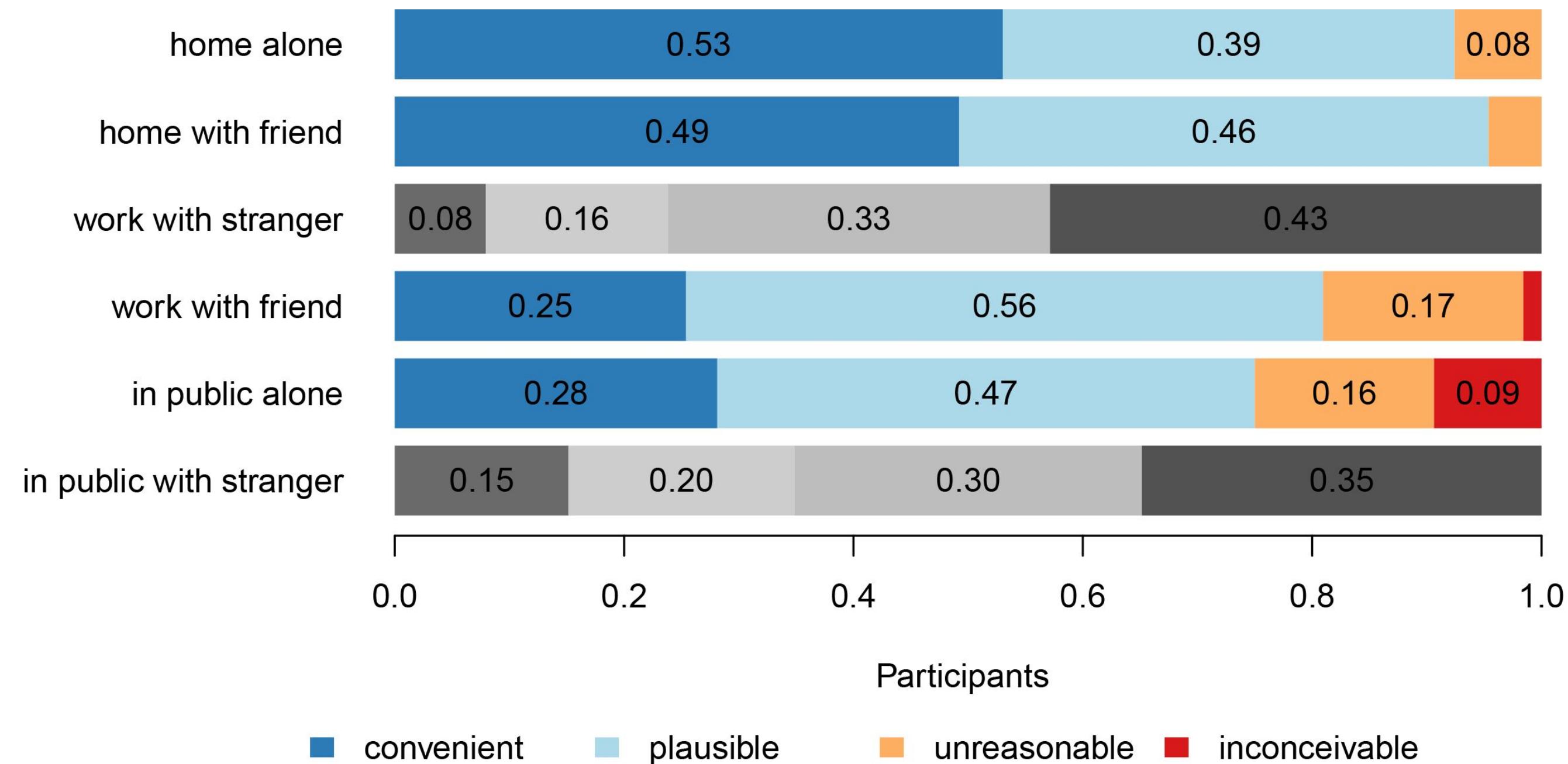
Fun motivations derived from: "Like Having a Really Bad PA": The Gulf Between User Expectation and Experience of Conversational Agents (Luger and Sellen, 2016)

Situations



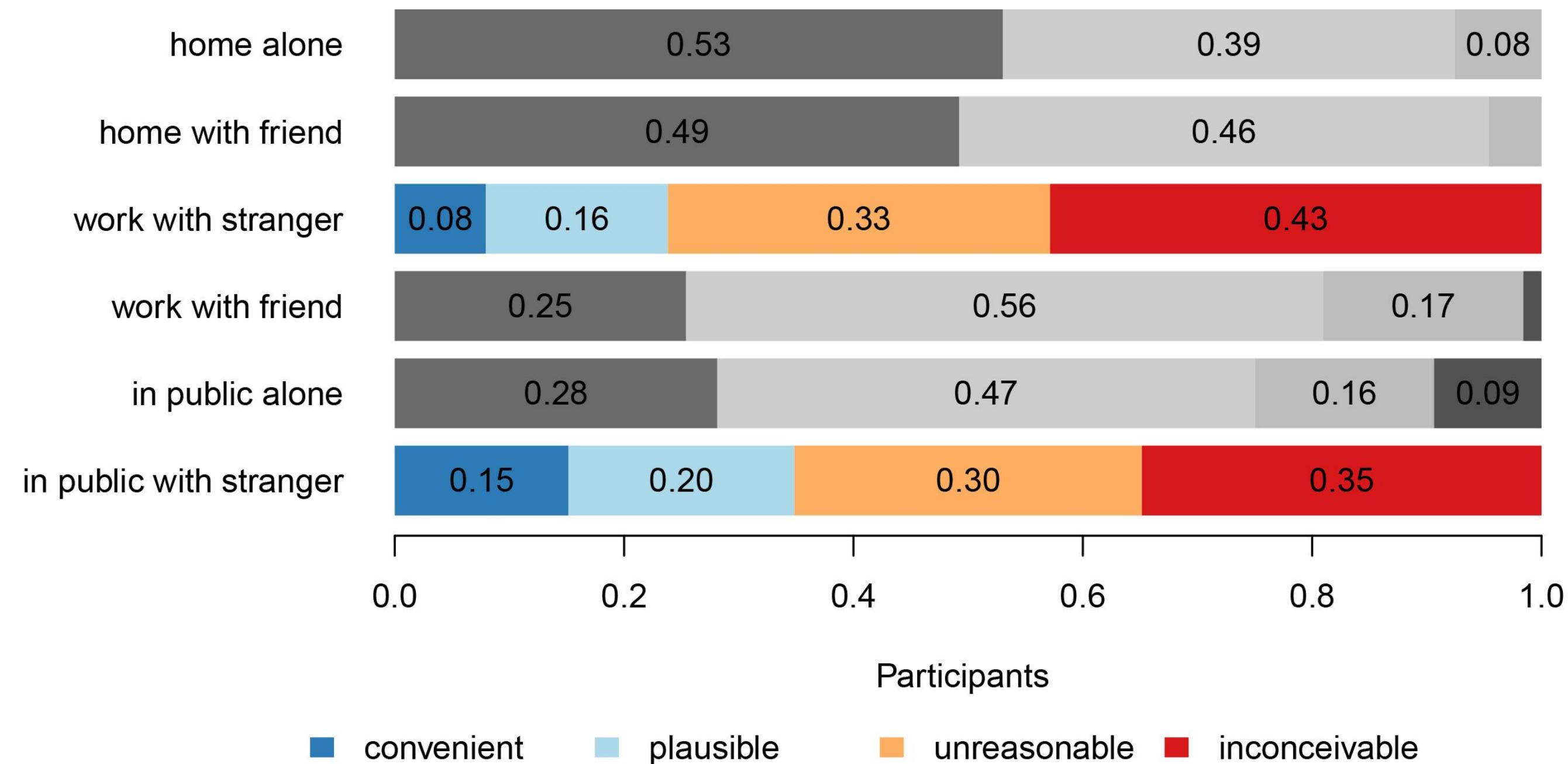
- Similar insights: “Evaluating the Social Acceptability of Voice Based Smartwatch Search” (Efthymiou and Halvey, 2016)

Situations



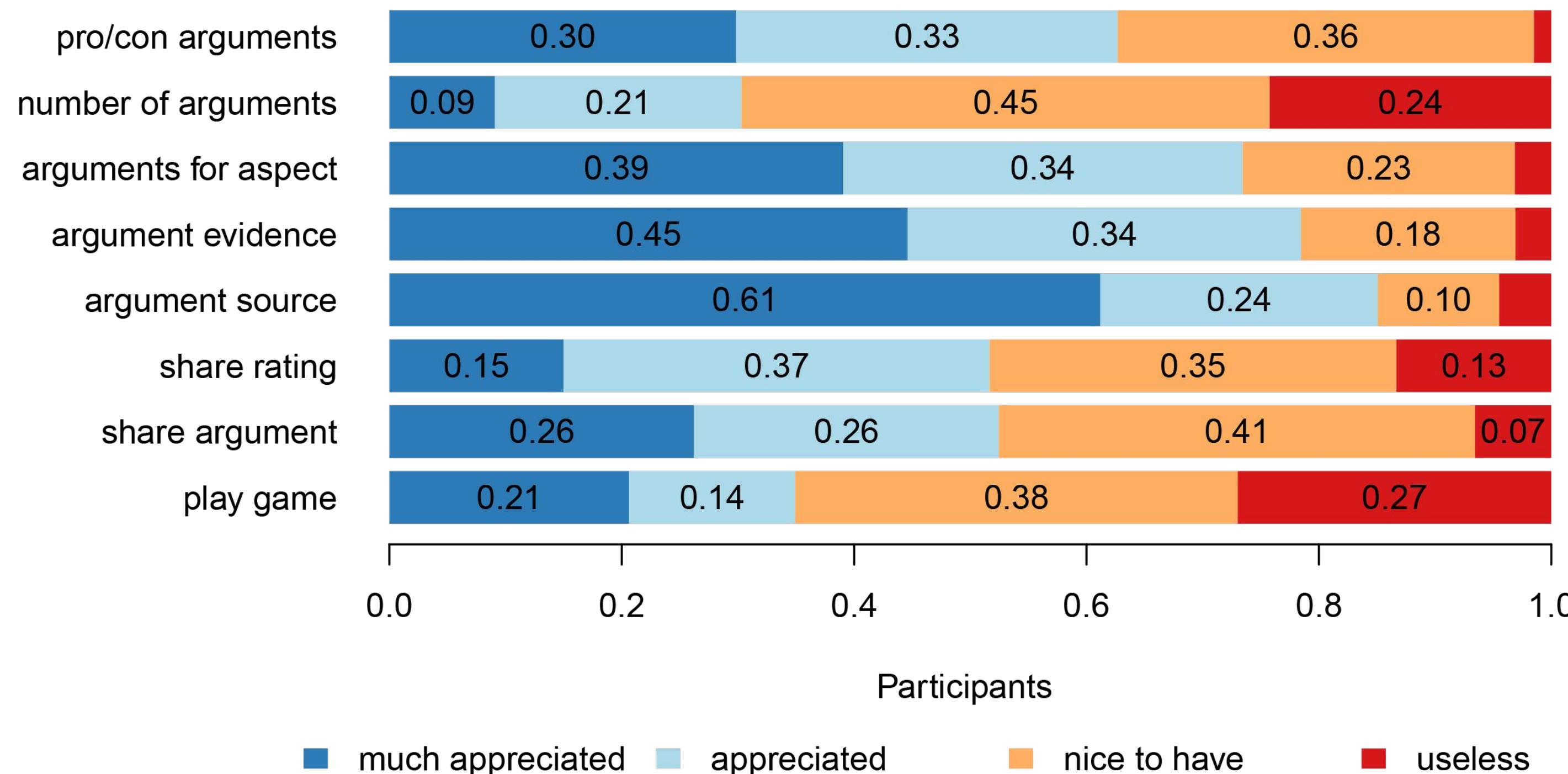
- Similar insights: “Evaluating the Social Acceptability of Voice Based Smartwatch Search” (Efthymiou and Halvey, 2016)

Situations

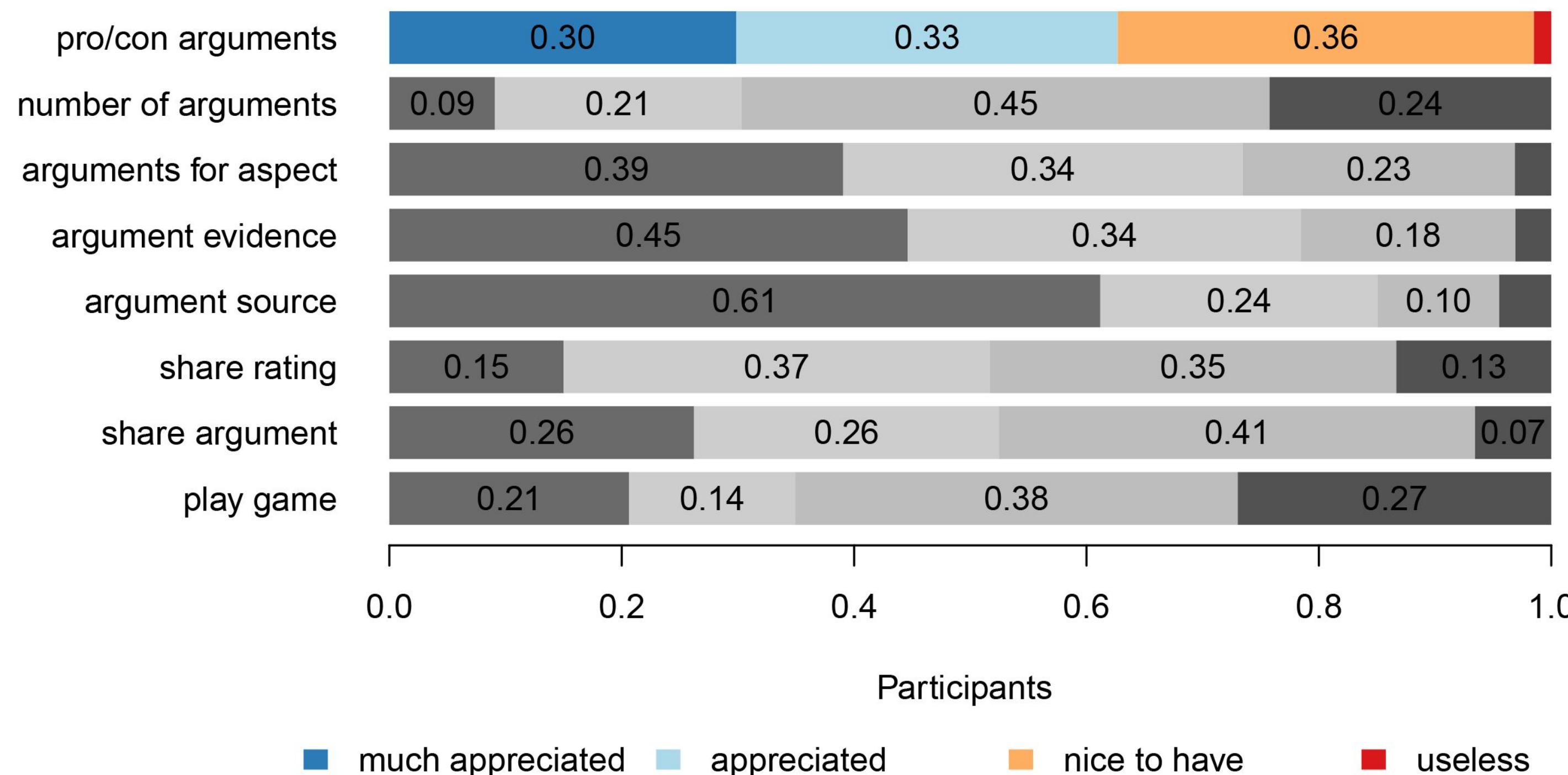


- Similar insights: “Evaluating the Social Acceptability of Voice Based Smartwatch Search” (Efthymiou and Halvey, 2016)

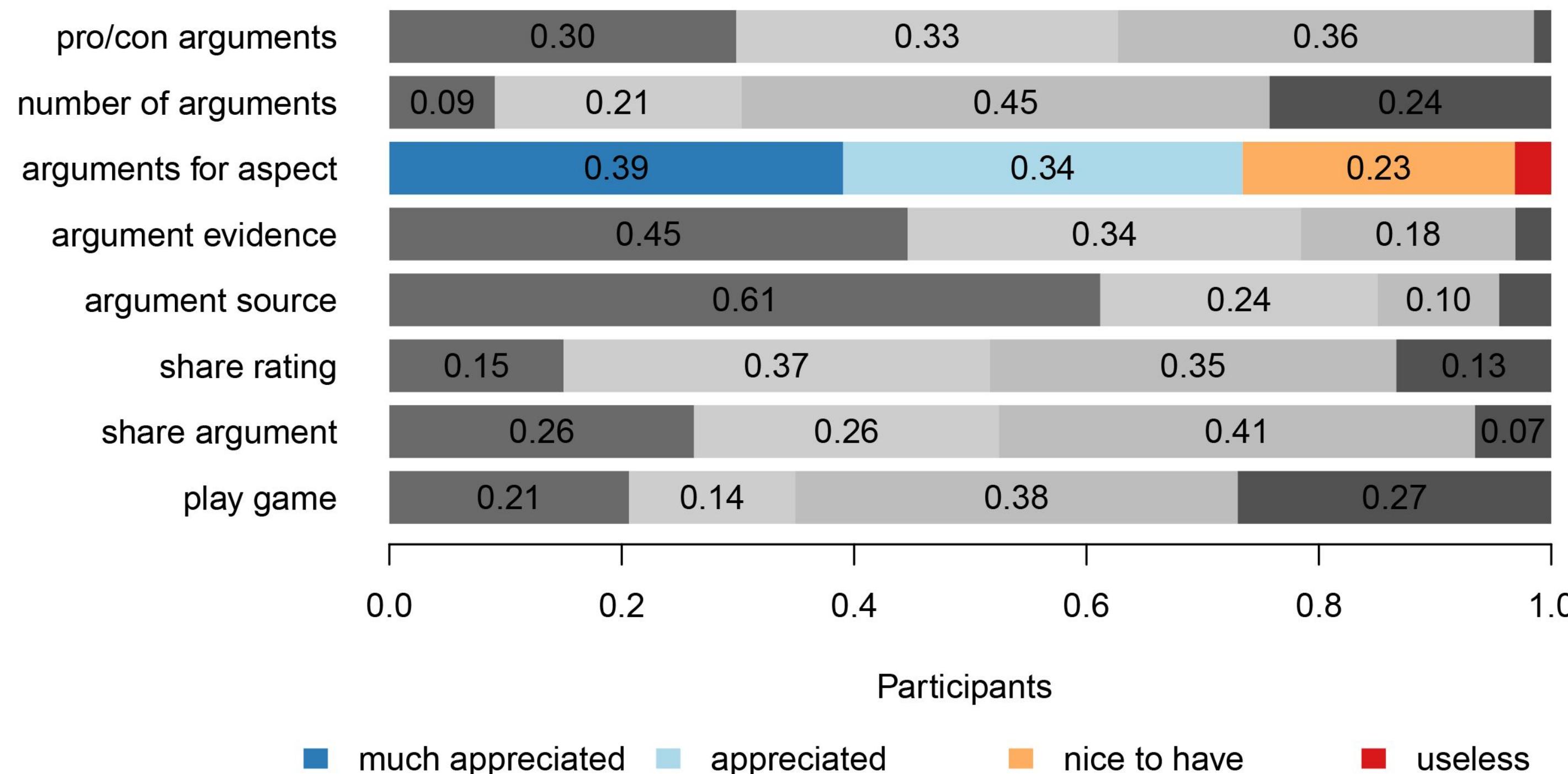
Features



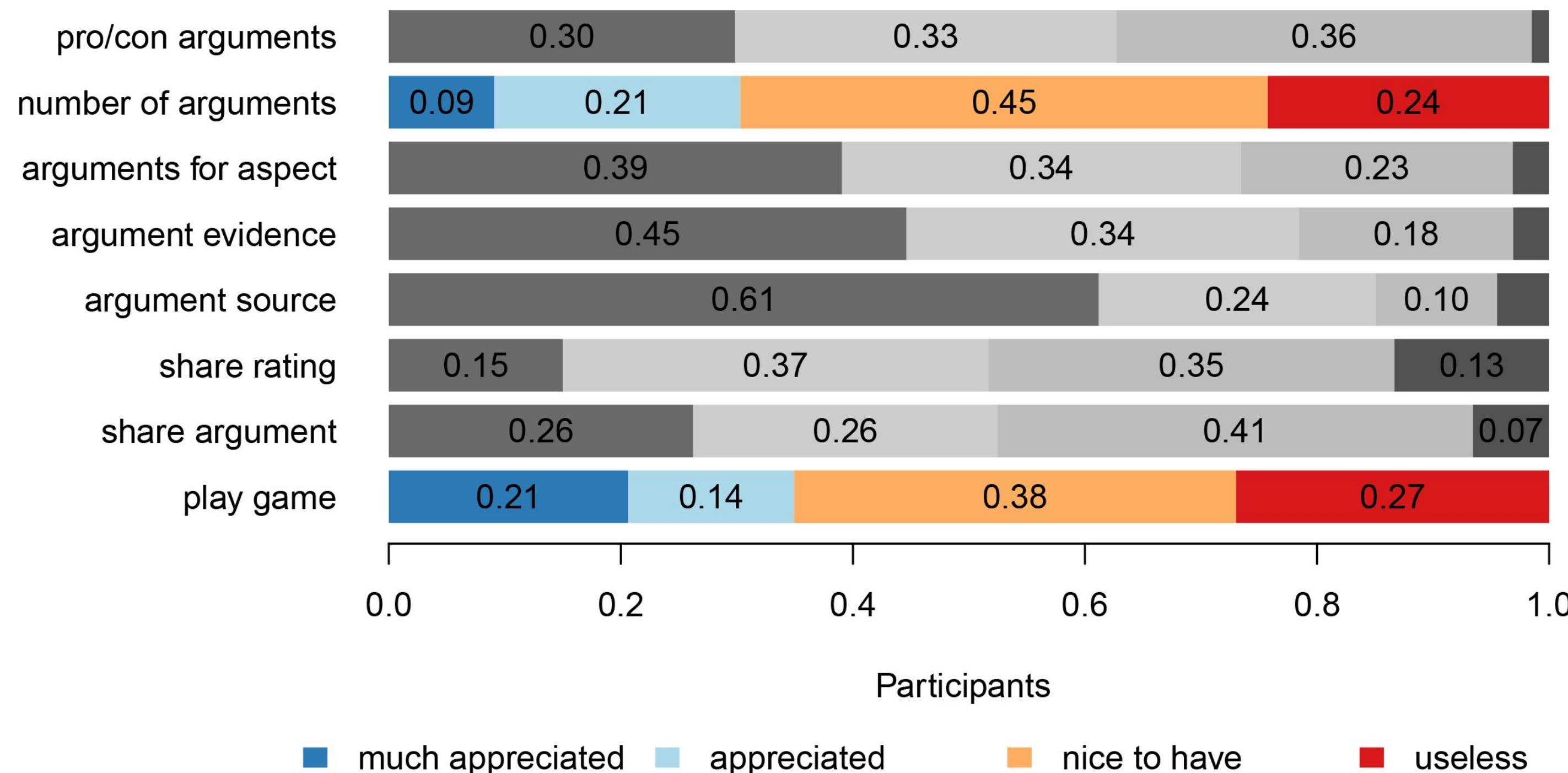
Features



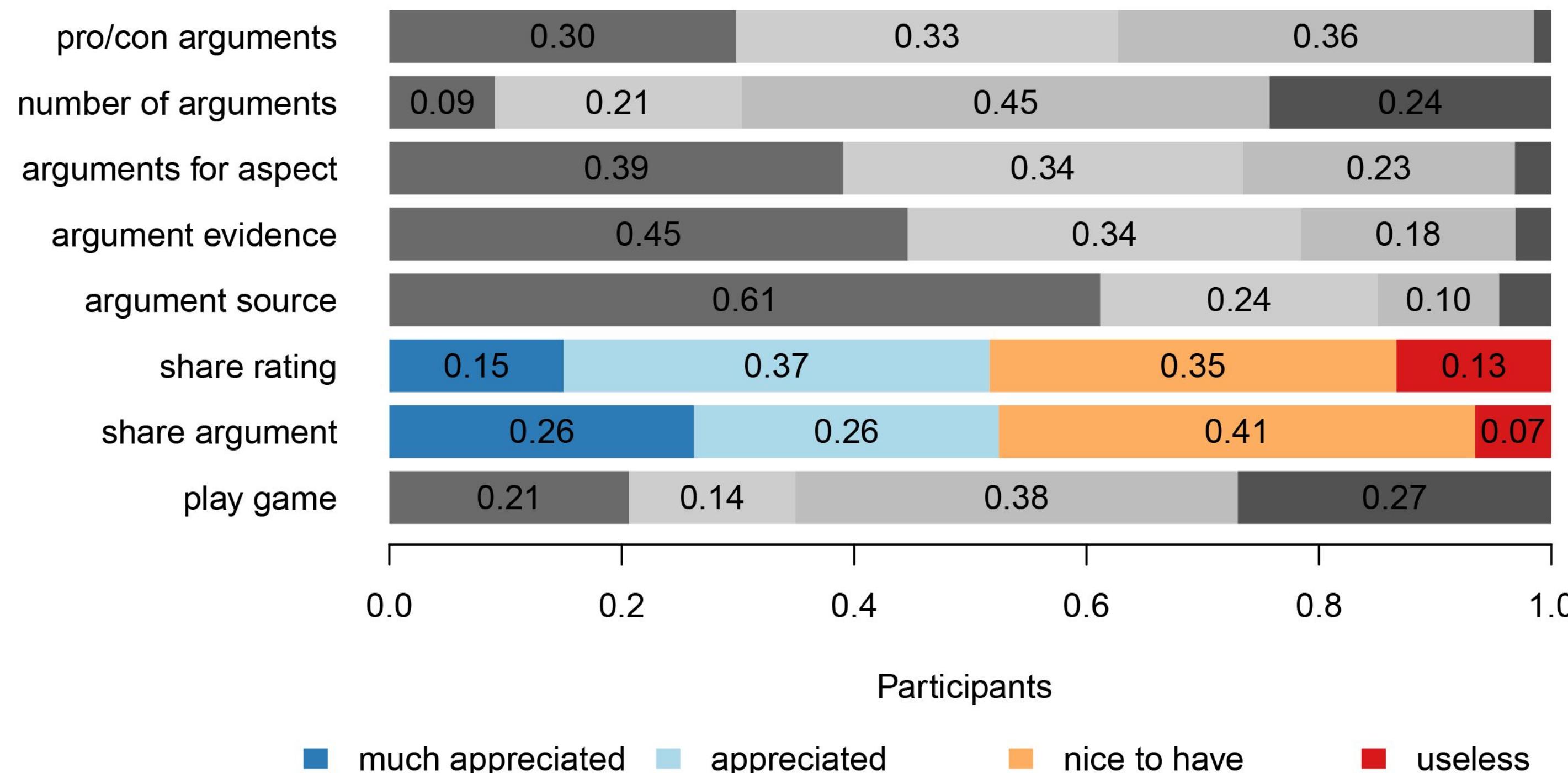
Features



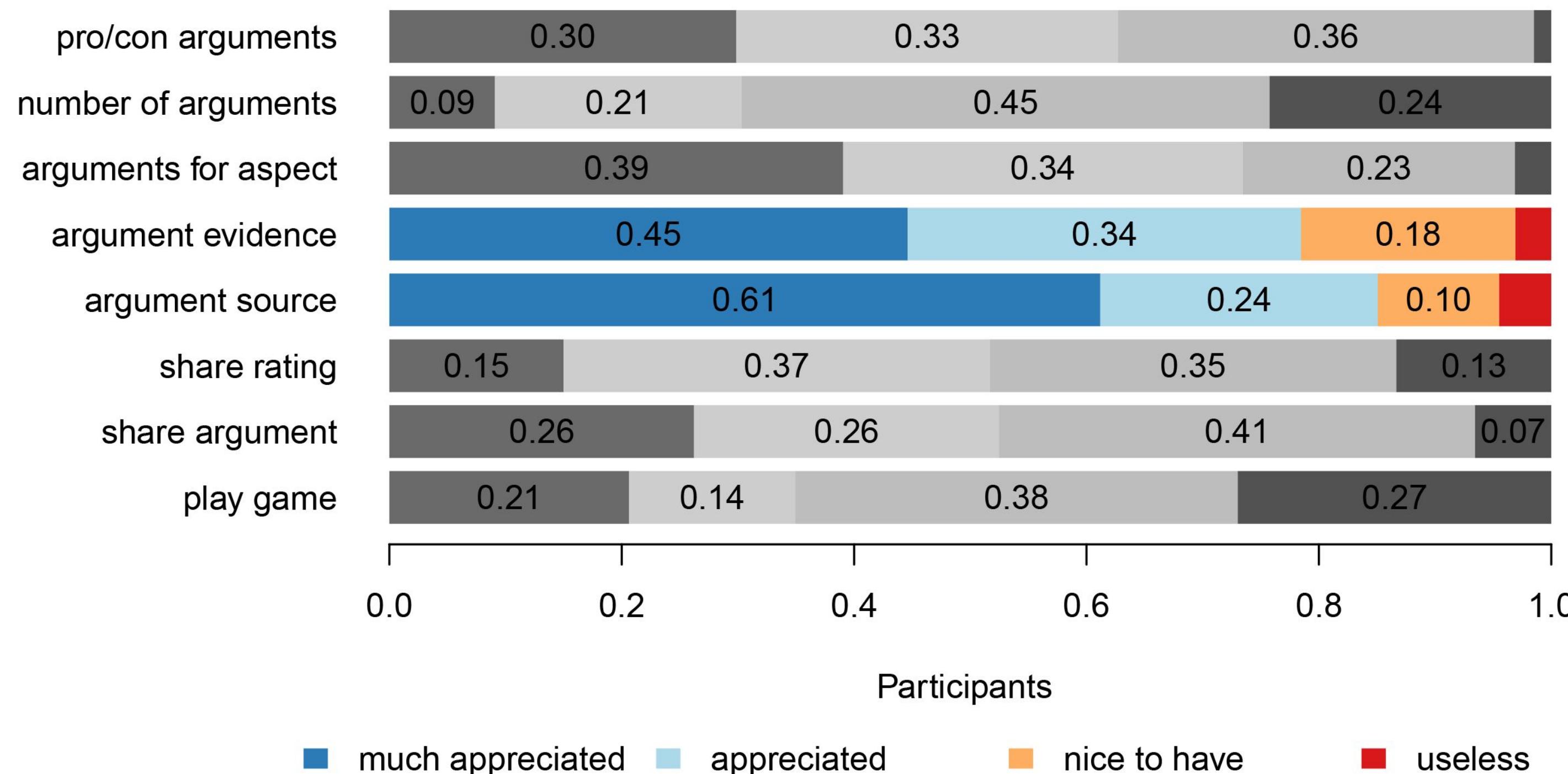
Features



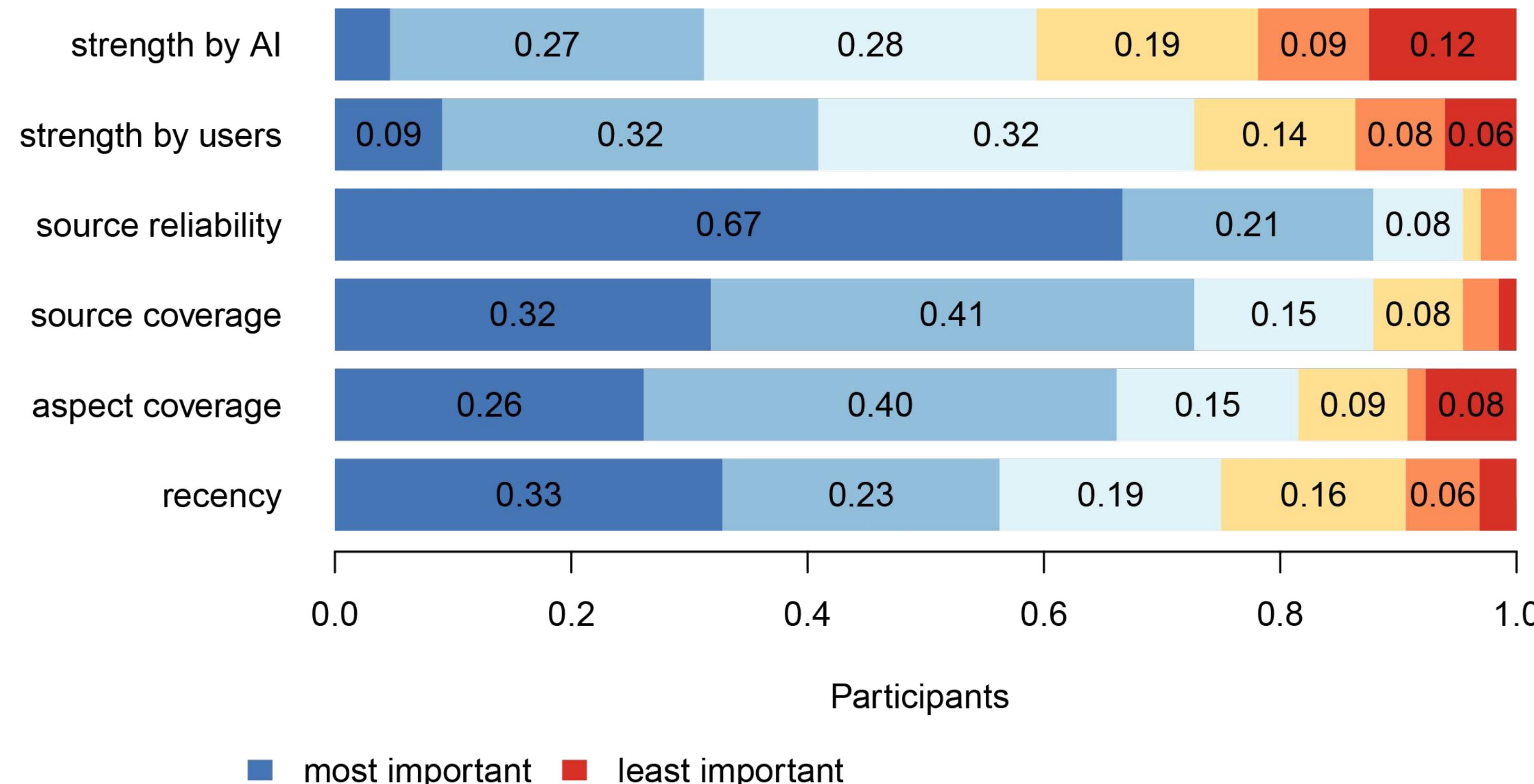
Features



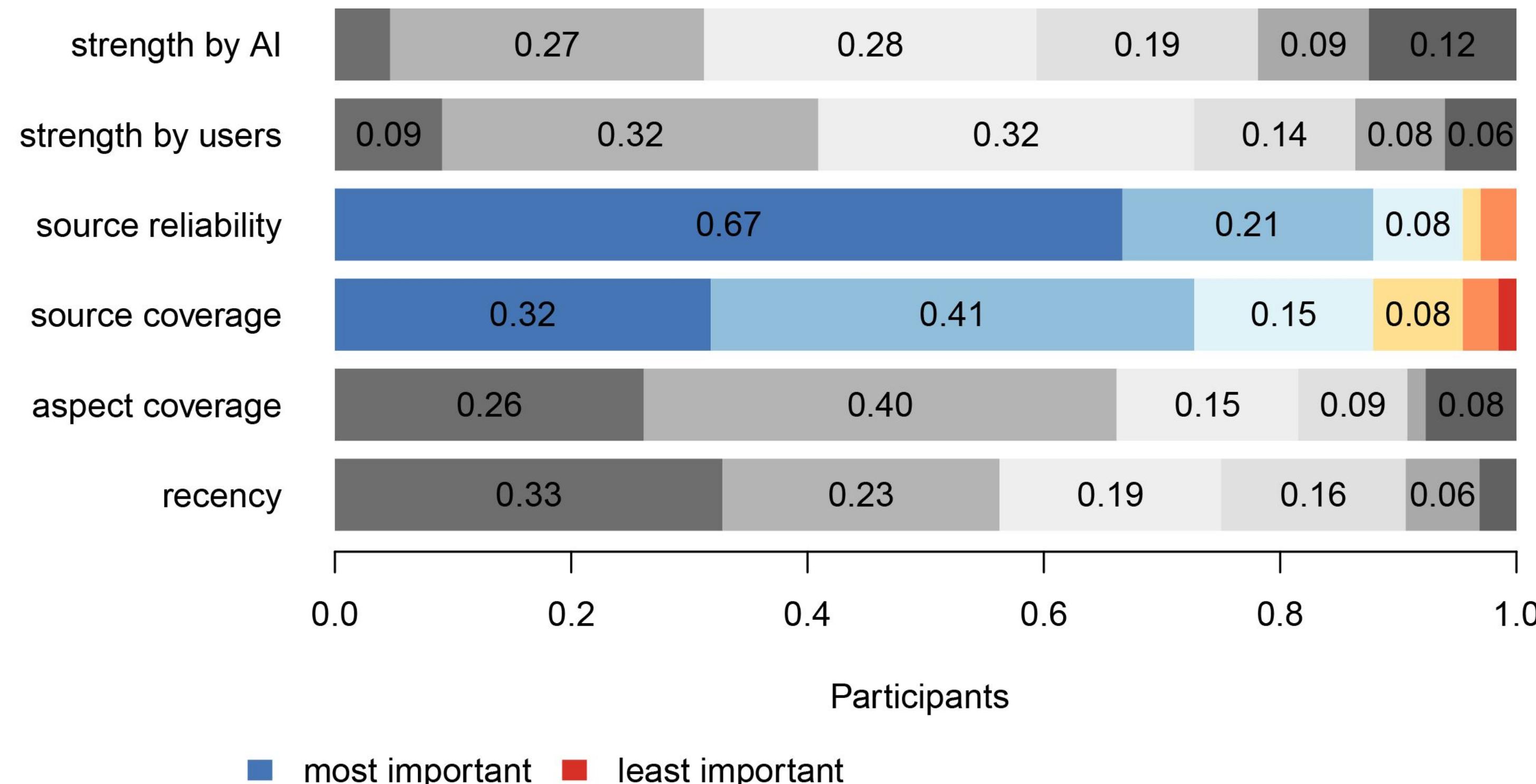
Features



Ranking Criteria

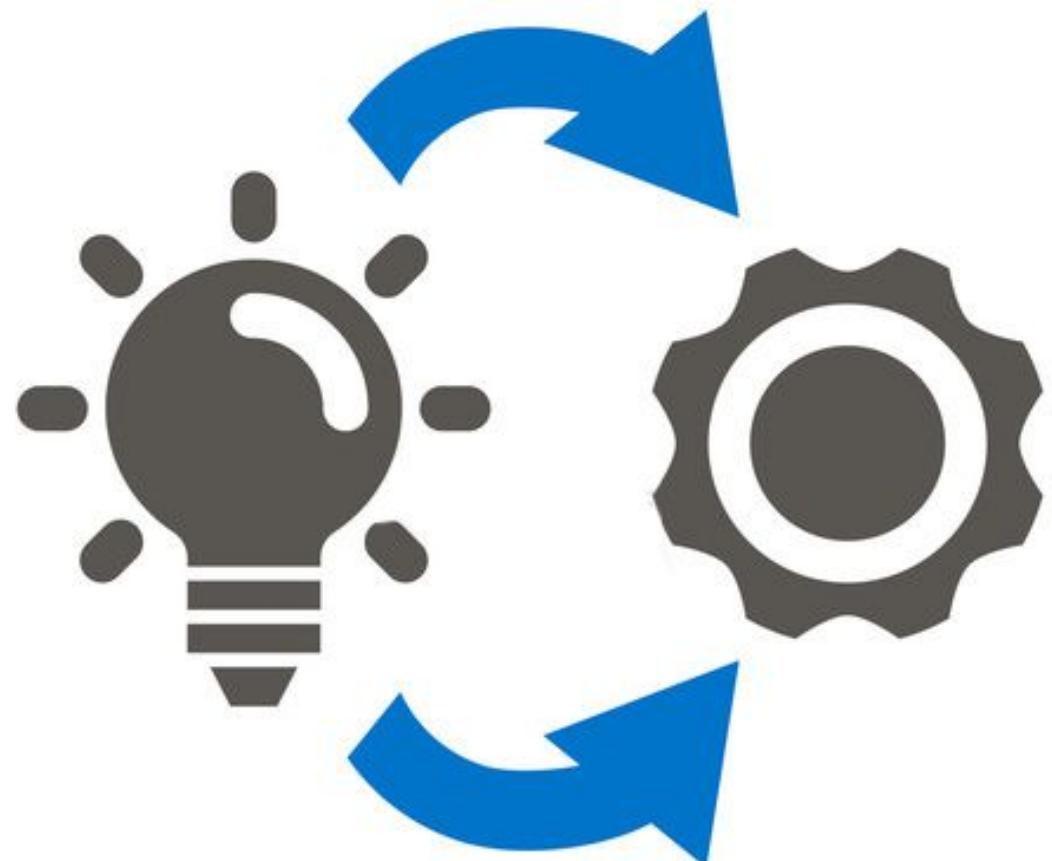


Ranking Criteria



Implementation and Evaluation

- Argument search engine not reliable enough
- Bad voice recognition
- Wrong matching of intents



Wizard of Oz Experiment

- Mock-up prototype
- Avoid problems in
 - speech recognition
 - intent matching
 - system errors



Variables

Motivations:

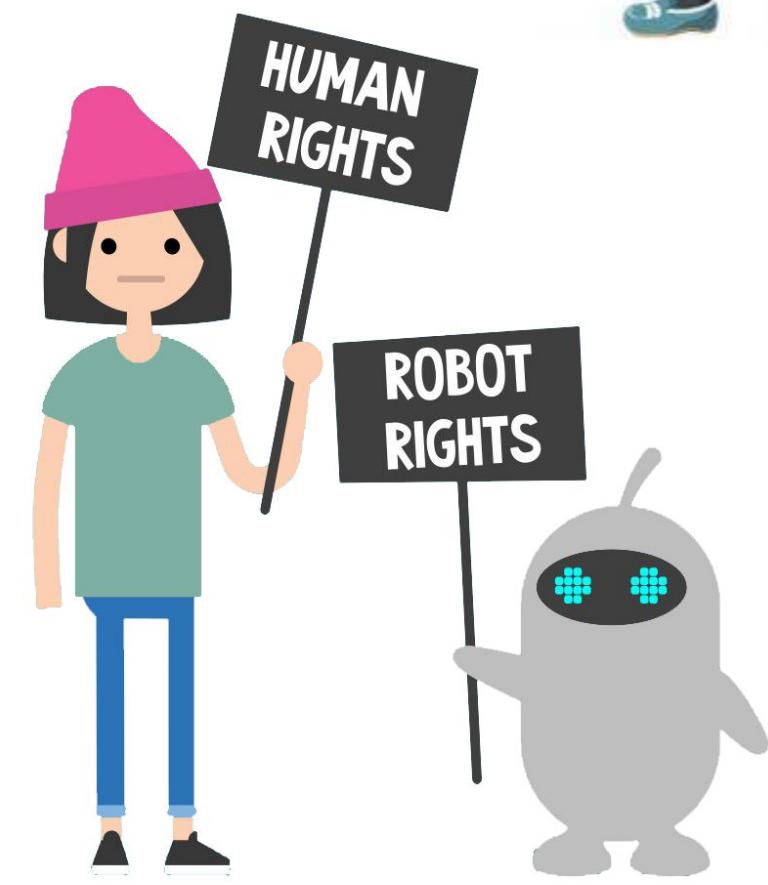
- **Making a decision**
- **Convince somebody**

Behaviour of the system:

- **Without category-guideline***
- **With category-guideline**

* "Investigating how conversational search agents affect user's behaviour, performance and search experience" (Dubiel et al., 2018)

Topics



Experimental set-up



Agent-side

- Prepared topics with arguments
 - Splitted in categories
 - Annotated with total numbers
- Behaviour rules
 - Conversational rules
 - Utterances for intents
 - How to present arguments



User-side



- Set-up:
 - Comfortable sofa
 - Voice interface on armrest
- Participants:
 - 12 male, 6 female
 - 18~30 years (13), 31~49 (5)
 - English level intermediate or proficient

Transcript

- Transcribed 72 audio records, classified with action tags
- 936 turns by the agent, 956 turns by the users
- 1.808 classified actions by the agent, 1.033 by the users

#	Transcribed recording of user (white) and agent voice (gray)	Corresponding tags
1	Alexa?	<user-activate>
2	Hello.	<agent-greetings>
3	[ahh] Tell me about the [uhm] pros and cons of vegetarians.	<user-open-topic>
4	Open topic “Humas should stop eating animal meat”. There are in total 6 pro and 8 con arguments for this topic. Which one do you want to hear?	<agent-open-topic> <agent-count-arguments> <agent-ask-pro-or-con>

Results

actions by agent	#	actions by users	#
Read pro arguments	204	Affirmation	247
No arguments left	178	Request pro arguments	126
Ask category	170	Negation	105
Count arguments	165	Open topic	77
Ask pro or con arguments	161	Request additional information	65
Read con arguments	160	Request con arguments	62
Ask more arguments	158	Activate	55
...		...	

Results

actions by agent	#	actions by users	#
Read pro arguments	204	Affirmation	247
No arguments left	178	Request pro arguments	126
Ask category	170	Negation	105
Count arguments	165	Open topic	77
Ask pro or con arguments	161	Request additional information	65
Read con arguments	160	Request con arguments	62
Ask more arguments	158	Activate	55
...		...	

Results

actions by agent	#	actions by users	#
Read pro arguments	204	Affirmation	247
No arguments left	178	Request pro arguments	126
Ask category	170	Negation	105
Count arguments	165	Open topic	77
Ask pro or con arguments	161	Request additional information	65
Read con arguments	160	Request con arguments	62
Ask more arguments	158	Activate	55
...		...	

Results

actions by agent	#	actions by users	#
Read pro arguments	204	Affirmation	247
No arguments left	178	Request pro arguments	126
Ask category	170	Negation	105
Count arguments	165	Open topic	77
Ask pro or con arguments	161	Request additional information	65
Read con arguments	160	Request con arguments	62
Ask more arguments	158	Activate	55
...		...	

Additional Information requests for...

Definitions:

“What does WWF stand for?”

→ encyclopedia

Additional Information requests for...

Definitions:

“What does WWF stand for?”

→ encyclopedia

Product details:

“How much is the average cost of an electric car?”

→ shops & product databases

Additional Information requests for...

Definitions:

“What does WWF stand for?”

→ encyclopedia

Product details:

“How much is the average cost of an electric car?”

→ shops & product databases

Other resources:

“Do you know how many people will be at the Zoo Erfurt tomorrow?”

→ blogs, scientific paper, statistics

Additional Information requests for...

Definitions:

“What does WWF stand for?”

→ encyclopedia

Product details:

“How much is the average cost of an electric car?”

→ shops & product databases

Other resources:

“Do you know how many people will be at the Zoo Erfurt tomorrow?”

→ blogs, scientific paper, statistics

Agent:

“What do you think of this topic?”

→ decision-making ability

Making a Decision (D) vs. Convincing Somebody (C)

	Car		Zoo		Study		Meat		AI		Uniform	
	D	C	D	C	D	C	D	C	D	C	D	C
pro	28	30	18	18	5	7	9	9	7	12	15	17
con	19	16	18	7	8	10	13	8	9	11	13	13
ratio	.60	.65	.50	.72	.38	.41	.40	.53	.44	.52	.54	.67

Making a Decision (D) vs. Convincing Somebody (C)

	Car		Zoo		Study		Meat		AI		Uniform	
	D	C	D	C	D	C	D	C	D	C	D	C
pro	28	30	18	18	5	7	9	9	7	12	15	17
con	19	16	18	7	8	10	13	8	9	11	13	13
ratio	.60▲	.65	.50▲	.72	.38▼	.41	.40▼	.53	.44▼	.52	.54▲	.67

Making a Decision (D) vs. Convincing Somebody (C)

	Car		Zoo		Study		Meat		AI		Uniform	
	D	C	D	C	D	C	D	C	D	C	D	C
pro	28	30	18	18	5	7	9	9	7	12	15	17
con	19	16	18	7	8	10	13	8	9	11	13	13
ratio	.60	.65▲	.50	.72▲	.38	.41▼	.40	.53▲	.44	.52▲	.54	.67▲

Making a Decision (D) vs. Convincing Somebody (C)

	Car		Zoo		Study		Meat		AI		Uniform	
	D	C	D	C	D	C	D	C	D	C	D	C
pro	28	30	18	18	5	7	9	9	7	12	15	17
con	19	16	18	7	8	10	13	8	9	11	13	13
ratio	.60	◀ .65	.50	◀ .72	.38	◀ .41	.40	◀ .53	.44	◀ .52	.54	◀ .67

Category-Guideline

Quantitative data:

- + slightly better ratings in every aspect

Qualitative data:

- + overview
- + comparison
- instruction
- number of categories
- felt limited

Overall Impression

- + fresh and new experience, very comfortable to do it hands-free
- + Nice flexible input
- overview + memory problems
- Skipping and navigation
- missing additional information
- speech synthesis

- First work which combines argument mining, explorative search and voice-based interface
- 1st study: online survey about motivational & situational aspects + possible features
- 2nd study: design of a mock-up prototype + evaluation with Wizard of Oz experiment
- User ratings and measurements of the experiments
- Transcript of 72 sessions between the human agent and the users

Insights

- Pre-Analysis:
 - Situation: at home, mostly alone or with friends
 - Motivation: preferred for tasks with low impact
- The Application:
 - Missing overview of the arguments
 - Memory and navigational problems
 - Possibility to request additional information

Future Work

- Comparison to argument search with web-interface
- Definition of the goal for exploratory tasks
- Including displays in form of home devices or smartphones
- Missing evaluation:
 - States and transitions between user and agent (Markov model)
 - Sentiment of the requests
 - Obstacles and solutions
 - Category selection

Thank you
for your attention!