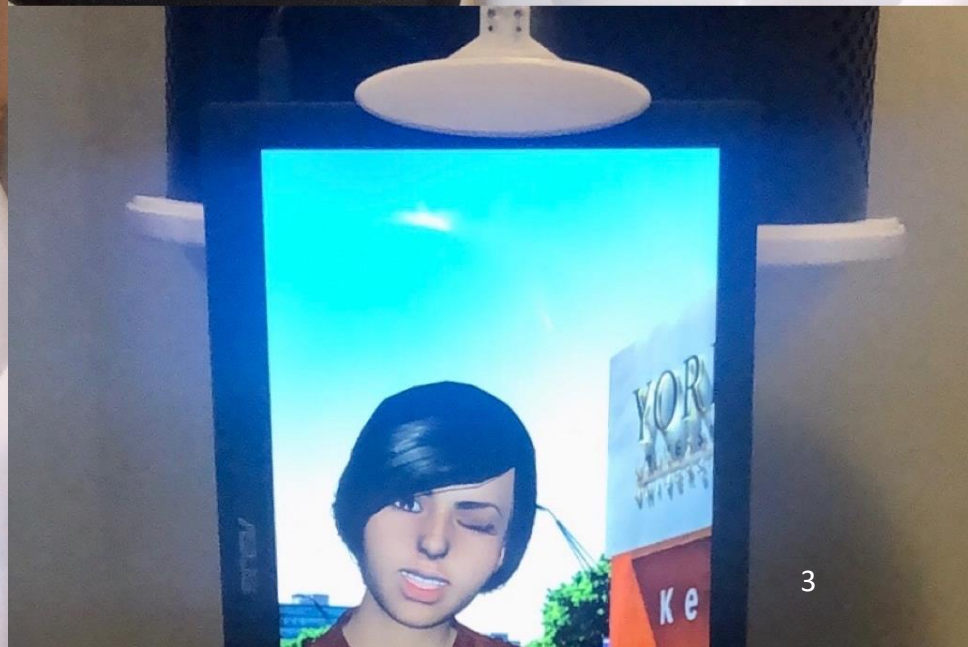
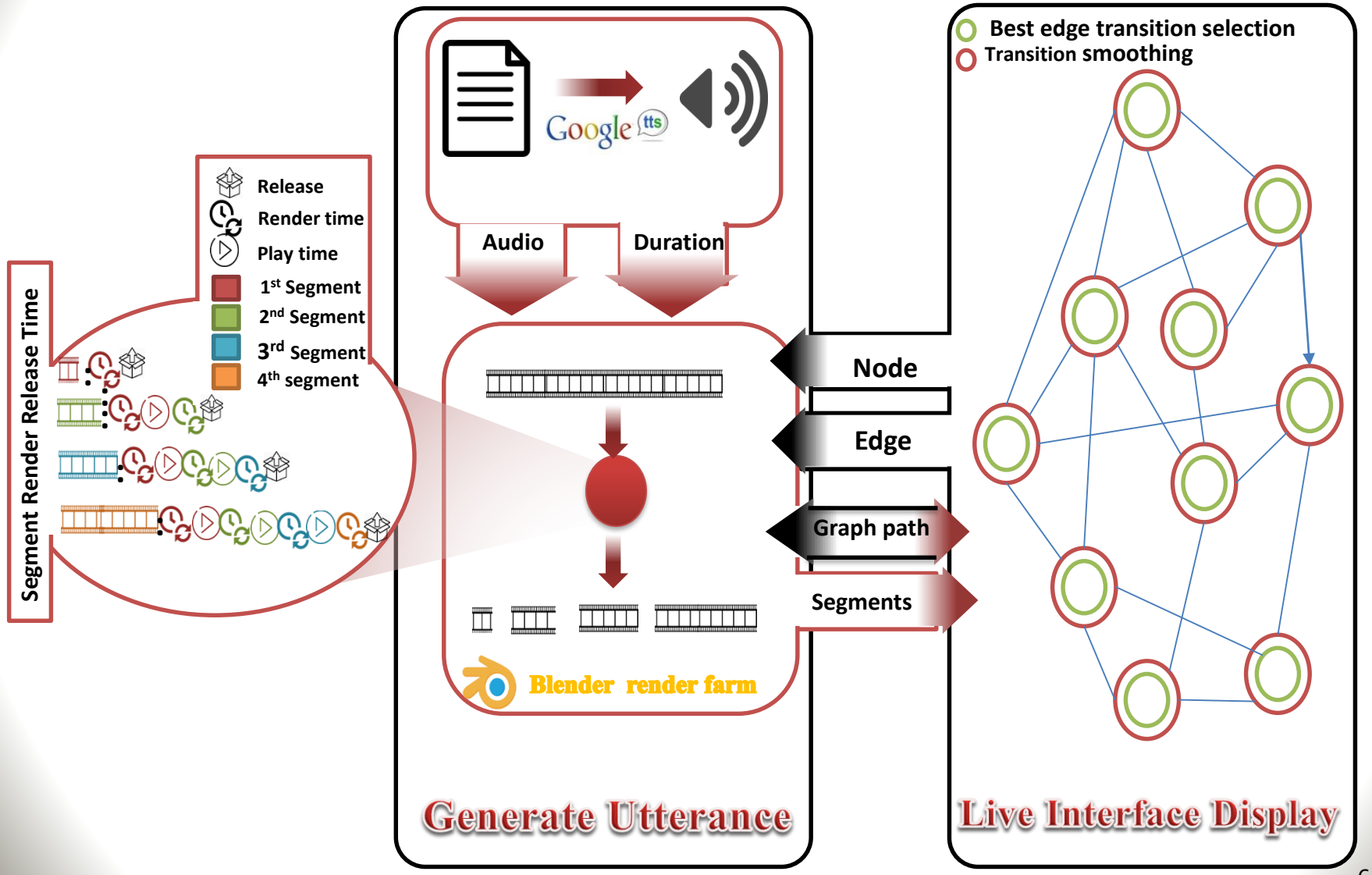
Enas Altarawneh, Michael Jenkin and I Scott MacKenzie
Lassonde School of Engineering, York University, Toronto, Canada
{enas,jenkin,mack}@eecs.yorku.ca

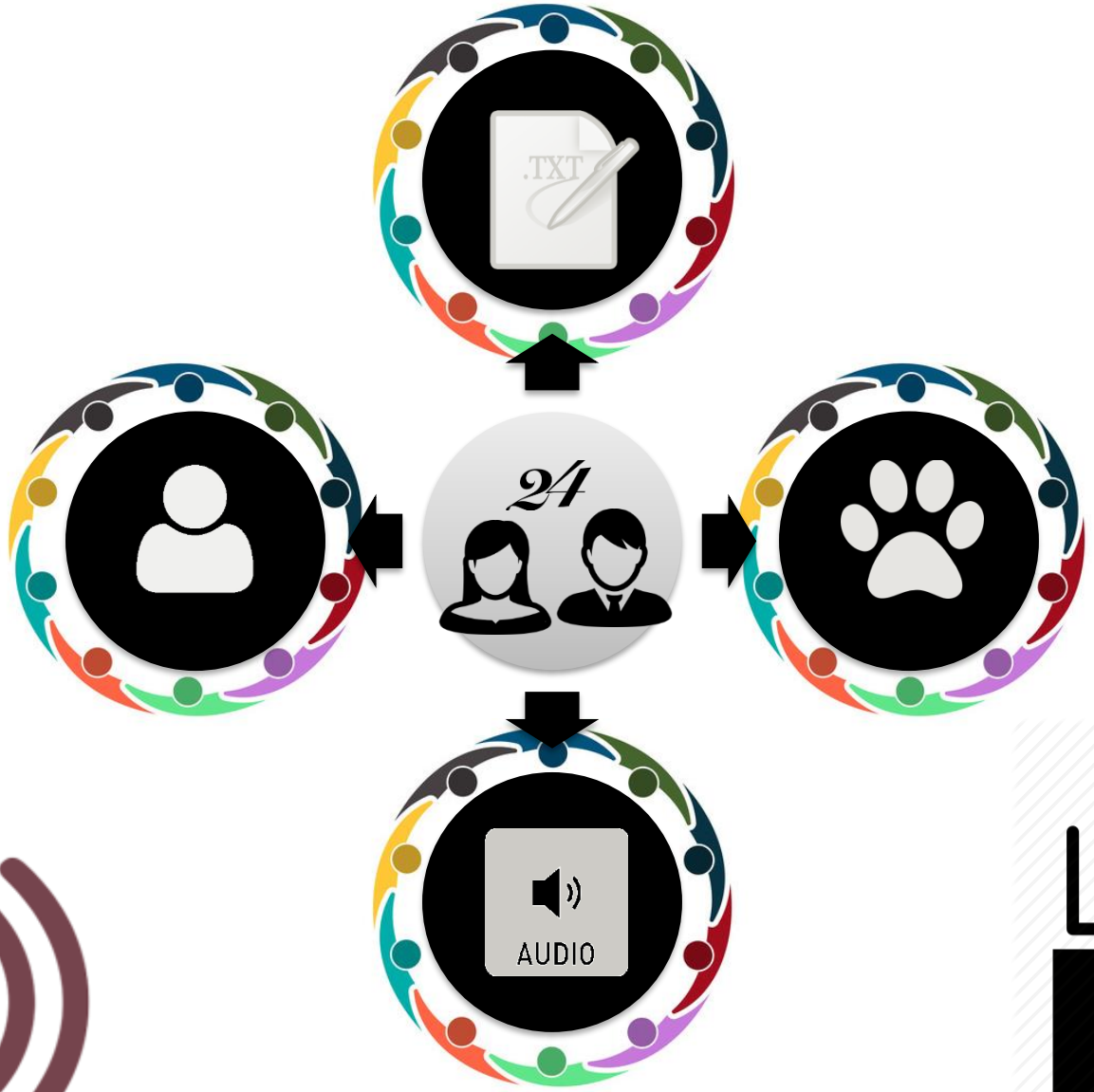
Leveraging cloud based technologies to talk to robots

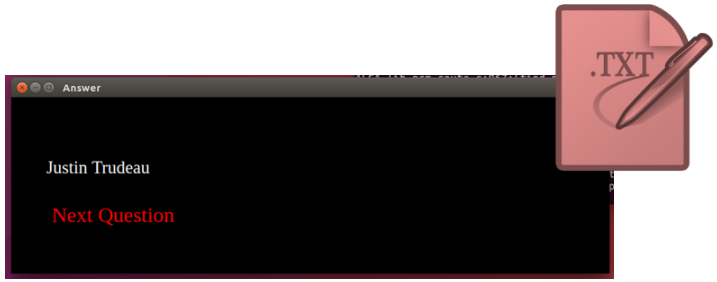
QuickTalk plugin



The Realistic Avatar Interface


Segment Render Release Time
Release
Render time
Play time
1st Segment
2nd Segment
3rd Segment
4th segment
Google tts
Audio
Duration
Blender render farm
Generate Utterance
Node
Edge
Graph path
Segments
Best edge transition selection
Transition smoothing
Live Interface Display

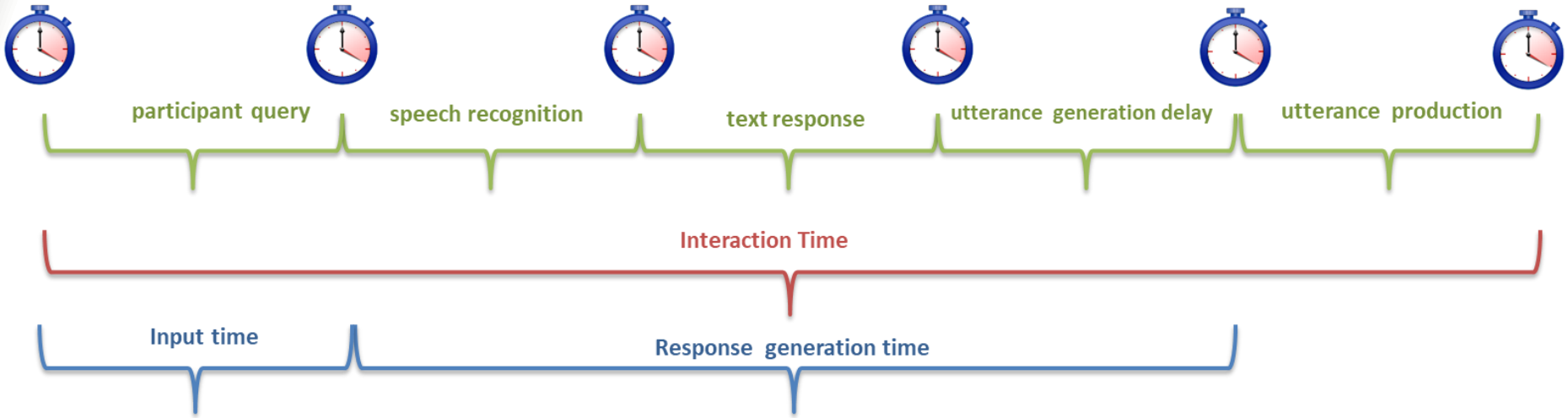

.TXT
24
AUDIO


Answer
Justin Trudeau
Next Question
.TXT

Answer
audio playing
AUDIO

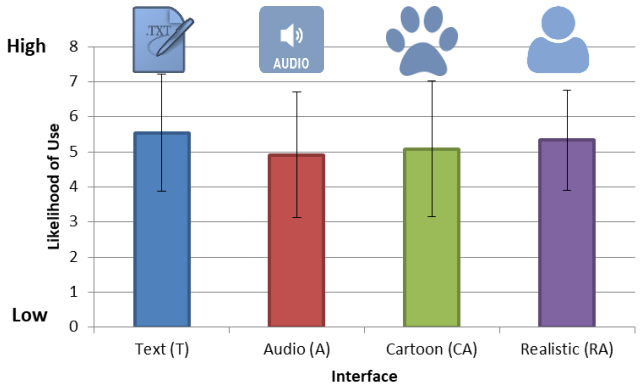
VLC media player

VLC media player

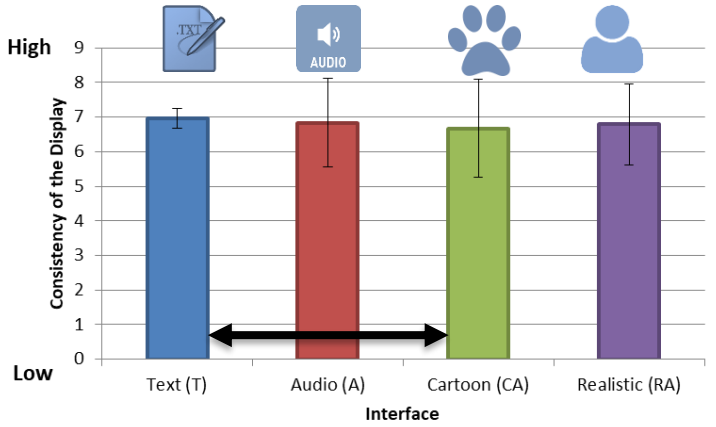

participant query
speech recognition
text response
utterance generation delay
utterance production
Interaction Time
Input time
Response generation time

FAILURE

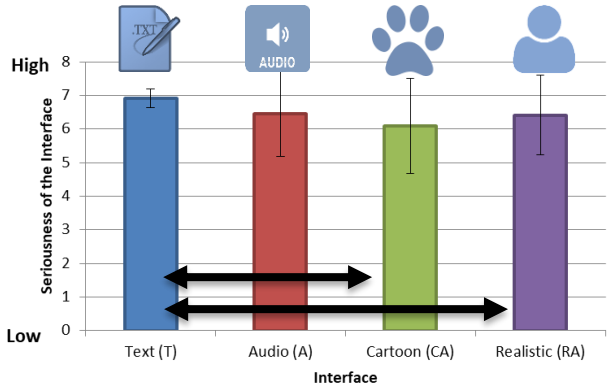
Category	Quality metric
Functionality	• Executes requested tasks • Accuracy of output • General ease of use
Humanity	• Convincing • satisfying • natural interaction
Affect	• Make tasks more interesting and fun
Ethics and behaviour	• Trustworthiness



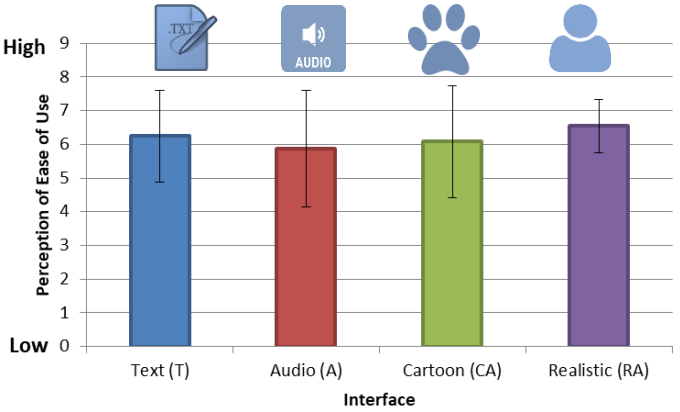
Likelihood of use in the future



Consistency of the display

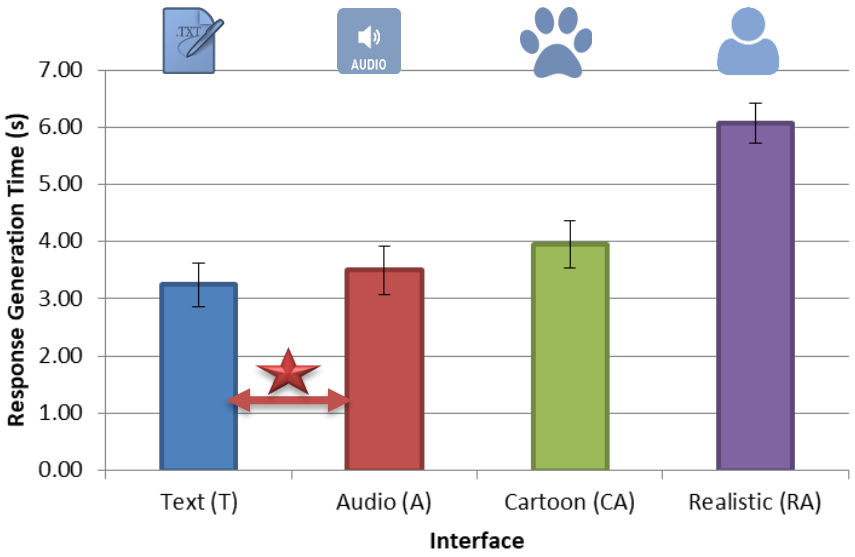


Seriousness of the interface

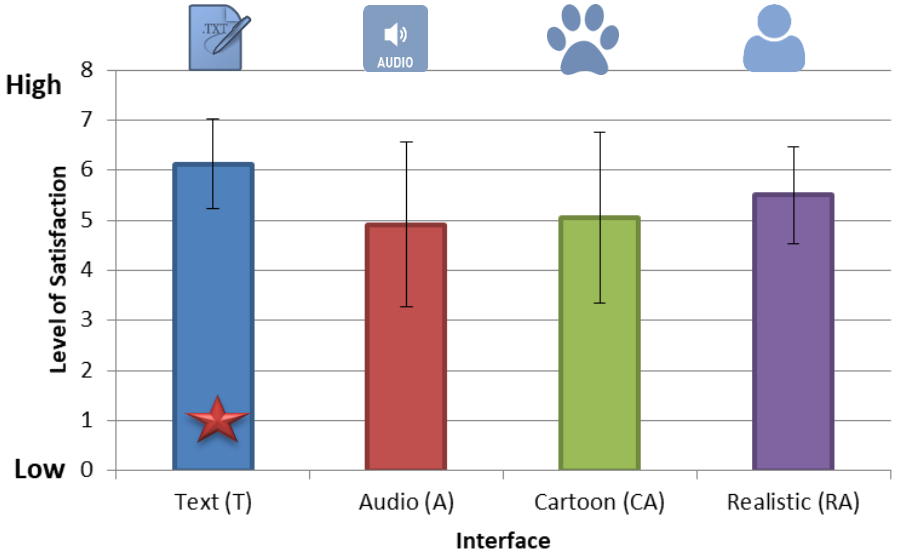


Ease of use

Error bars : SD± 1

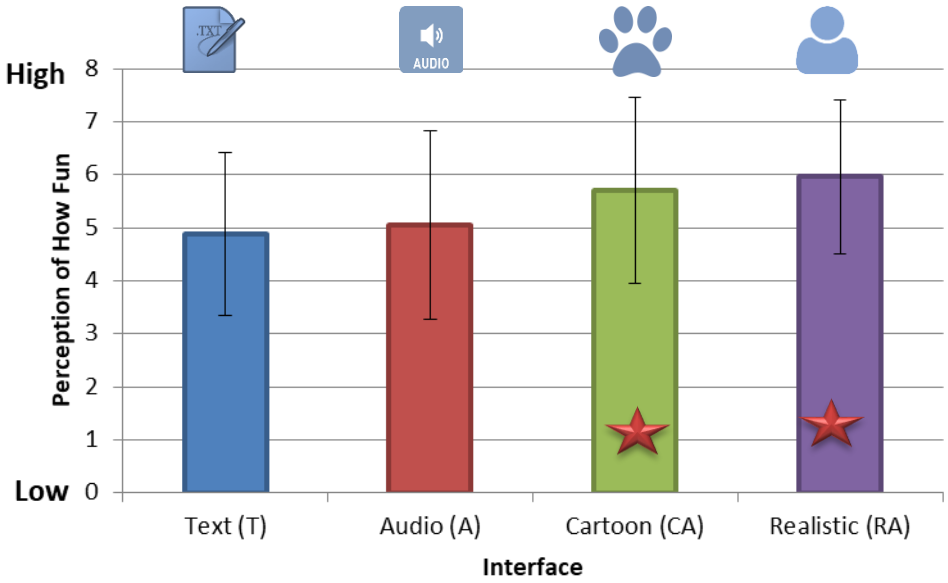


Response time by interface



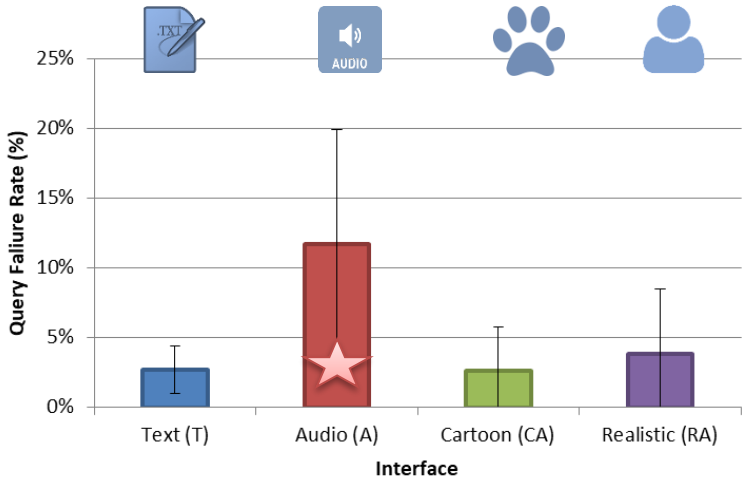
User satisfaction with response time

Error bars : SD± 1

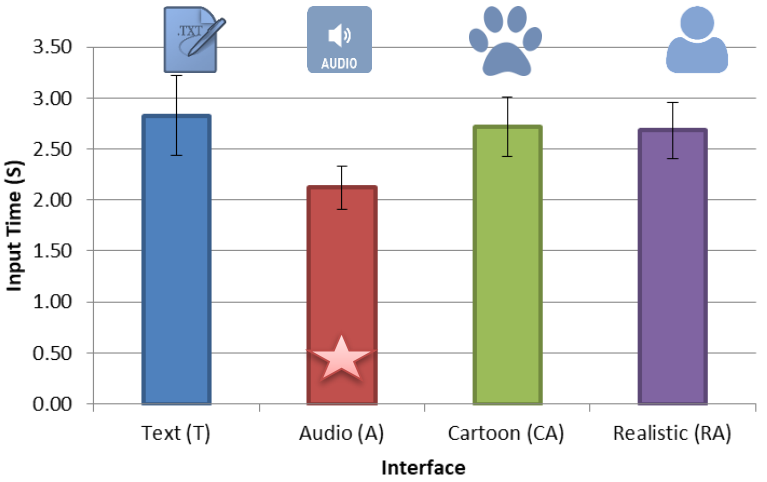


How fun the interface is

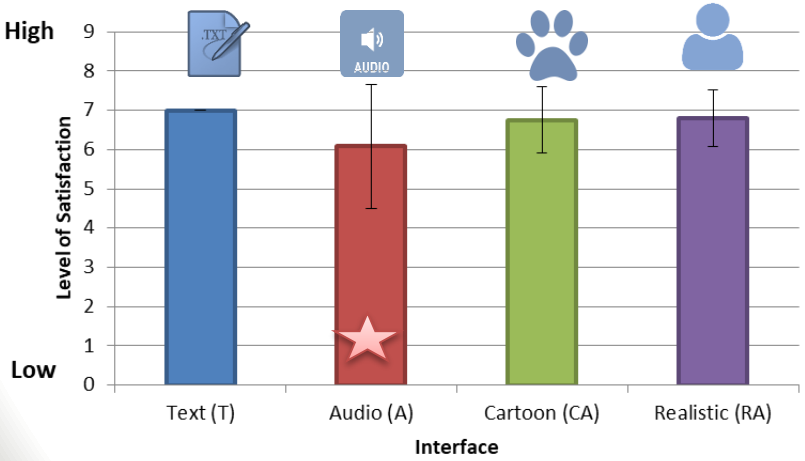
Error bars : SD± 1



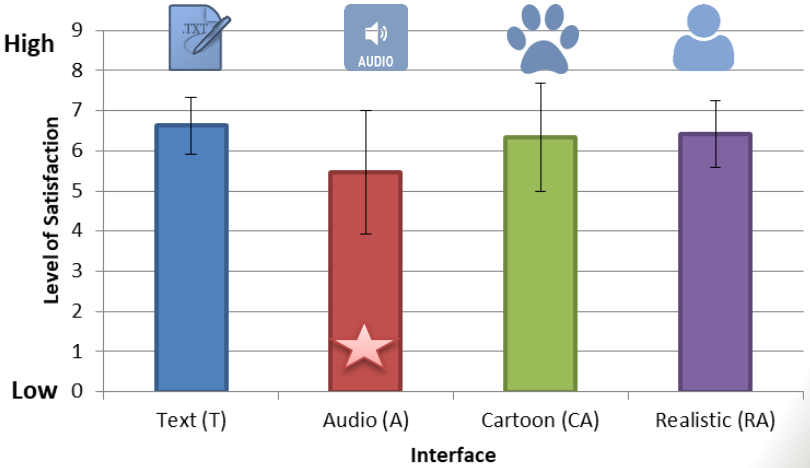
Query failure rate



Input time

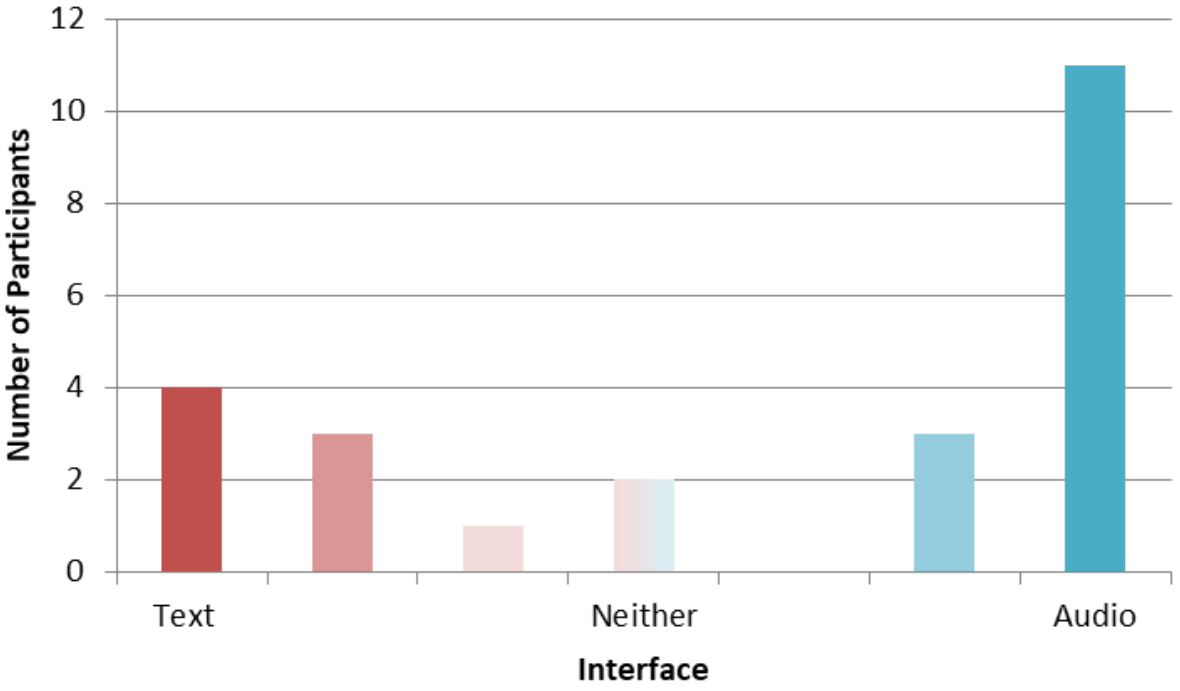


User perceived accuracy

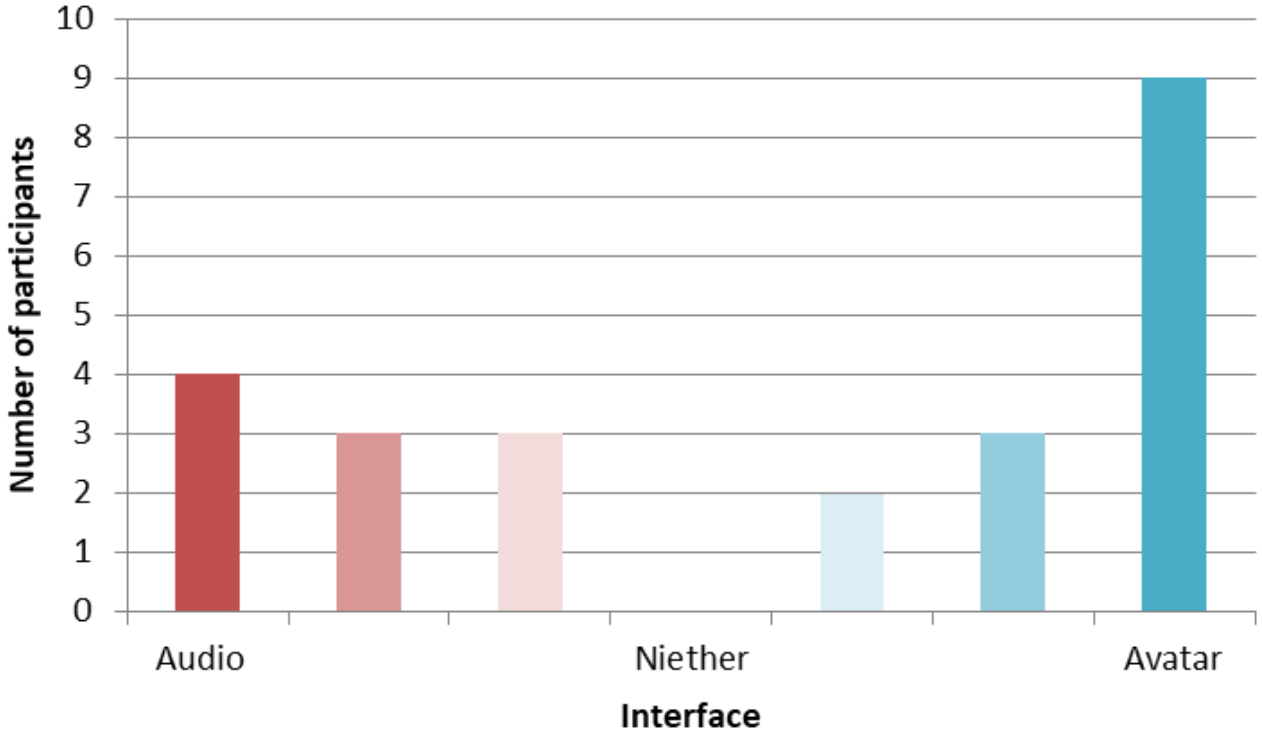


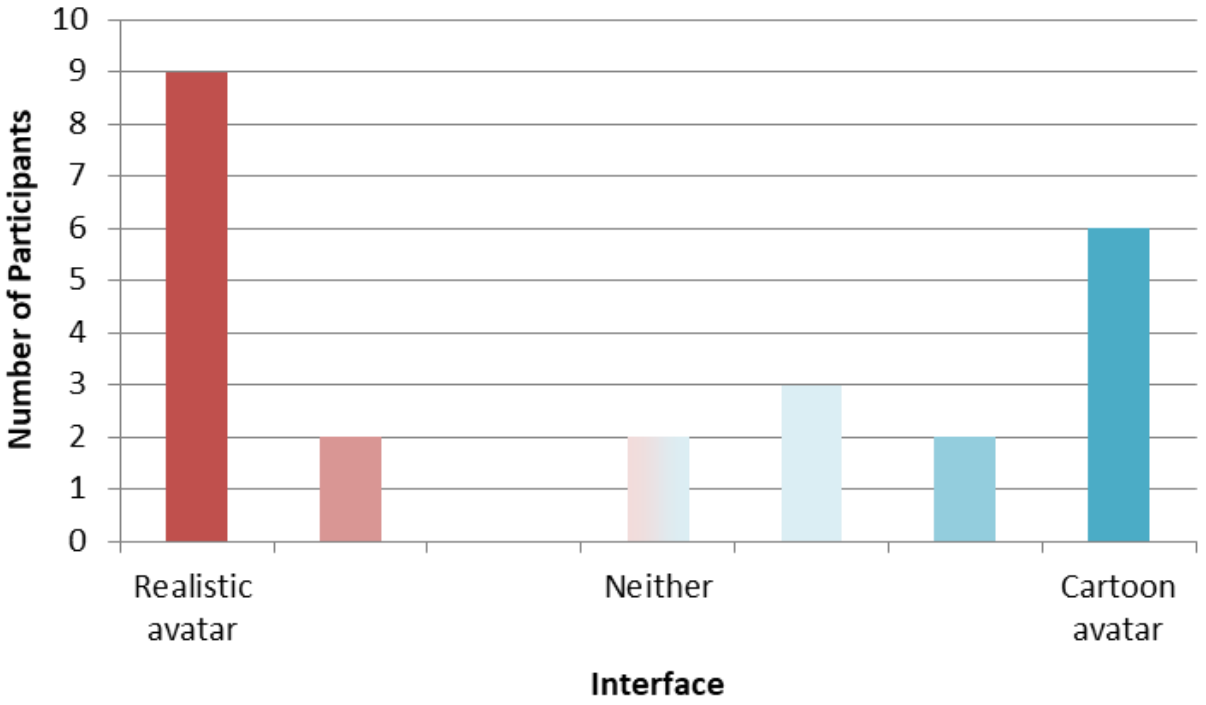
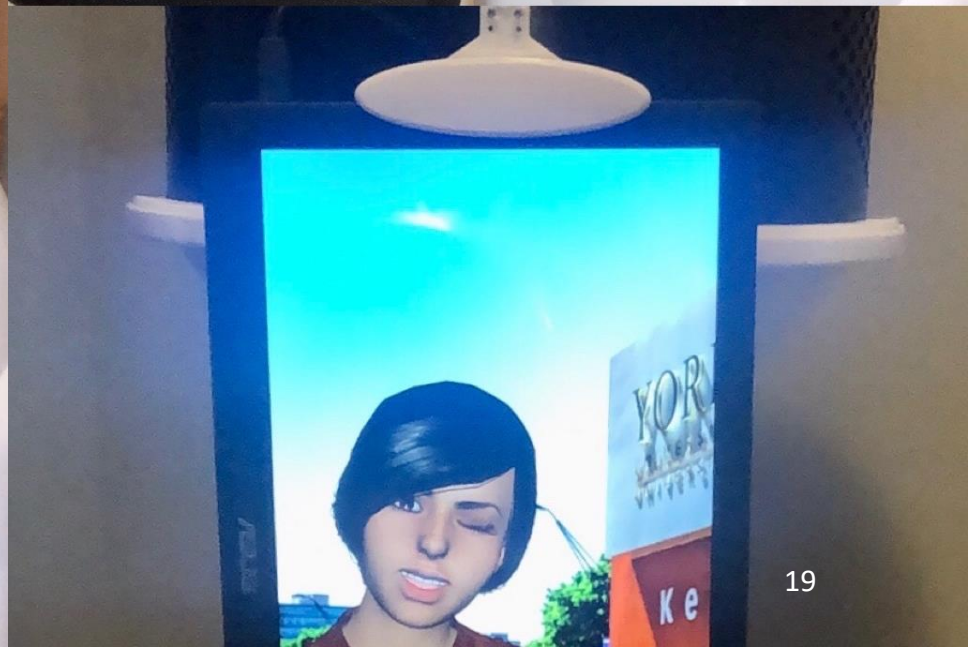
User satisfaction with response

Error bars : SD± 1


12
10
8
6
4
2
0
Number of Participants
Text
Neither
Audio
Interface

AUDIO


10
9
8
7
6
5
4
3
2
1
0
Number of participants
Audio
Niether
Avatar
Interface


10
9
8
7
6
5
4
3
2
1
0
Number of Participants
Realistic avatar
Neither
Cartoon avatar
Interface