

Predicting People's Opinions: Sounds Easy?

Quality of Experience Prediction for Speech and Audio Applications

Dr Andrew Hines
School of Computer Science
University College Dublin



How do you know if you have developed a good app?



William Sealy Gosset
(1876-1937)





Why is “Student” is famous and not Gosset?

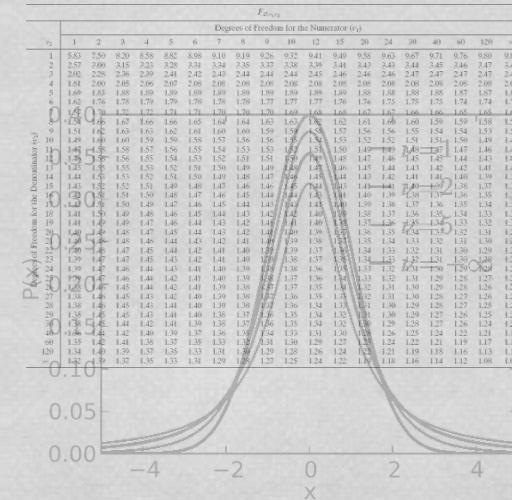
William Sealy Gosset



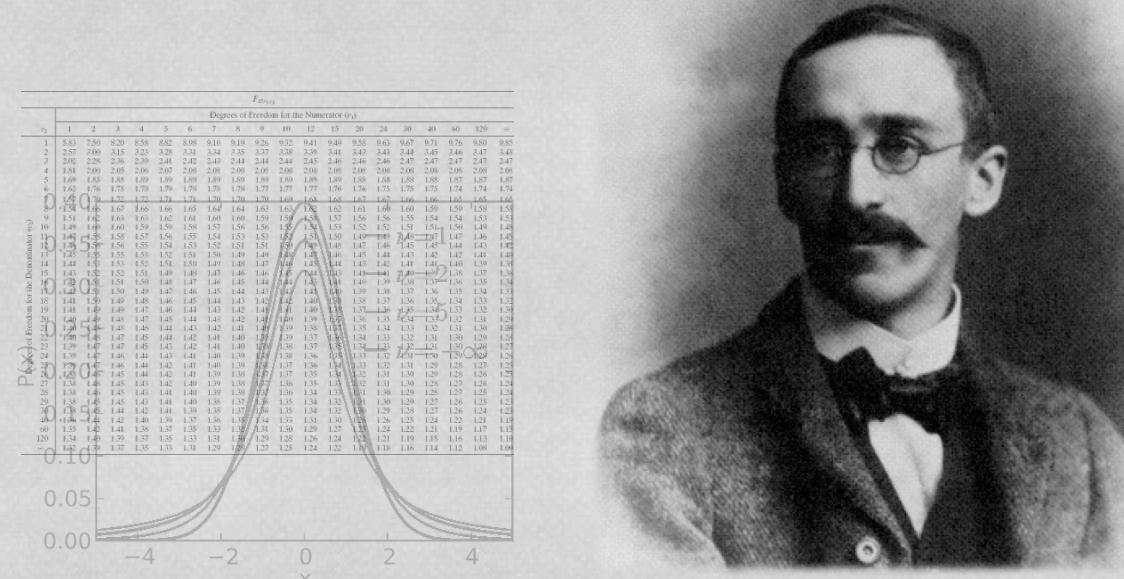
- Born 1876
- *Oldest of 5 children*
- *Father a Colonel in Royal Engineers*
- *Rejected from military academy due to eyesight*
- *Winchester and New College Oxford*
- In 1908 he published two paper's under the pseudonym Student
- Student's t test
- Dealing with real data inductively
- **Use small samples to predict general outcomes**
- Difference:
 - parameter standard deviation, σ
 - sample standard deviation, s

Gosset's Story

- Moved to Dublin for work in 1899
- Company he worked for run by two families
- Revolutionised their industry by instilling a scientific approach to their business
- Like tech giants of today provided accommodation, food onsite
- Encouraged socialising together



- Company head agreed to allow publication subject to the employees name not appearing. They would merely be designated “Pupil” or “Student”
- Who was this company?







GUINNESS® Quality of Experience



Brewer Qualitative Measures

“rub” of hops

“texture” of barley: “milky” or “steely”

Gosset expressed barley value per acre

Using t-tables

Quoted probability as odds as if betting at a horse race!

Measured and Assessed

Crop variety

Farm location

Seasons

Malting assessments

Brewing results

Student (1908a). The probable error of a mean. *Biometrika* 6 1-25.

Student (1908b). Probable error of a correlation coefficient. *Biometrika*. 6 302-310

Box, J. F. (1987). Guinness, Gosset, Fisher, and Small Samples, *Statistical Science* 2(1) 45-52



What are the QoE factors for a VoIP call?



What factors would you use
to evaluate your app?

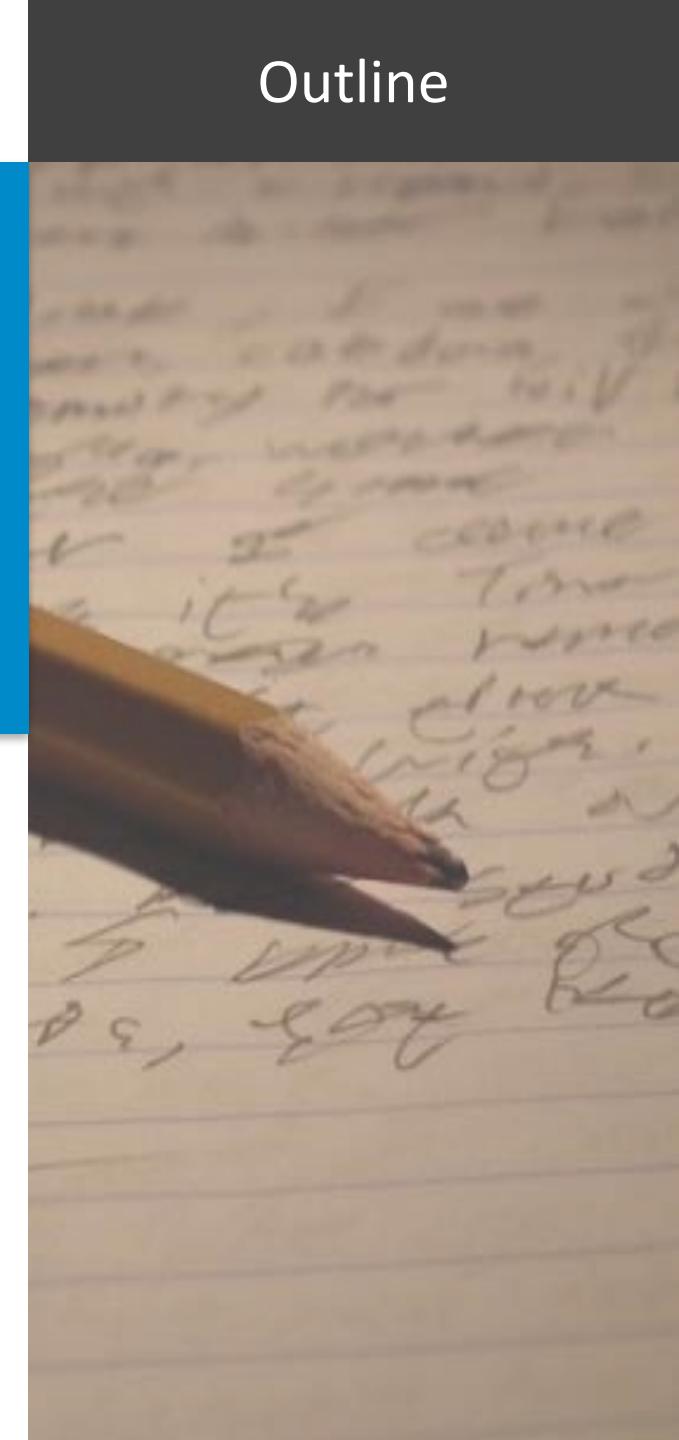
How would you measure them?

User
Experience
(UX)

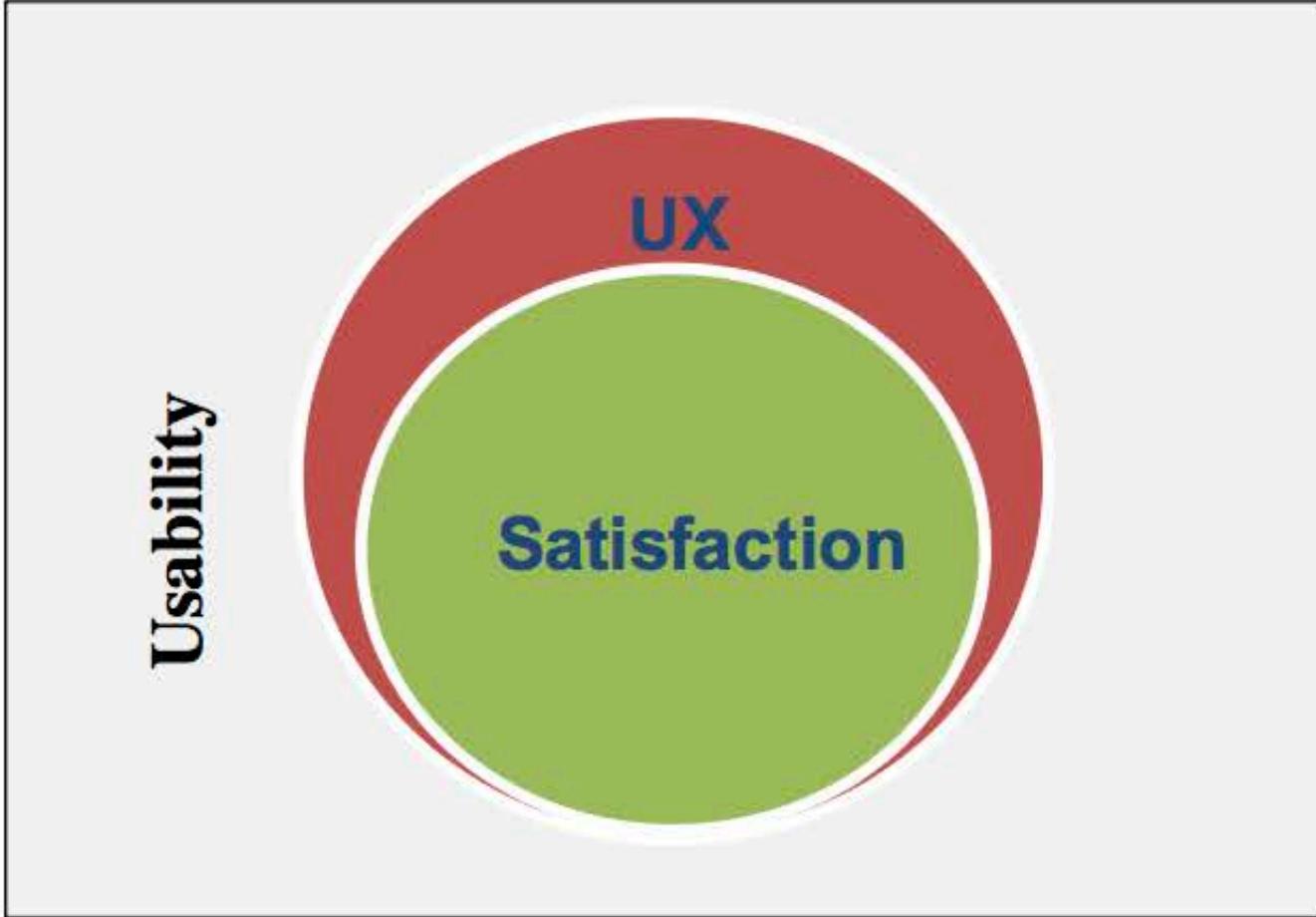
Predicting People's Opinions: Sounds Easy?

Quality of Experience Prediction for Speech and Audio Applications

User Experience
Quality of Experience
QoE Design Factors

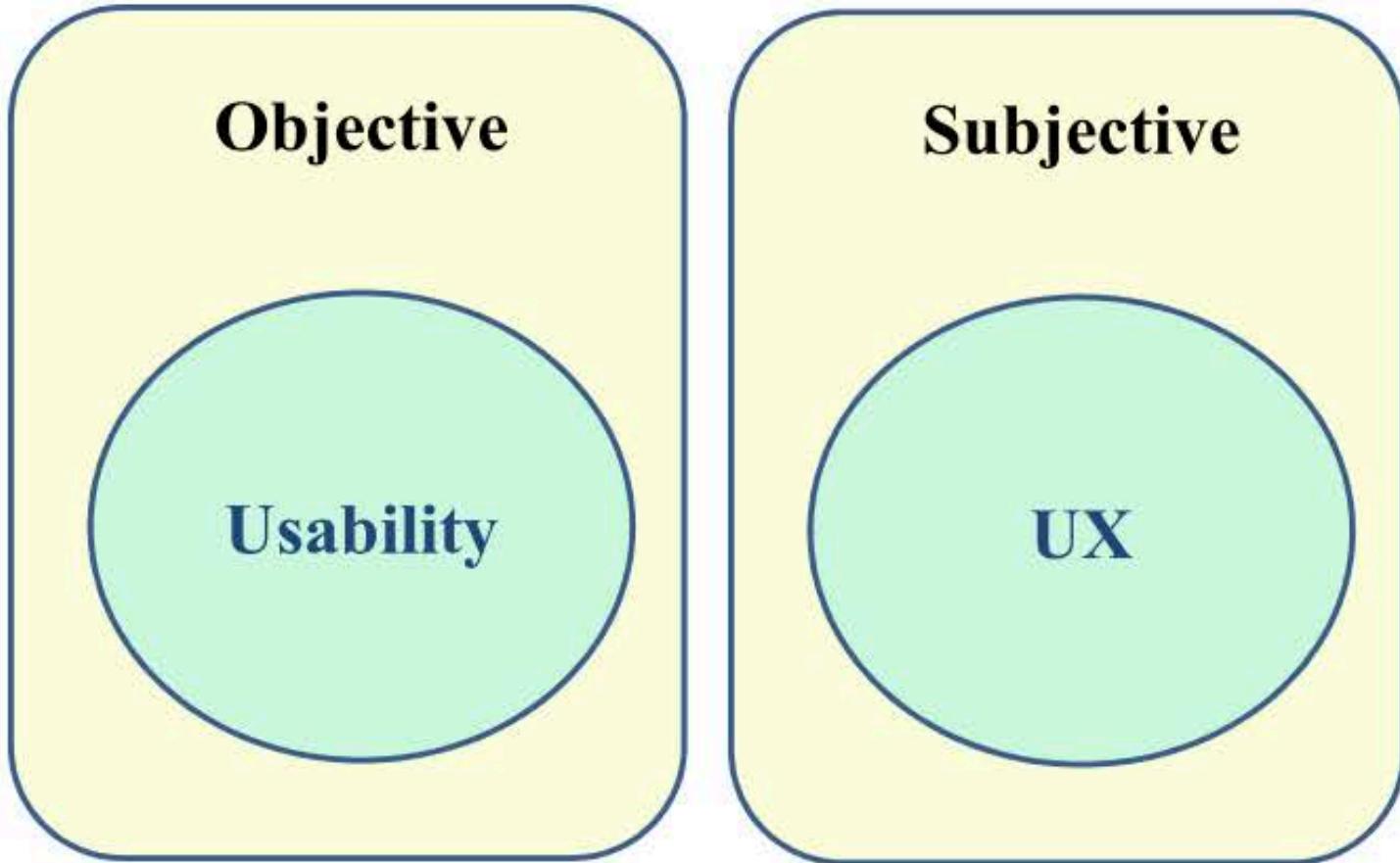


UX as a subset of usability



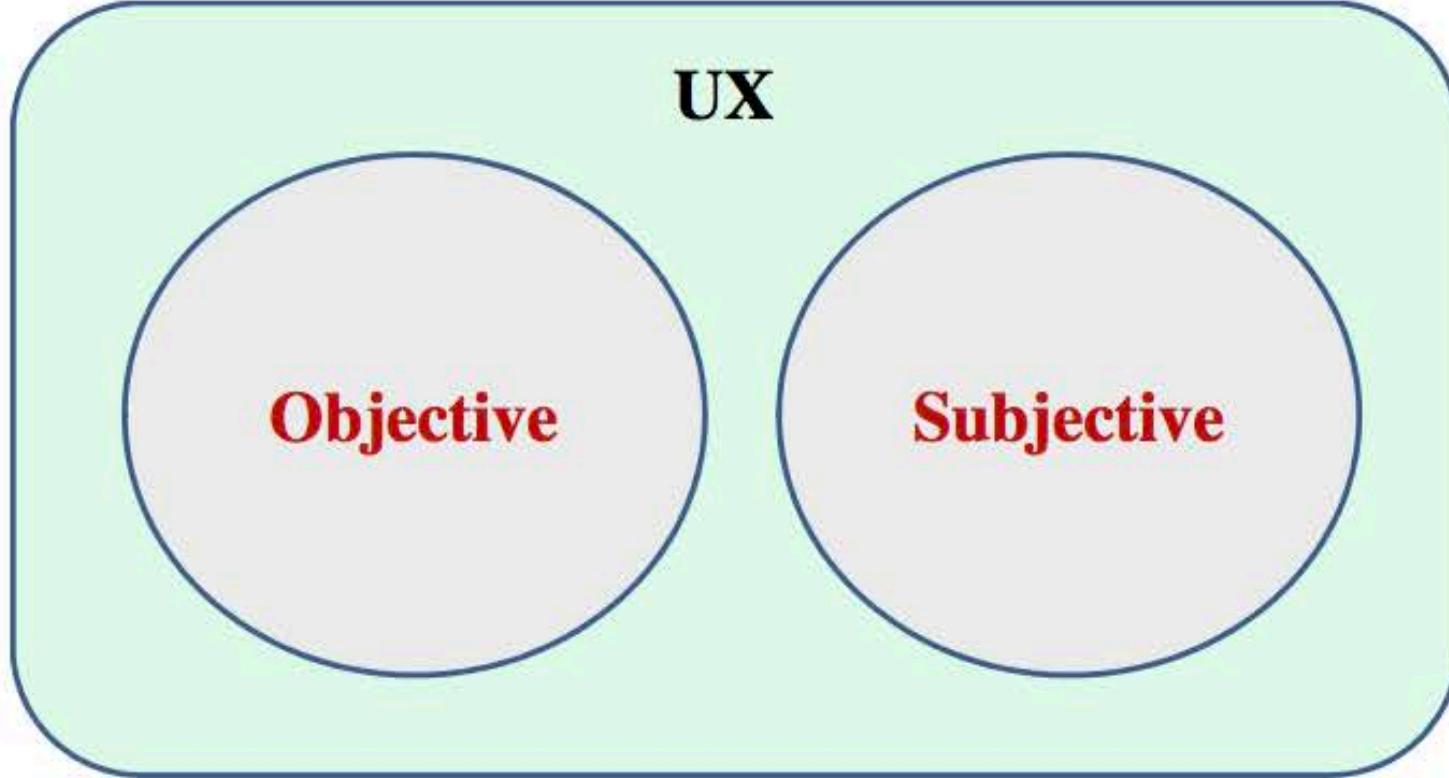
User
Experience
(UX)

UX distinct from usability: UX as subjective



User
Experience
(UX)

Umbrella term for objective and subjective (ISO 9241–210)



User
Experience
(UX)

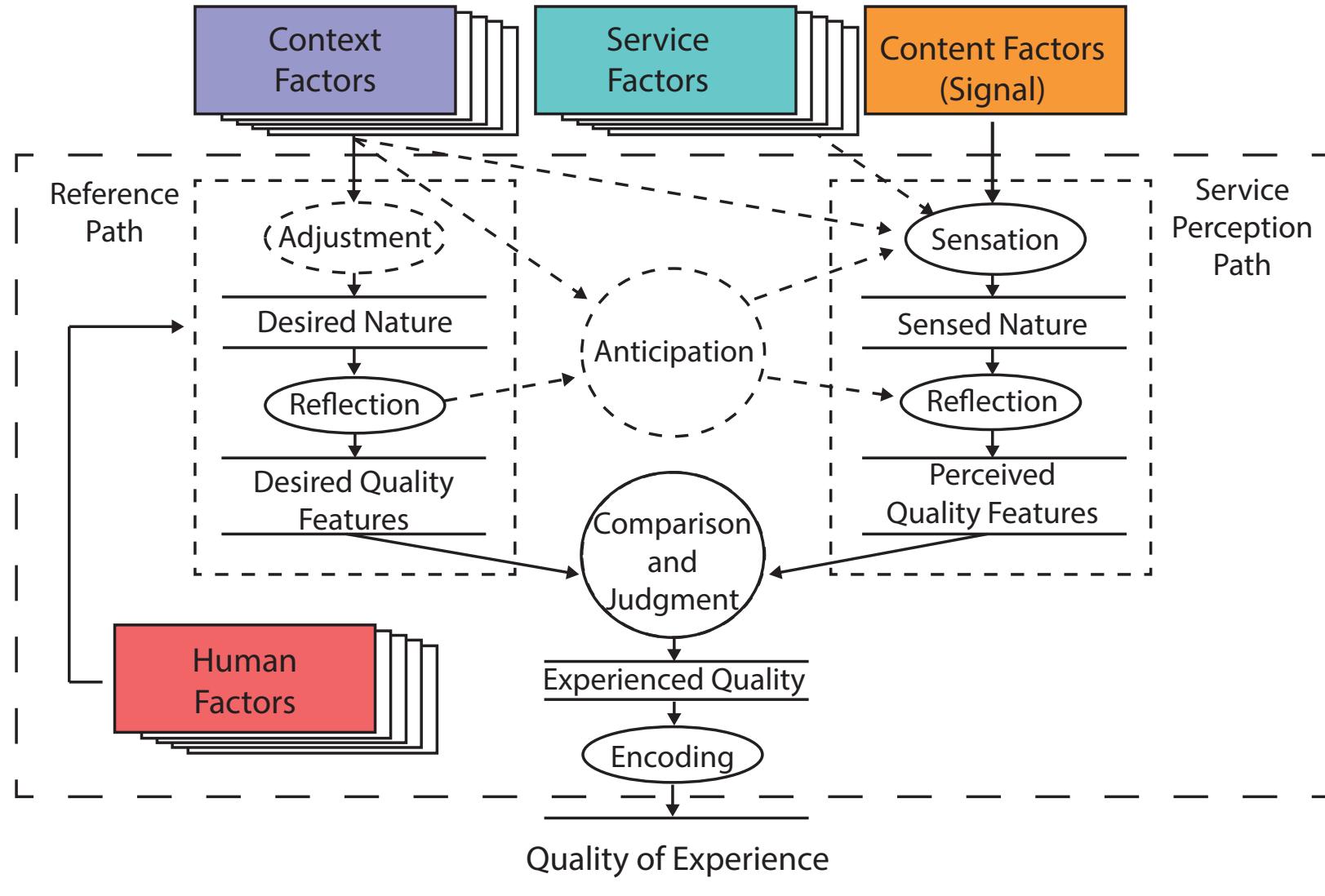
User Experience (UX)

Quality of Experience (QoE)

	QoE	UX
Origins	Telecommunications	Human-Computer Interaction (HCI)
Driving Force	Technology-centered	Human-centered
Theoretical Basis	Measurement and instrumentation approaches Historical lack of theoretical frameworks	Non-instrumental research Theoretic background in hedonic psychology
Measurement and Evaluation	Quantitative and Empirical	Qualitative and Interpretative
Experience and Perceptions	Focus on “quality formation” and perception of quality	Focus on “experience” concept

Table adapted from: Wechsung, Ina; Moor, Katrien De (2014). Möller, Sebastian; Raake, Alexander, eds. *Quality of Experience*. T-Labs Series in Telecommunication Services. Springer International Publishing. pp. 35–54.

The degree of delight or annoyance of a user of an application or service



Quality of Experience (QoE)

What factors would you use
to evaluate your app?

How would you measure them?

Quality of
Experience
(QoE)



Predicting Quality of Experience (QoE) For Audio Streaming

Network (Service)

e.g. delivery path/delay/buffering

Encoding (Service)

Bitrate and Codec

Content

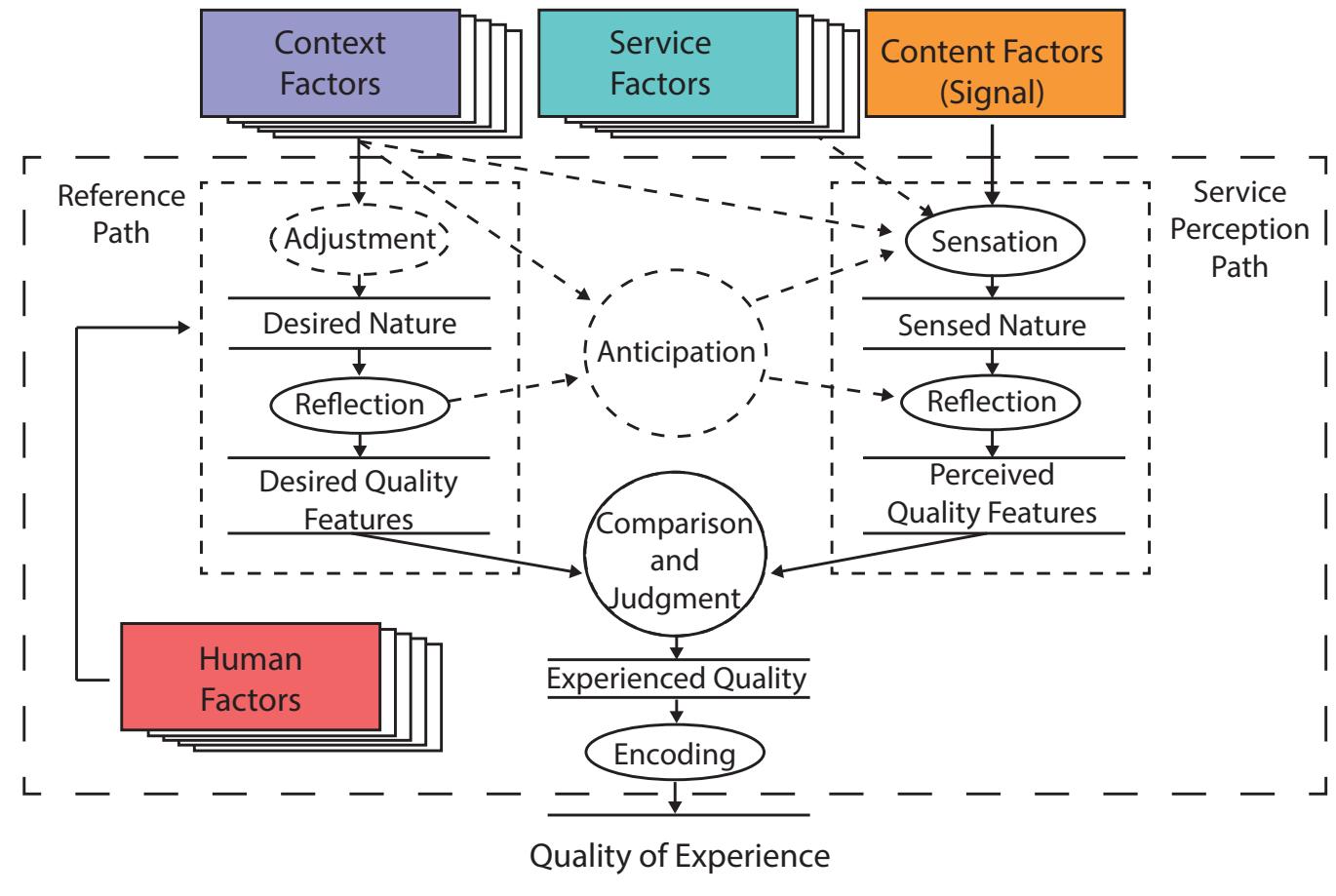
video/audio/combination

Other Factors (Human)

mood, time of day, need a coffee?

Playback Device (Context)

mobile, home theatre, headphones

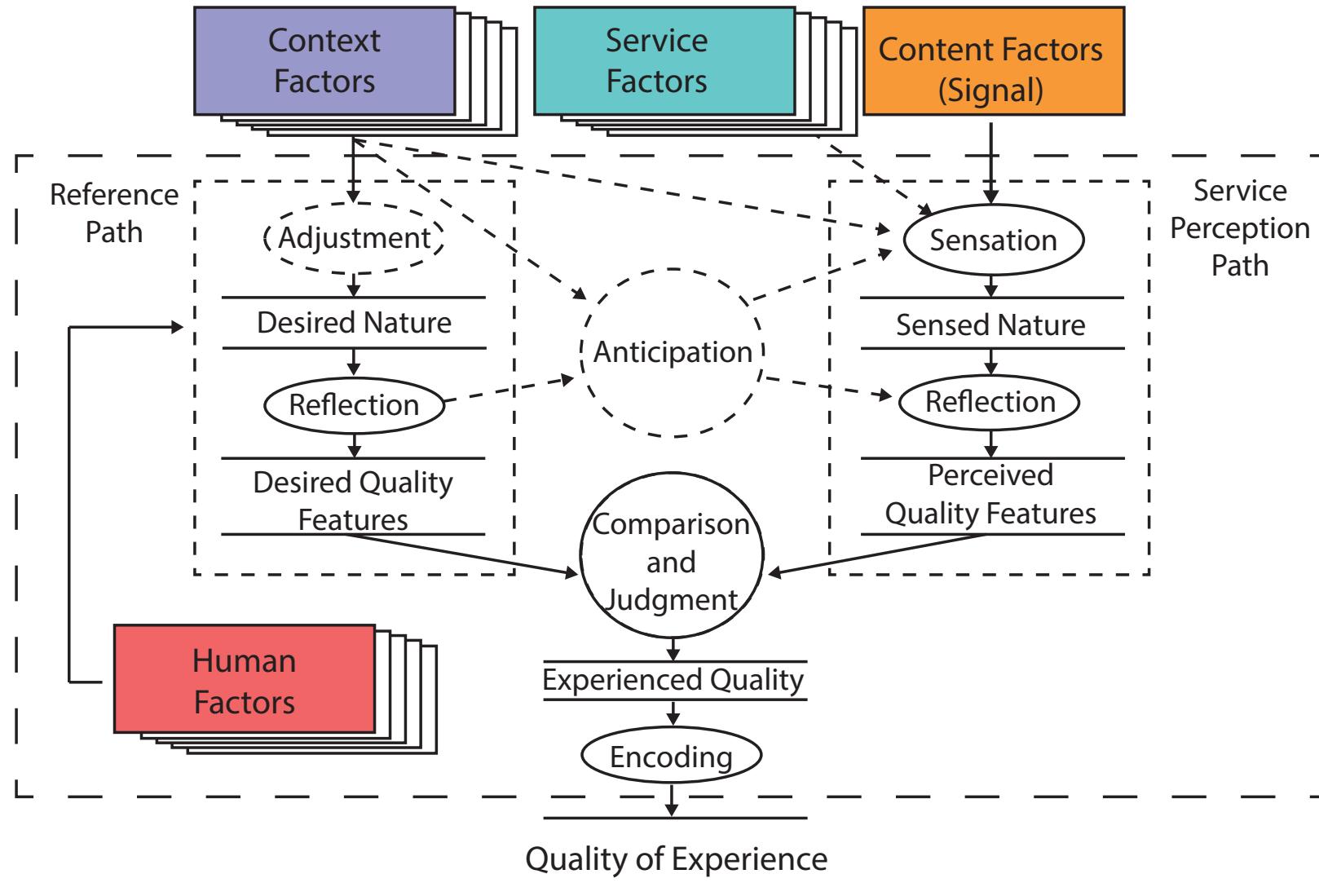




*“Horses for Courses” -
Match the Application to
the Model*

Intelligibility
vs
Fidelity

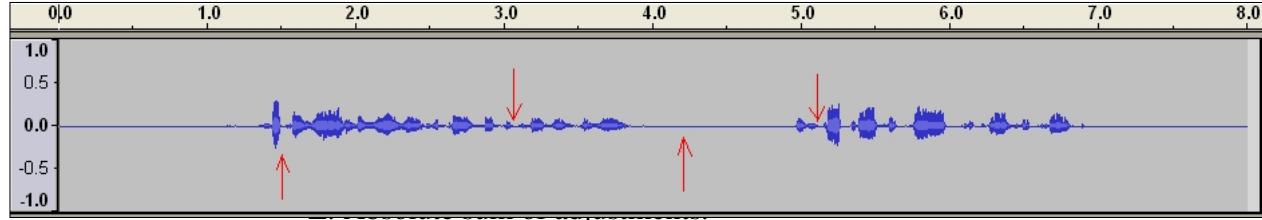
Network Service Factors



Quality of Experience (QoE)

Network Service Factors

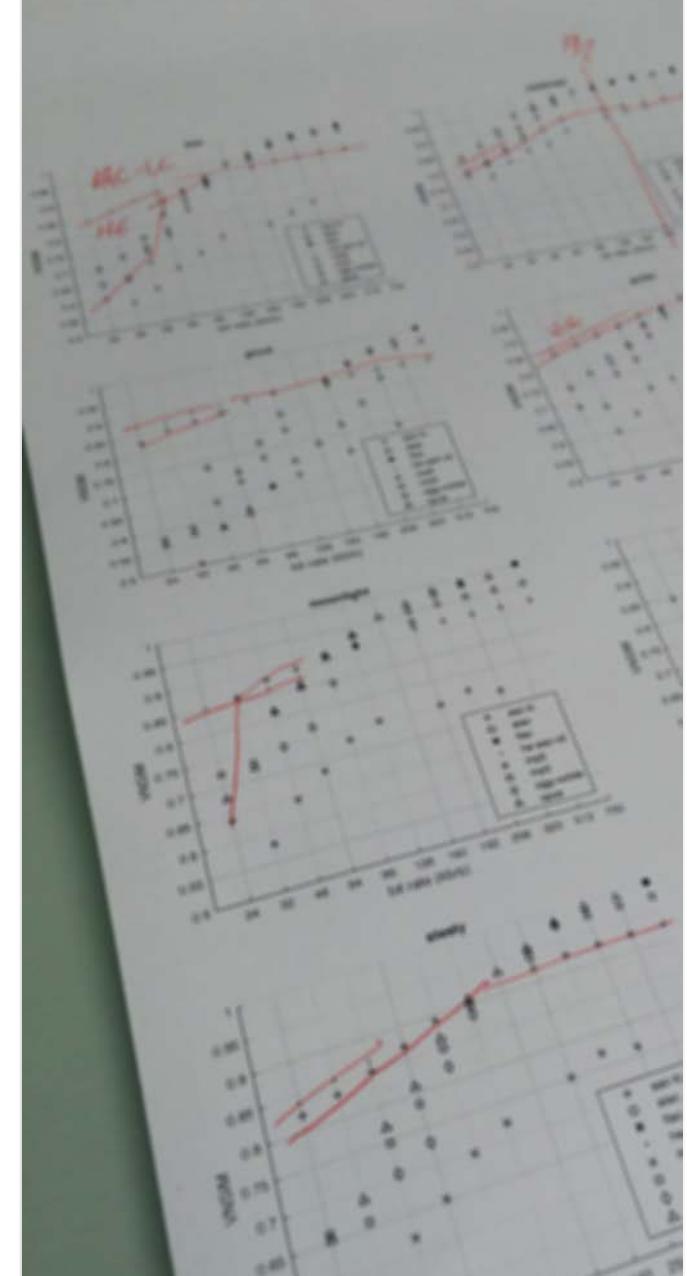
Case Study: Network Delay Adjustment Analysis



Test conditions	1st	2nd	3rd	4th	Σ
Ref	0	0	0	0	0
1	2	-2	3	-3	10
2	4	-4	-4	4	16
3	3	-3	-6	6	18
4	5	-5	-5	5	20
5	3	-6	-7	10	26
6	16	-12	-8	4	40
7	10	-17	-6	13	46
8	10	-15	-10	15	50
9	8	-23	-3	18	52
10	5	10	-30	15	60
11	-15	15	-15	15	60
12	-25	22	-8	11	66

P. Pocta, H. Melvin, [A. Hines](#). An Analysis of the Impact of Playout Delay Adjustments introduced by VoIP Jitter Buffers on Listening Speech Quality. *Acta Acustica united with Acustica*, 101(3):616–631, 2015

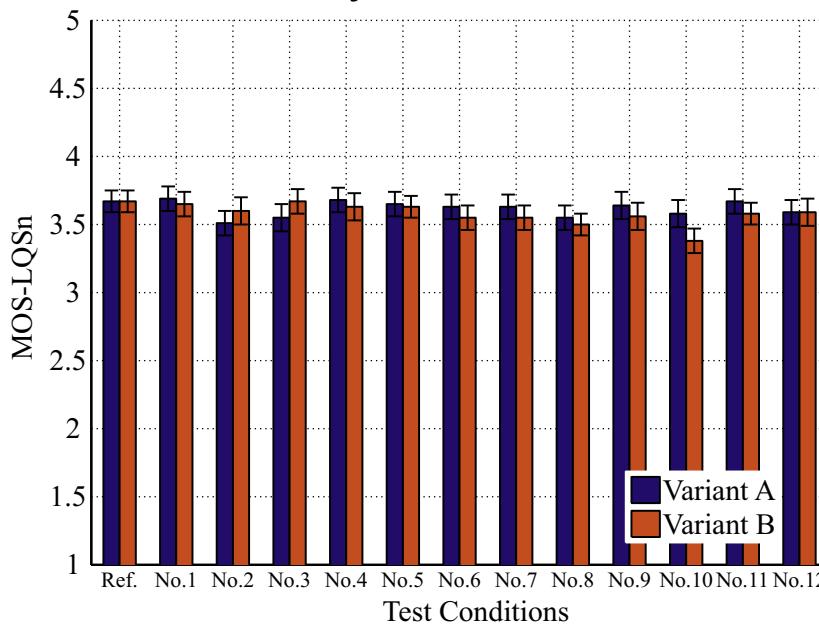
[A. Hines](#), P. Pocta, H. Melvin. Detailed Analysis of PESQ and VISQOL Behaviour in the Context of Playout Delay Adjustments Introduced by VOIP Jitter Buffer Algorithms. In *Quality of Multimedia Experience (QoMEX), Fifth International Workshop on*, Klagenfurt am Worthersee, Austria, 2013. (Winner of Best Paper Award)



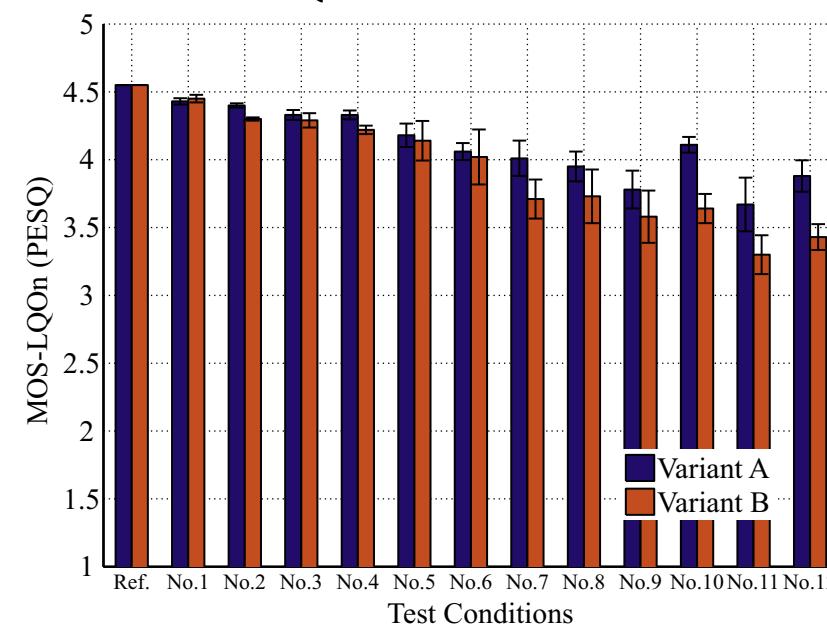
Model Assessment

Case Study: Network Delay Adjustment Analysis

Subjective Tests



PESQ Model Predictions



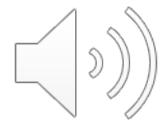
QoS
is not
QoE

- No quality impact detected by listeners
- ITU Standard model (PESQ) predicted significant quality differences
- Prediction were an **accurate QoS measurement but not a QoE measurement**

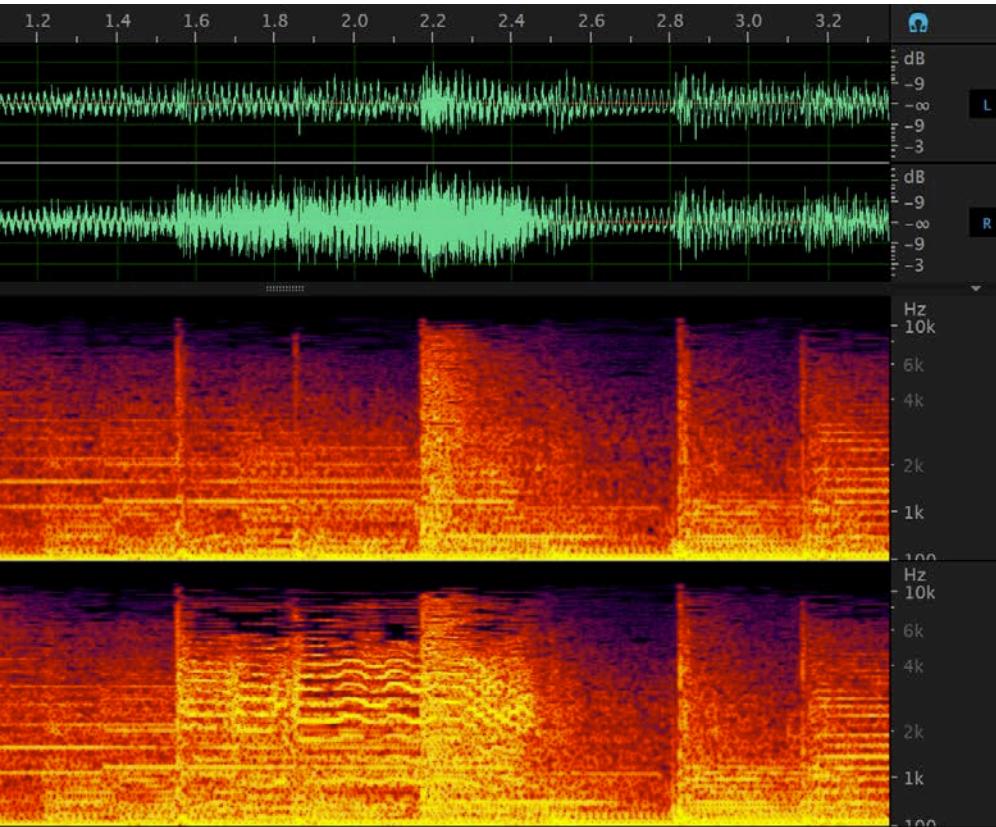
Content, Service and Context: Audio QoE for YouTube



user/premium content | transcoding service | delivery context



Content Factors



Importance of stereo

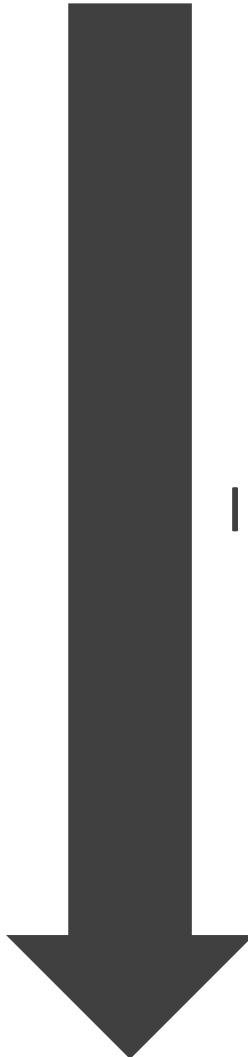
Music
Video



Home
Video



Silent
Game
Capture



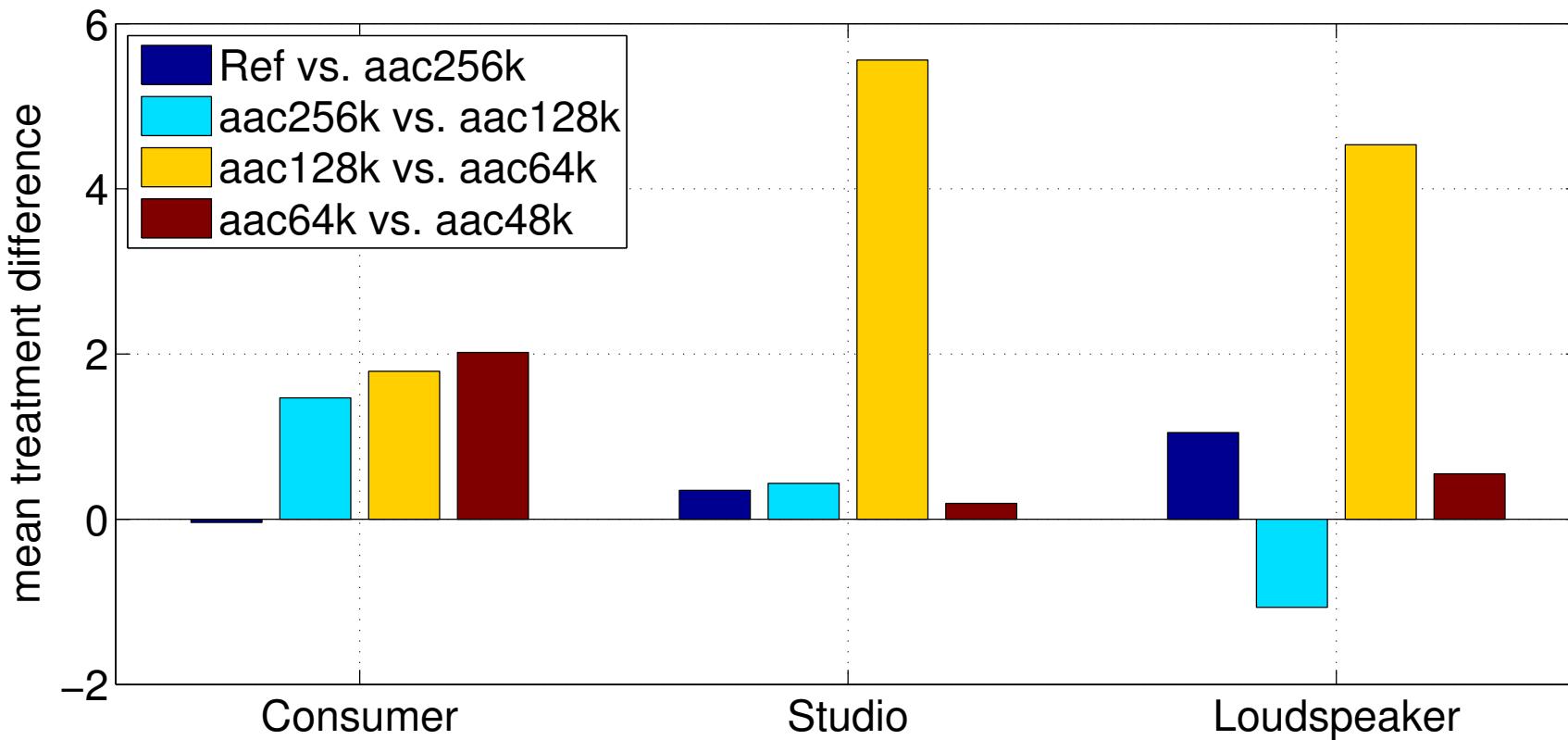
Audio
Quality
Importance

Content Factors

Service Factors

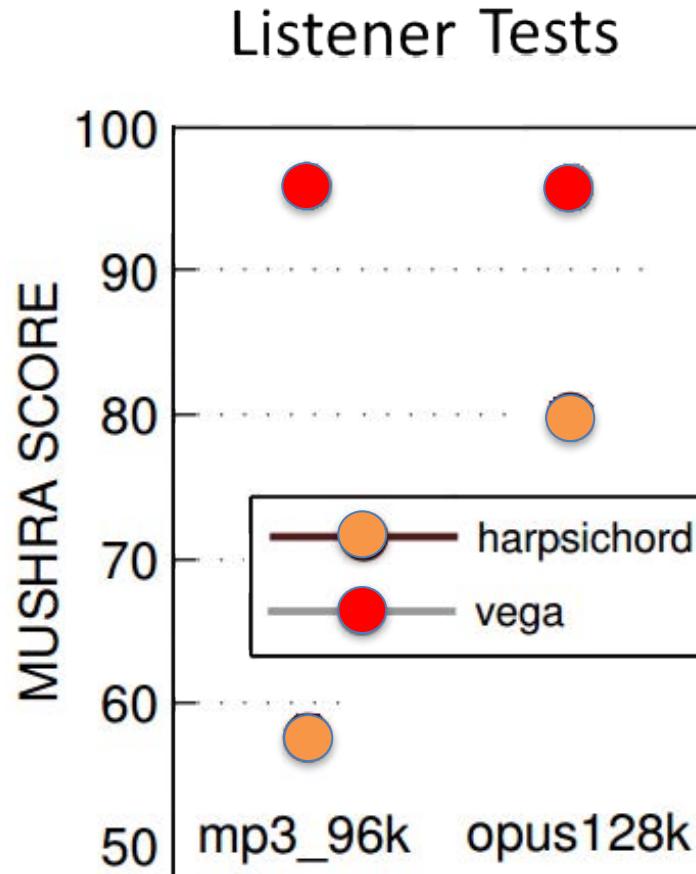
**Spend your
Bytes
Wisely**

Listening Equipment And Environment



Impact of Content on Encoding

Encoding
Quality
Varies with
Content

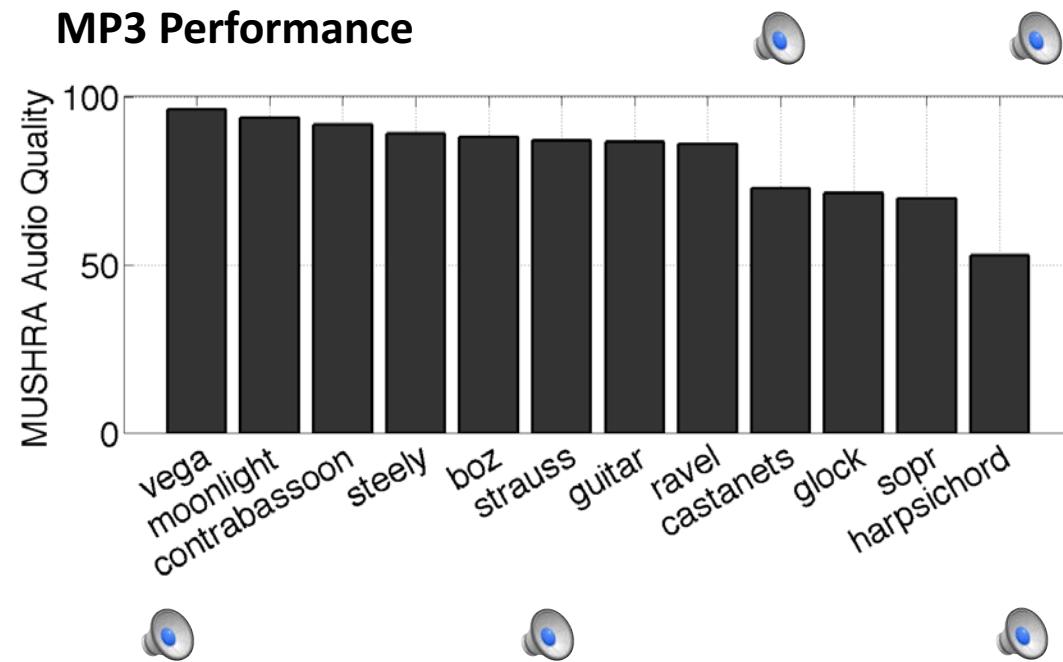


Subjective Tests

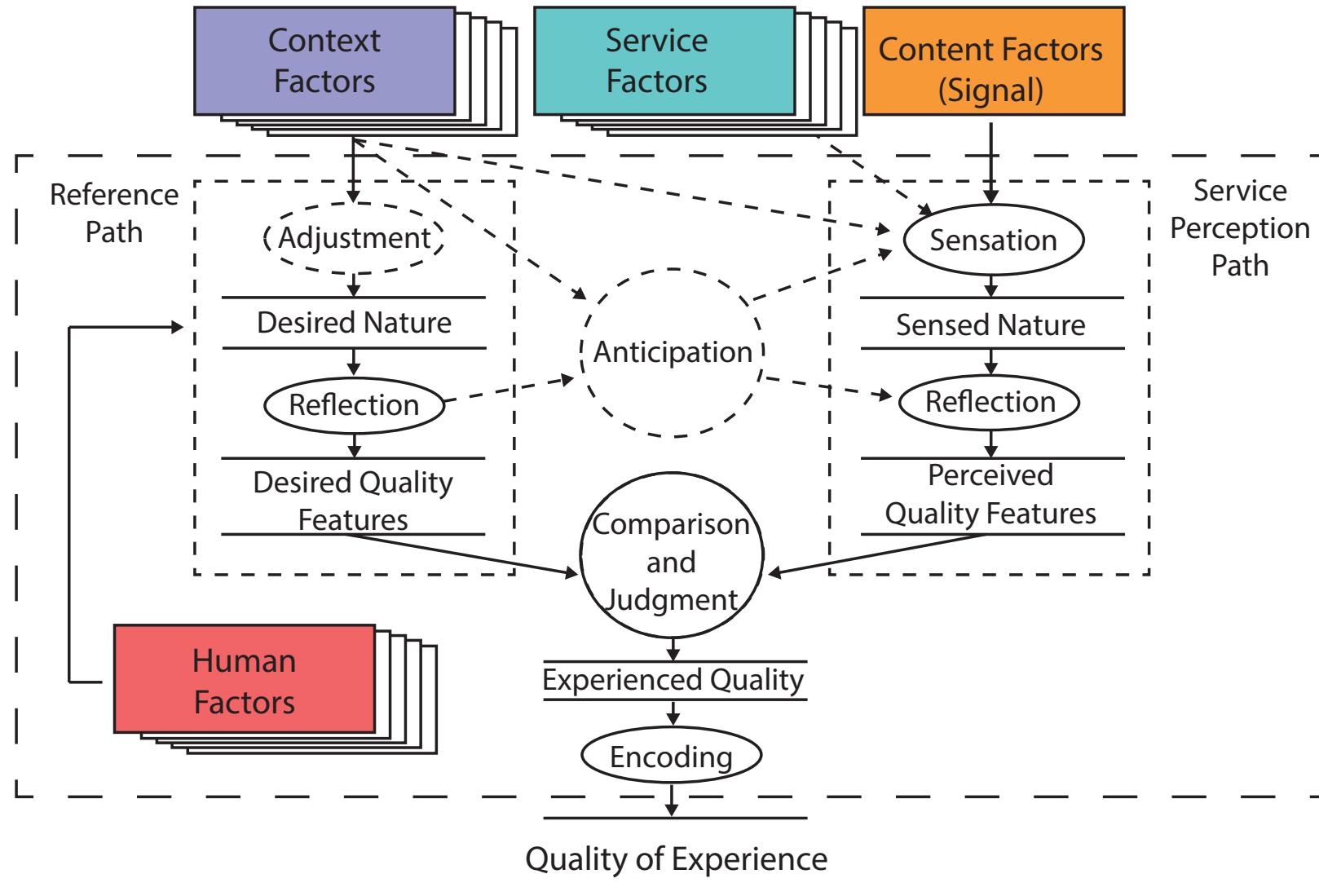
Labelling, Bias and Tools

Content Factors

Service Factors



Human Factors: Biases



Quality of Experience (QoE)

Subjective Biases

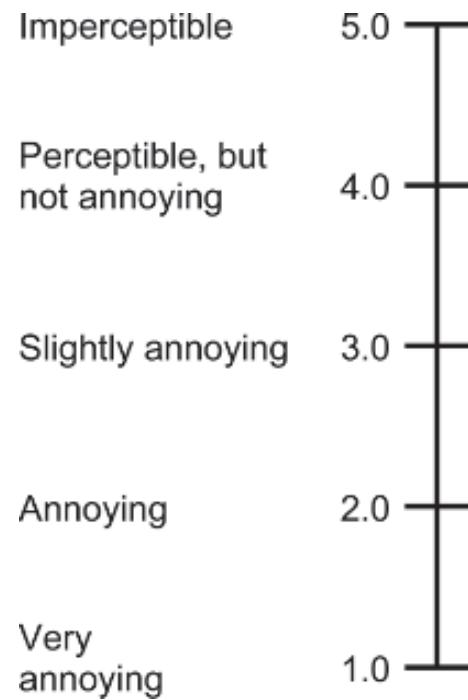


Fig. 1. ITU-R impairment scale [4], [19]

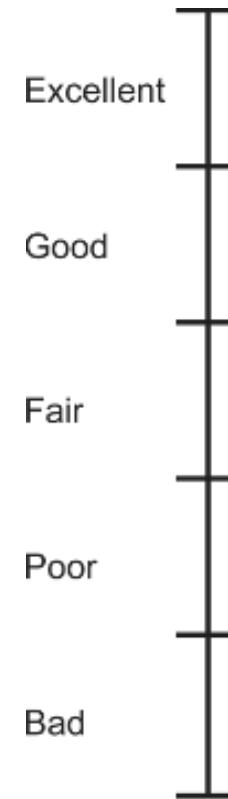
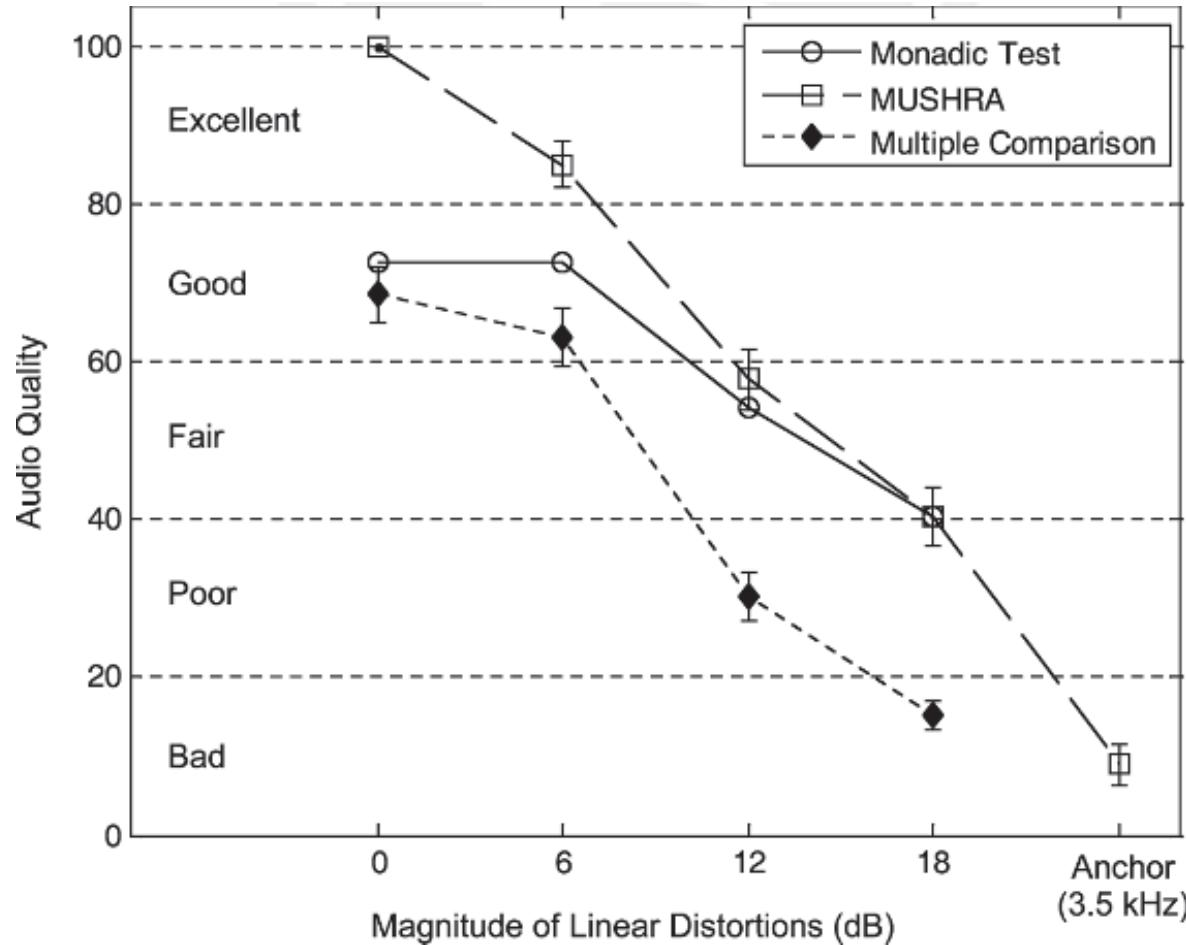


Fig. 2. ITU-R quality scale [5][19].

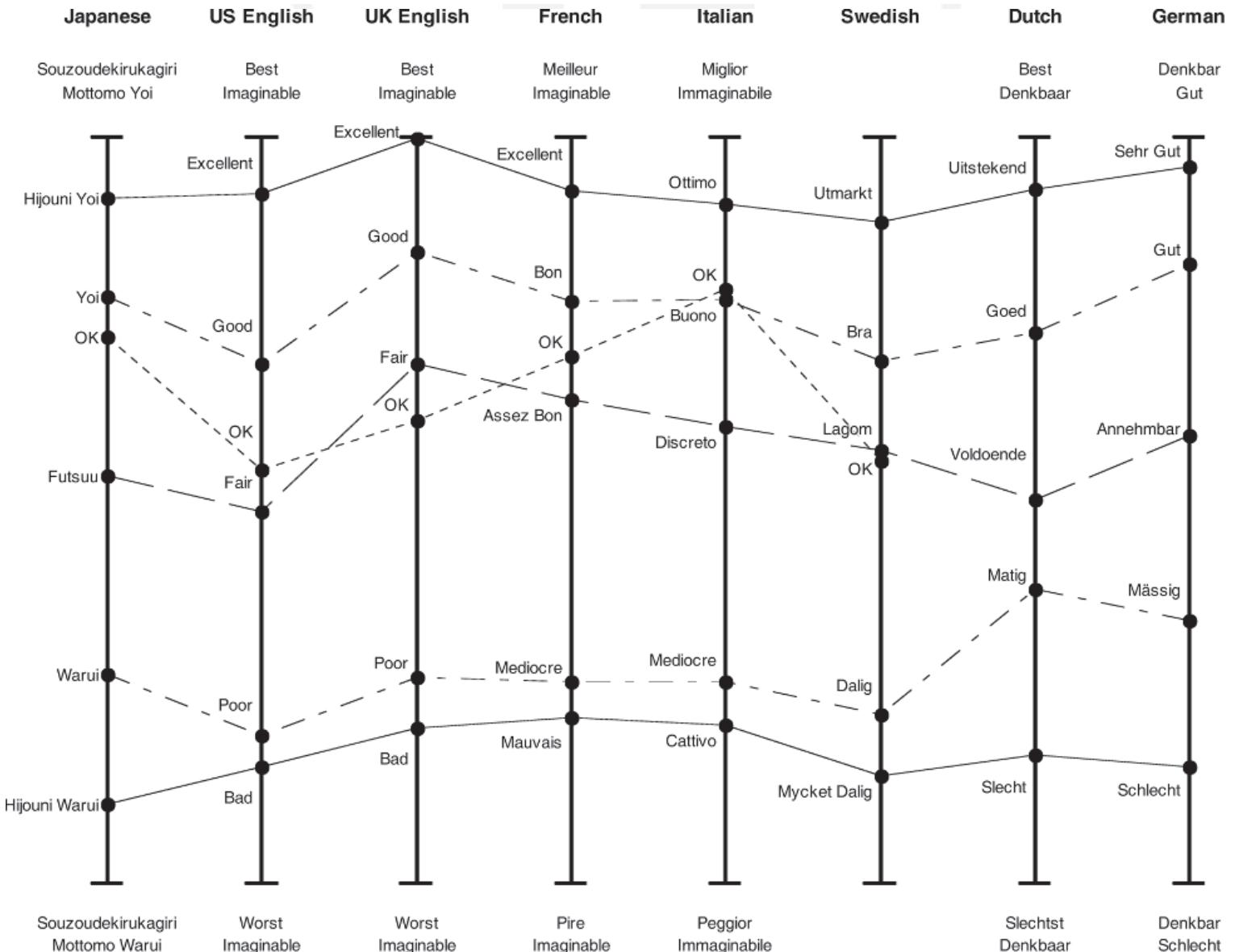
What are
the 5 points
on your
Likert
scale?

Method Bias



Your
(standard?)
methods

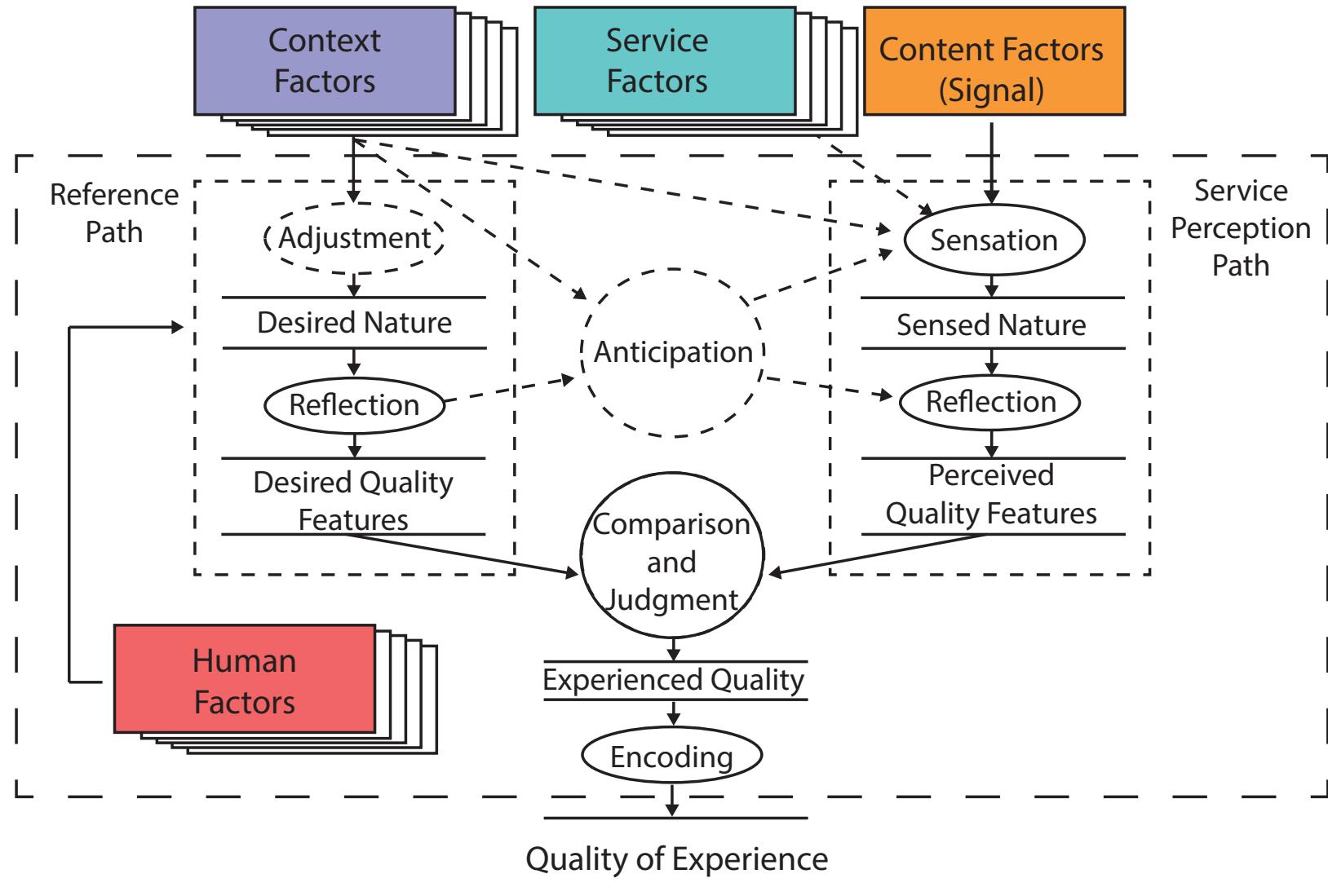
Label Bias



Your Labels

Fig. 12. Combined results of scaling of labels used in quality evaluation. (Data taken from [7], [26], [59]–[62]; see text for details.)

Putting it all together



Quality of Experience (QoE)

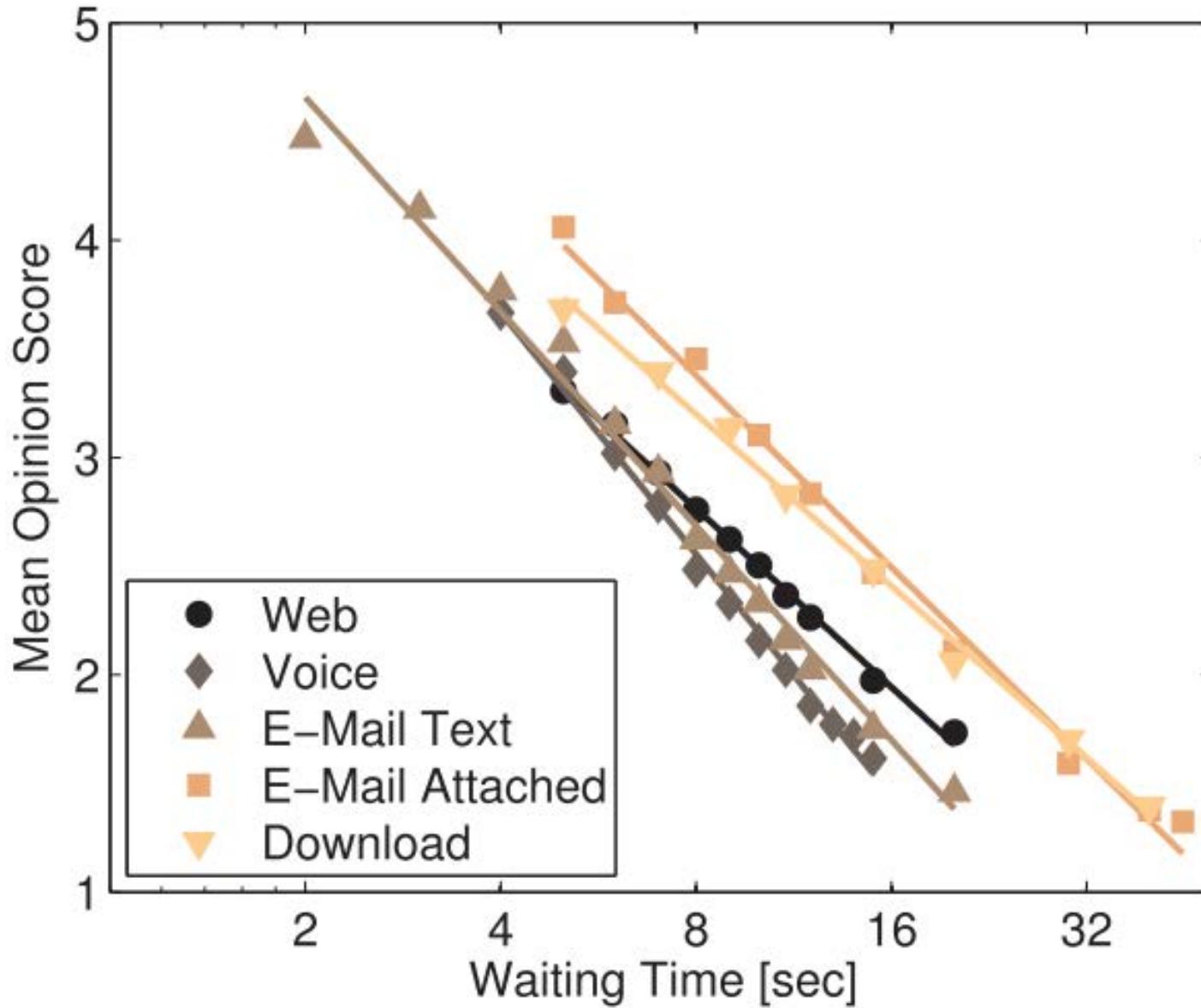
Human Factors

Service Factors

Content Factors

Context Factors

Waiting
Time



Quality of Experience (QoE)

What to measure and how to measure: as tricky as actually measuring!



Context Factors

Content Factors



“OK”

“Fair”

“Lagom”



Want to know more? Visit qxlab.ucd.ie for:

QoE machine learning models for QoE prediction
PhD and Research Assistant opportunities
Summer Opportunities