HELLENIC REPUBLIC
National and Kapodistrian University of Athens
EST. 1837

CHIGREECE CHAPTER 2025
Connect. Engage. Inspire.

Voice-based interactions in Immersive Interactive Digital Narratives: Cognitive and Affective Dimensions of the User Experience

PhD Student: **Olga Chatzifoti**, olchatz@di.uoa.gr
Supervisor: **Dr. Maria Roussou**
Department of Informatics and Telecommunications (Di),
National and Kapodistrian University of Athens (NKUA)

Definitions

Voice-based interactions : Human-computer interactions that use voice as input modality, supported by transcription and language processing technology.

Immersive : Immersion, the objective aspects provided by the technology, in this study specifically with the use of Virtual Reality head-mounted displays.

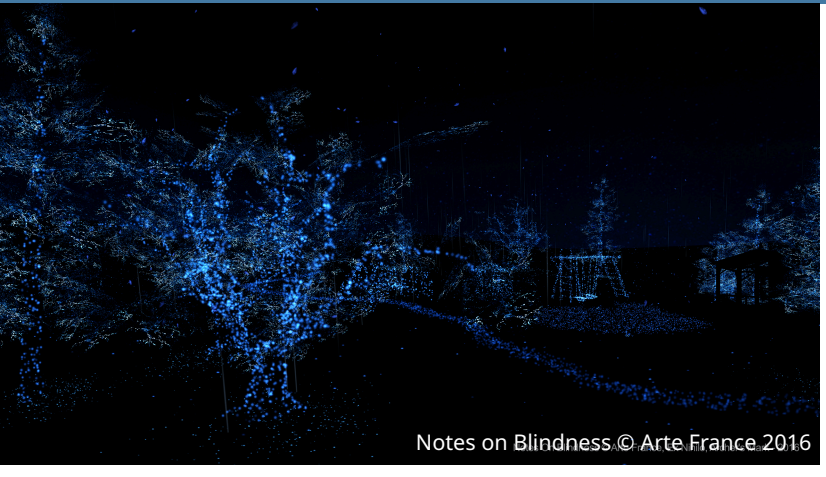
Interactive Digital Narratives : IDN as an expressive digital narrative system, shaped by participant interaction, producing instantiated outcomes.

This research focuses on the mechanics of story-telling and dialogue, which are core components of iIDNs.

Storytelling : *to verbally narrate a story*

Dialogue : *to converse with a non-playable character (NPC)*

Thought experiment



Notes On Blindness is a short VR Documentary where abstracts visualizations unfold in sync with the voice recorded diary of writer John Hull as he progressively lost his vision.

This is a story told in first person with a very powerful function for voice.

Thought experiment

- *How would the user experience be affected if instead of **passive listening**, the user could **actively use their own voice** to unfold the narrative?*

- *How would the input modality of voice affect the user's **cognitive and affective relationship with the story**, and in turn the overall immersive experience?*

Research questions

How do **voice-based interaction methods** (VBIs), when applied for **storytelling** and **dialogue** mechanics, affect the user experience with regard to...

- the user's **sense of presence, co-presence and embodiment?** ***(RQ#1)***
- the user's **emotional response to the content?** ***(RQ#2)***
- the user's **critical engagement with the content?** ***(RQ#3)***

Experiment design

Comparative user study
between vocal and silent user experiences, scoped for:

storytelling

why? less confounding variables compared to dialogue with conversation agent, more rigid experiment conditions

role-playing

why? roleplaying offers a safe expressive space without personal exposure

Linear Narrative

why? no decision making on the flow of story compared to branching narratives, less variability

READ A MONOLOGUE AS THE CHARACTER

Comparative user study

The user is invited to read a monologue page by page in three conditions:

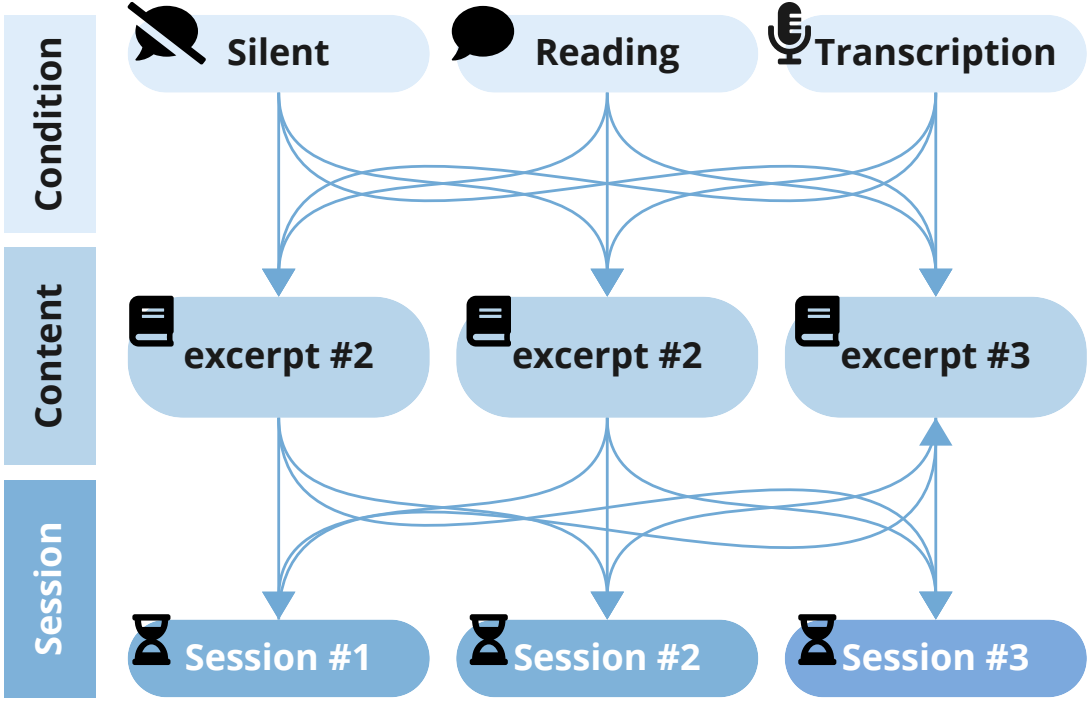
- a **silent** baseline,
- a **vocal test** with no speech technology and no system feedback,
- and a **second vocal test** with real-time speech transcription and visual feedback (highlight) of the spoken text.

			
	Condition #1	**Condition #2**	**Condition #3**
	baseline	test #1	test #2
	**Reading silently**	**Reading out loud**	**Reading out loud**
	no speech technology	no speech technology	**visual feedback of transcription**
	Press button to turn page	Press button to turn page	Press button to turn page

Experiment design

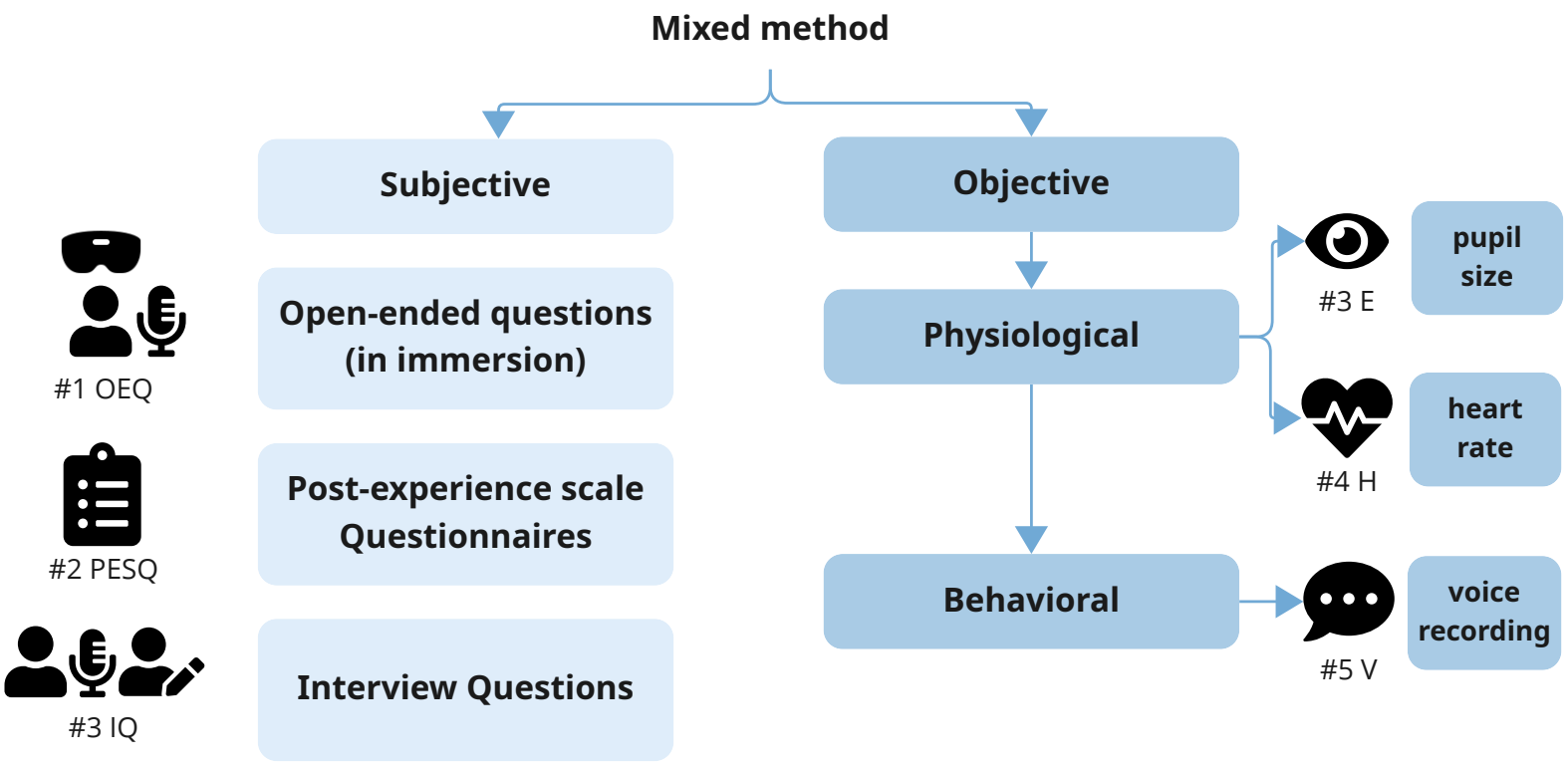
Within-subjects design to control for inter-participant differences.

Randomization against sequence bias and content familiarity bias.

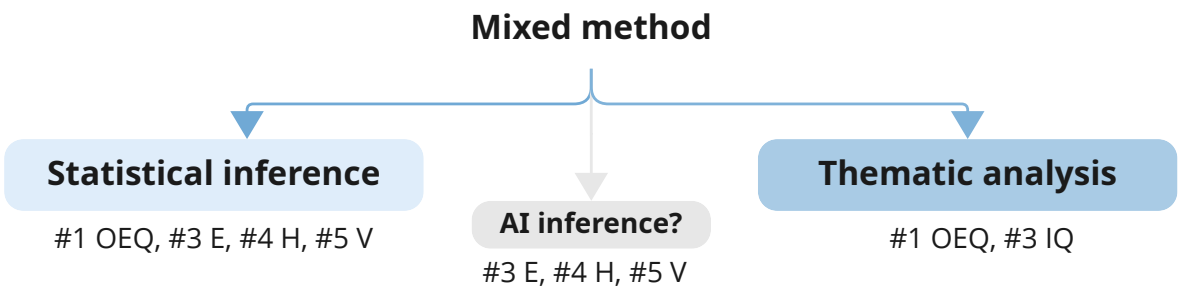


Three 5' excerpts of "The Eyes Have It" a public domain short story by Philip K. Dick written in first person (monologue with role-playing elements)

Data Gathering



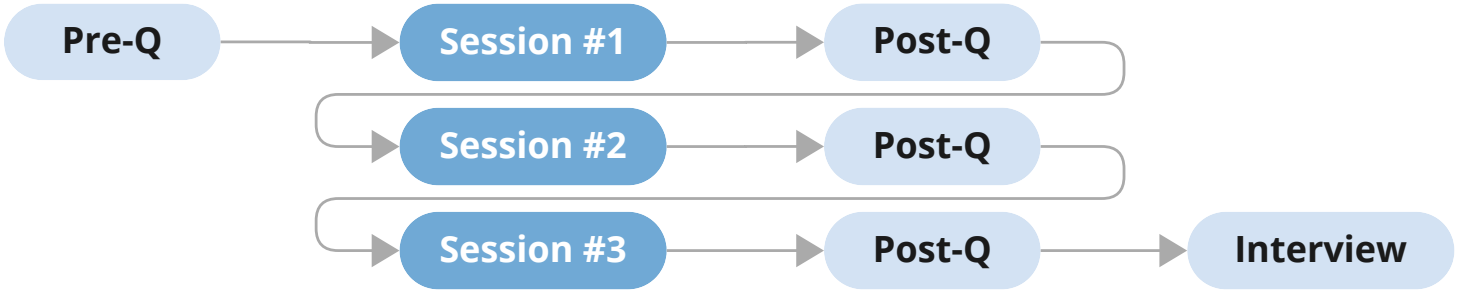
Data analysis



Relationship between RQ and data gathering methods

RQ#1	presence	**#2 PESQ** Single-item, SWAS 2
	co-presence	**#2 PESQ** Single-item
	embodiment	**#2 PESQ** Single-item
RQ#2	emotional engagement	**#1 OEQ** Q3, **#2 PESQ** SWAS 3, **#3E+#4H, #5 V**
RQ#3	critical engagement	**#1 OEQ** Q2, **#2 PESQ** SWAS 1, **#3E**

Session structure

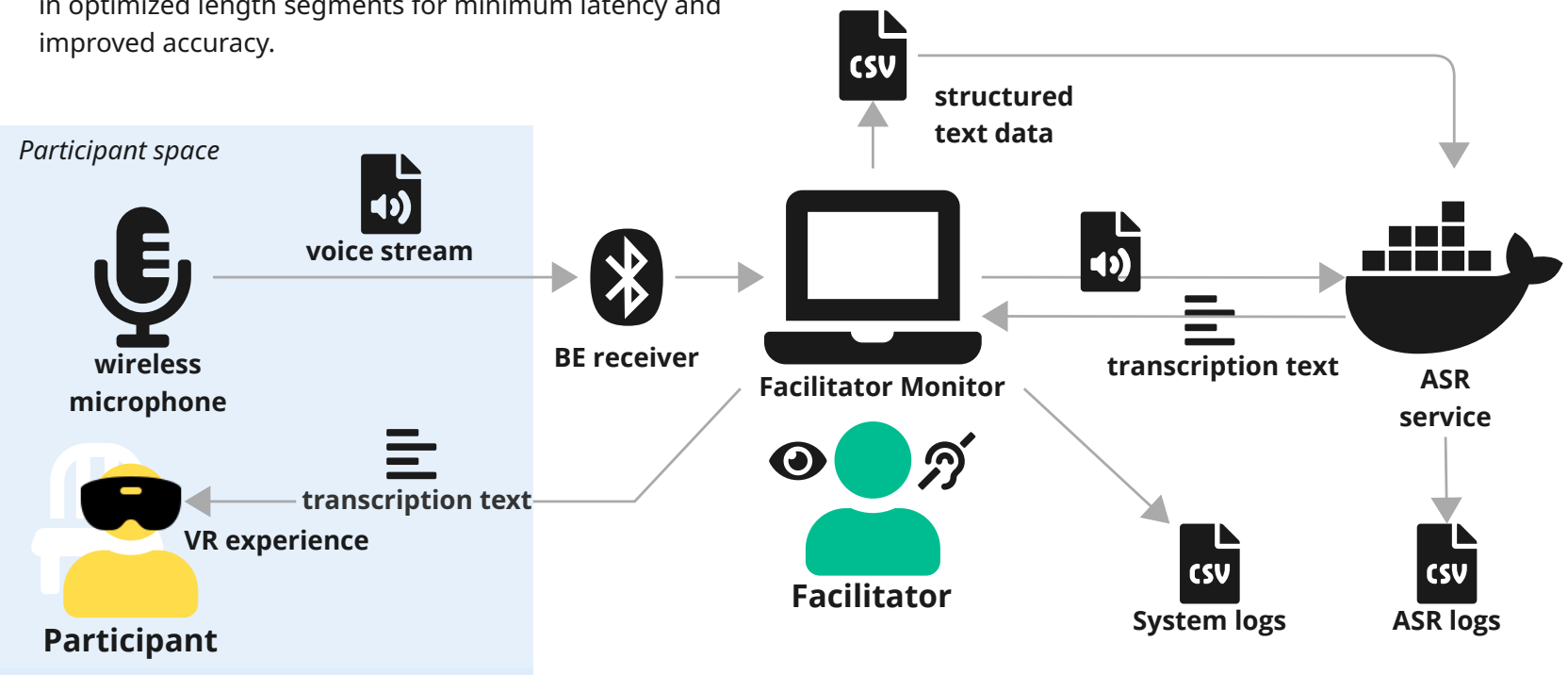


Session Phases

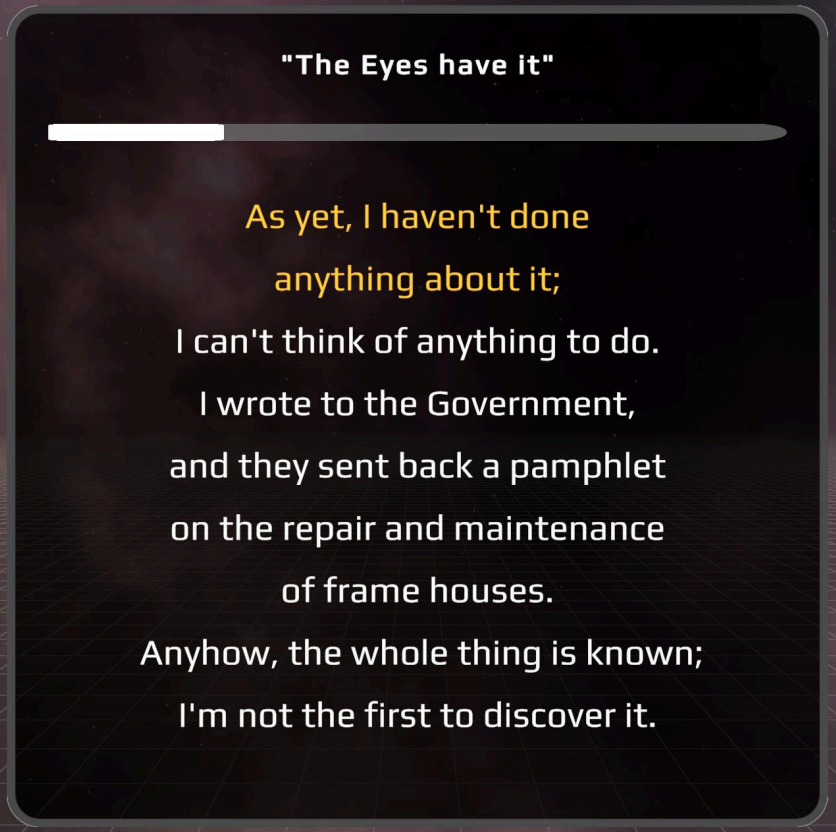
Phase name	Participant action	Function of data	Type of data
Calibrate	Do nothing	Baseline for eye analysis	
Instructions	Read instructions	Baseline for voice analysis	
Main activity	Read text	Test data	
Reflection	Open-ended questions	Self-reporting data, baseline data	

Technology

Use of **voice streaming** with **real-time transcription** and **matching** of raw transcription text to reference text in optimized length segments for minimum latency and improved accuracy.



Immersive experience



Screenshot of the immersive environment with abstract visualization and minimal design.

XR Device Simulator

Open Questions

- How does the impact of voice based interactions on the user experience relate to the **level of immersion** (non-XR, AR, VR)?
- What are best practices for assessing emotional and cognitive responses using **physiological data**, specifically eye and heart rate, for VBIs?
- What are other suitable **assessment methods** for critical engagement or emotional responses in an experiment setting, potentially from other research domains, like media studies?
- How does the **ecological validity** of this experience design affect the reliability the experiment results?

Voice-based interactions in Immersive Interactive Digital Narratives:
Cognitive and Affective Dimensions of the User Experience

Thank You!

PhD Student: **Olga Chatzifoti**, olchatz@di.uoa.gr
Supervisor: **Dr. Maria Roussou**
Department of Informatics and Telecommunications (Di), University of Athens (NKUA)