"This is My Fault", Really? Understanding Blind and Low-Vision People's Perception of Hallucination in Large Vision Language Models

Yilin Tang[1], Yuyang Fang[1], Tianle Wang[2], Liuqing Chen[1]*, Lingyun Sun[1]

[1]Zhejiang University, [2]Beijing University of Posts and Telecommunications

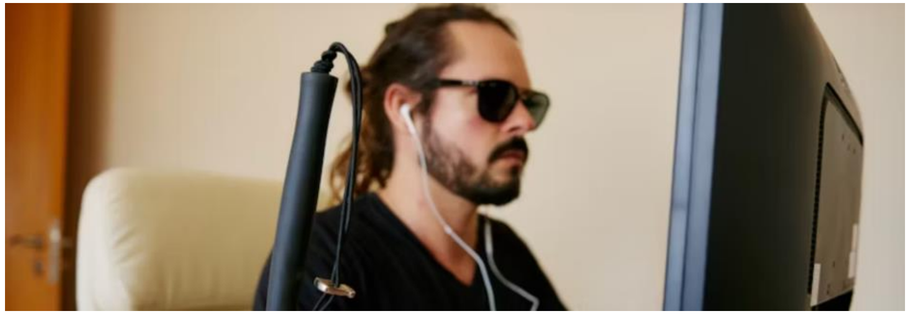
* Corresponding author: chenlq@zju.edu.cn

UIST'25

浙江大学 ZHEJIANG UNIVERSITY ICI Lab

Research Background

- Blind and low-vision (BLV) people encounter challenges in comprehending visual content within their daily lives.



- Visual question-answering (VQA) tools powered by large visual language models (LVLMs) are widely used by BLV individuals.

- LVLMs can generate hallucinations: nonsensical, unfaithful, undesirable textual output given the image inputs

Research Gap

Current research uses general datasets and tasks to evaluate and mitigate AI hallucination.

E.g. MSCOCO dataset

E.g., Tasks involve users without disabilities

CAN NOT be directly applied to BLV-VQA context

- Images posed by BLV people often have quality problems, such as blurriness, poor lighting, and sub-optimal framing.
- Visual queries posed by BLV people cover various usage scenarios and exhibit high lexical diversity.
- BLV people struggle to assess the accuracy of AI responses and rely on them to make decisions.

Research Questions

RQ1: What types and scenarios of hallucinations arise in the context of VQA for BLV individuals?

RQ2: How do BLV individuals perceive hallucinations and the risks associated with them?

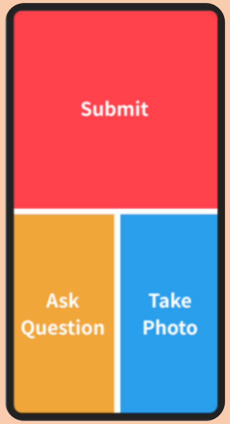
RQ3: What kind of solutions are desired by BLV individuals to help them mitigate the influence of hallucinations?

Study 1: Understanding Hallucinations in the Context of VQA for BLV Individuals: Types, Scenarios, and Proportions

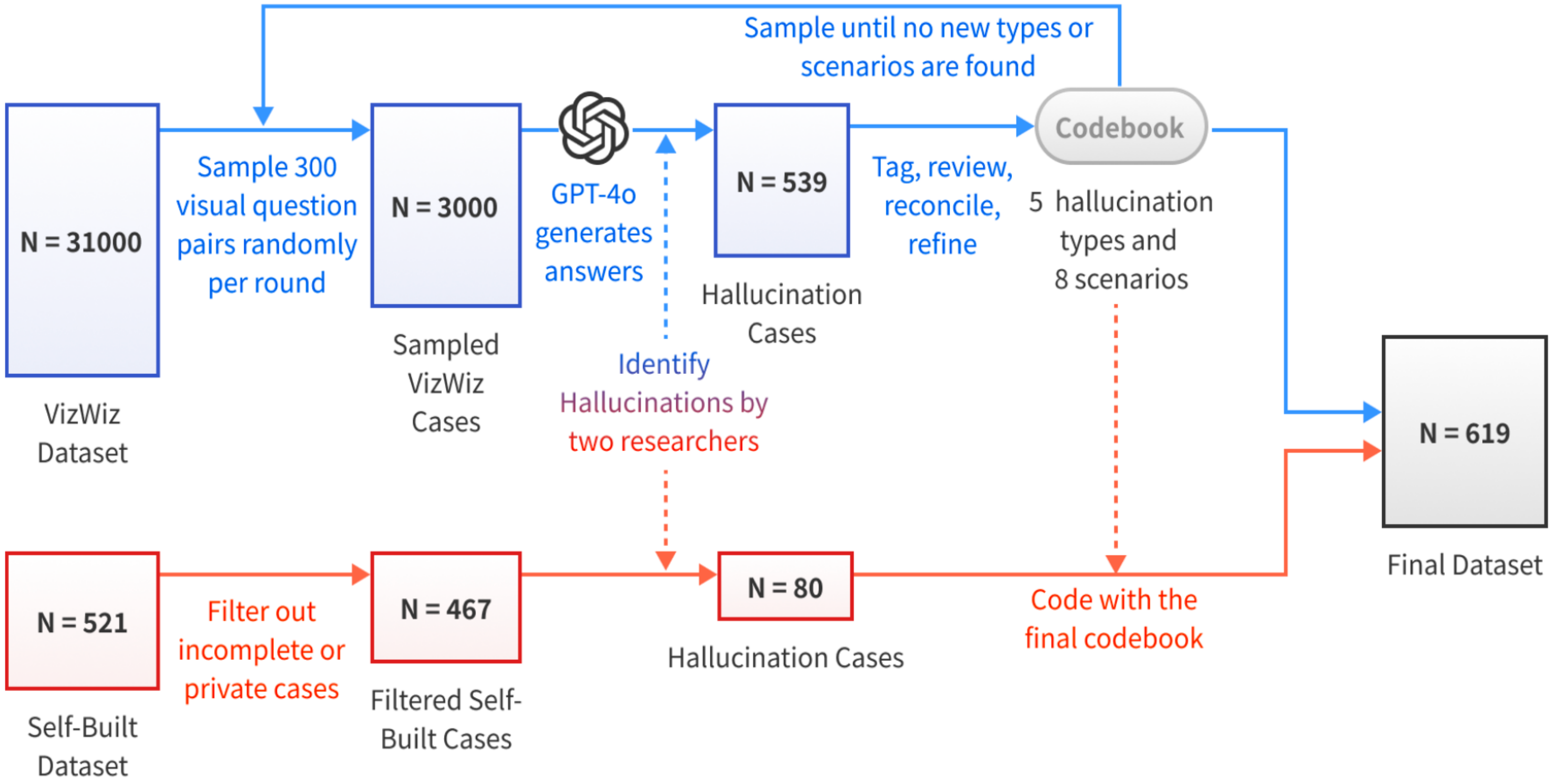
VizWiz

An existing dataset includes the VQA cases posed by BLV people **to human volunteers**

AskVision



A self-built dataset that represents the VQA cases posed by BLV **to LVLMs**.



External phenomena ✅
Underlying mechanisms ❌

Finding 1.1— Five types of hallucination

- **Among the 3,467 cases, 619 cases involved hallucinations, with probabilities exceeding 17%.**

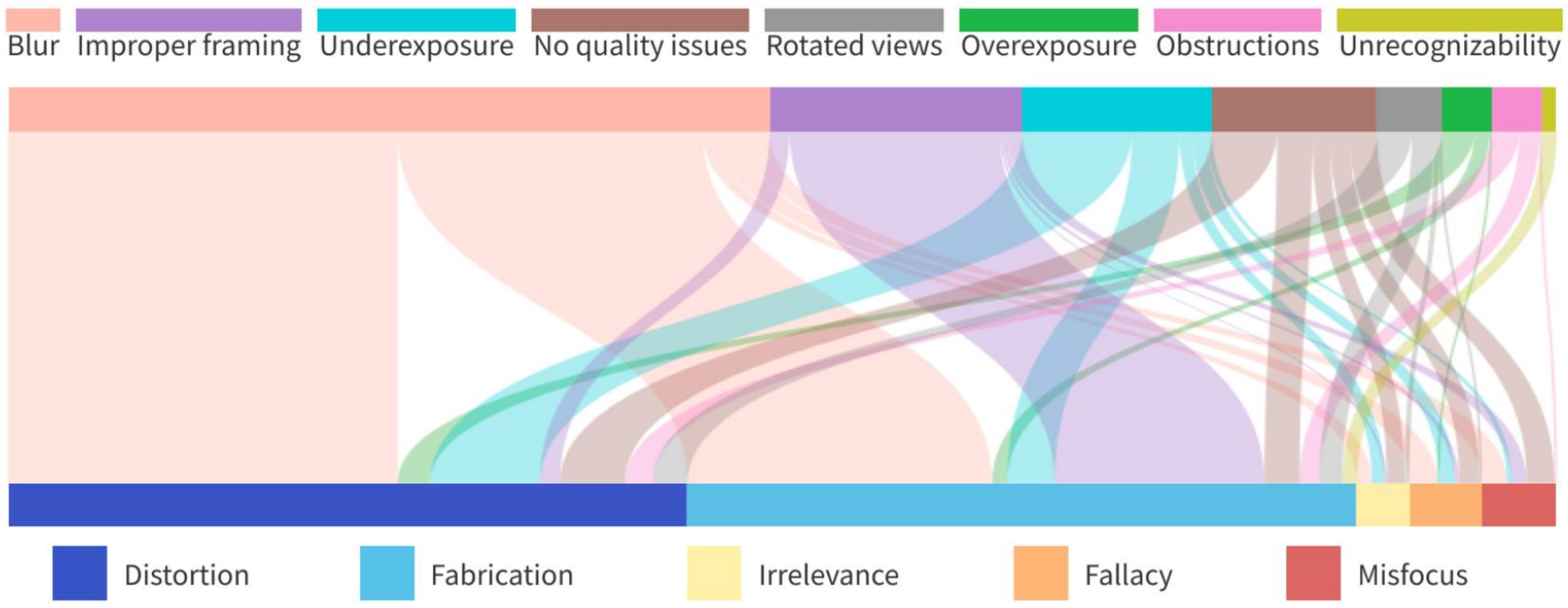
Distortion	LVLMs provide answers that are **INCONSISTENT** with the verifiable information in images for questions that **COULD BE ANSWERED** based on the images
Fabrication	LVLMs **MAKE UP** answers when the image information is **INSUFFICIENT** for the questions
Irrelevance	LVLMs' answers that do not conflict with the image but are **OFF-TOPIC**
Fallacy	LVLMs answer incorrectly when they need to **REASON** based on information in the image
Misfocus	LVLMs attempt to answer the input question, but **ERRONEOUSLY FOCUSING** on parts of the image that are unrelated to the intent of the question

were found to cover existing hallucination taxonomies.

reflect the visual challenges faced by BLV individuals during VQA, as they cannot accurately capture and ask questions.

Finding 1.2— Two potential causes of hallucination

- **Image quality**



- Images with framing and unrecognizable issues predominantly led to Fabrication
- Images with blur, bright and dark issues mostly resulted in Distortion.
- Images without quality issues have a 10.36% chance of causing hallucinations.

- **Content word numbers in visual questions**
  - More informative visual questions are prone to result in Fabrication
  - Less informative questions are more likely to result in Misfocus.

Finding 1.3— Eight scenarios of hallucination

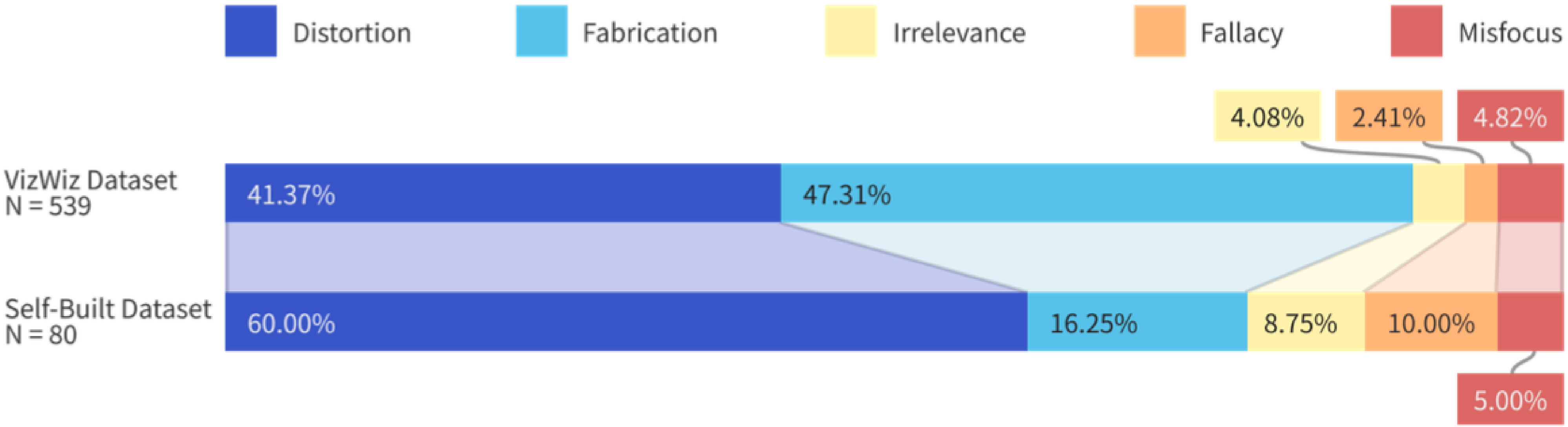
- **The probabilities of the 5 types of hallucinations vary across the 8 scenarios.**

Scenarios	P_S		P_{HS}		Hallucination (P_{HT})				
	in VizWiz	in self-built	in VizWiz	in self-built	Distortion	Fabrication	Irrelevance	Fallacy	Misfocus
House Maintenance	21.60%	32.12%	15.99%	14.77%	**45.63%**	37.86%	6.80%	0.00%	9.71%
Cooking and Dining	34.07%	14.13%	19.49%	30.30%	27.14%	**64.82%**	2.01%	2.01%	4.02%
Healthcare	7.90%	14.35%	18.49%	14.71%	40.91%	**47.73%**	4.55%	4.55%	2.27%
Finance	1.63%	2.36%	30.00%	54.55%	**46.67%**	26.67%	6.67%	20.00%	0.00%
Entertainment	10.07%	3.00%	20.46%	14.29%	38.71%	**50.00%**	4.84%	1.61%	4.84%
Work and Education	10.43%	9.85%	16.50%	10.87%	**59.62%**	21.15%	7.69%	5.77%	5.77%
Outdoor	3.43%	15.63%	11.54%	19.18%	41.67%	**50.00%**	8.33%	0.00%	0.00%
Dressing	10.87%	8.57%	15.95%	2.50%	**71.15%**	26.92%	0.00%	0.00%	1.92%

Note: P_S is calculated as the ratio of the number of cases in a specific scenario to the total number of cases in the dataset. P_{HS} is calculated as the ratio of the number of hallucinations in a specific scenario to the total number of cases in that scenario. P_{HT} is calculated as the ratio of the number of a specific hallucination type in a scenario to the total number of hallucinations in that scenario.

- Visual tasks vary across scenarios, which may account for the different distributions of hallucinations.
- Among all scenarios, Distortion and Fabrication are the most common types of hallucination.
- Distortions are more likely to arise when LVLMs are required to describe the basic characteristics of objects. Fabrications are more frequently observed when LVLMs need to analyze detailed information.

Finding 1.4— Comparing Results from VizWiz and Self-Built Datasets



- The types of hallucinations in VizWiz and self-built datasets were found to be the same
- The distribution of different types of hallucination varies.
- BLV individuals pose more concise and less informative questions to LVLMs.
- BLV individuals are more likely to pose images without quality flaws to LVLMs.

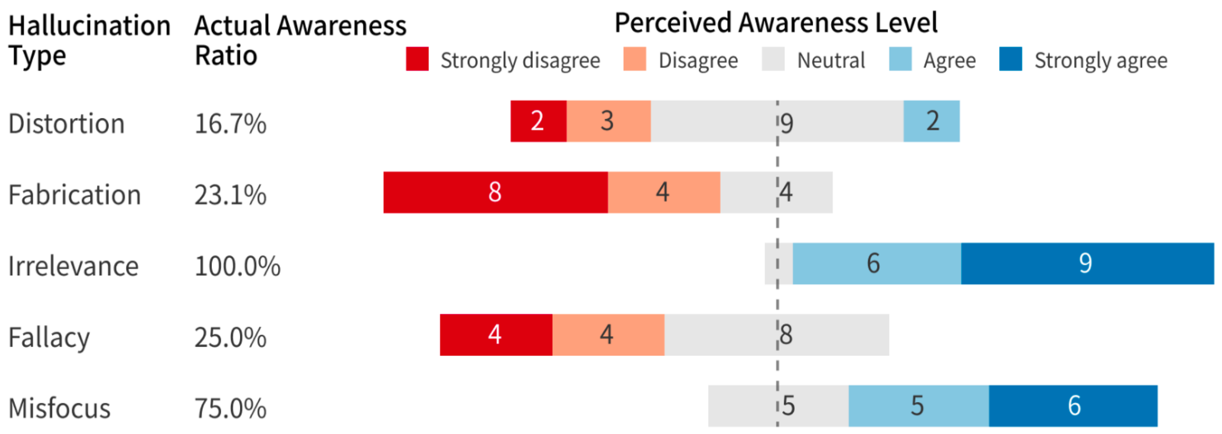
Study 2: Understanding BLV Individuals' Perceptions of Hallucinations and the Risks Associated with Them

- **Conducted semi-structured interviews with 16 BLV individuals.**

ID	G.	Age	Edu.	Occupation	Diagnosis	Onset	Vis.Exp.	Vis. Acc.(Tools)	AI. Acc. (Tools)
P01	W	51	H.S.	Masseur	Retinitis Pigmentosa	31	ATB	3 (V. O., Be My Eyes)	1
P02	M	60	J.H.S.	Community Coordinator	Trauma	21	ATB	3 (V. O., Be My Eyes)	1
P03	W	42	H.S.	Customer Service	Optic Neuropathy	23	AB-LC	3 (V. O.)	1
P04	W	71	P.S.	Retired	Glaucoma	0	CTB	3 (TBack)	1
P05*	M	61	H.S	Retired	Retinal Detachment	16	ALV	3 (TBack)	3 (DouBao, SeeingAI)
P06*	W	28	B.E	Liberal Professions	Hereditary Retinal Diseases	10	AB-LC	3 (V. O., NVDA)	3 (Be my AI, ChatGPT-4V)
P07*	M	22	H.S.	Masseur	Glaucoma	0	CTB	2 (V. O.)	2 (ChatGLM)
P08*	W	49	J.H.S.	Masseur	Trauma	20	ATB	1	1
P09*	M	32	B.A.	Piano tuner	Retinoblastoma	0	CB-LC	3 (V. O.)	1
P10*	M	40	B.E.	Programmer	Optic Atrophy	0	CTB	3 (V. O., NVDA)	2 (ChatGLM)
P11*	W	53	M.A.	Professor in art	Neuromyelitis Optica	33	ALV	3 (V. O.)	2 (DouBao)
P12*	M	37	B.E.	Online sales	Retinitis Pigmentosa	0	CB-LC	3 (V. O., Be My Eyes)	2 (DouBao)
P13*	W	70	J.H.S.	Retired	Diabetic Retinopathy	45	ATB	2 (TBack)	1
P14*	W	69	H.S.	Retired	Glaucoma	58	AB-LC	2 (TBack)	1
P15*	W	53	B.A.	Soprano of a choir	Cataract	0	CTB	2 (V. O.)	1
P16*	M	70	P.S.	Fortune teller	Trauma	3	ATB	3 (V. O.)	1

Finding 2.1— BLV Individuals' Awareness of Hallucinations

- Detected less than 1/3 of hallucination cases.
- Were deceived by the LVLM's rhetoric
- Blame themselves rather than models
- Hallucination types affected their awareness
- Detection strategies of hallucinations

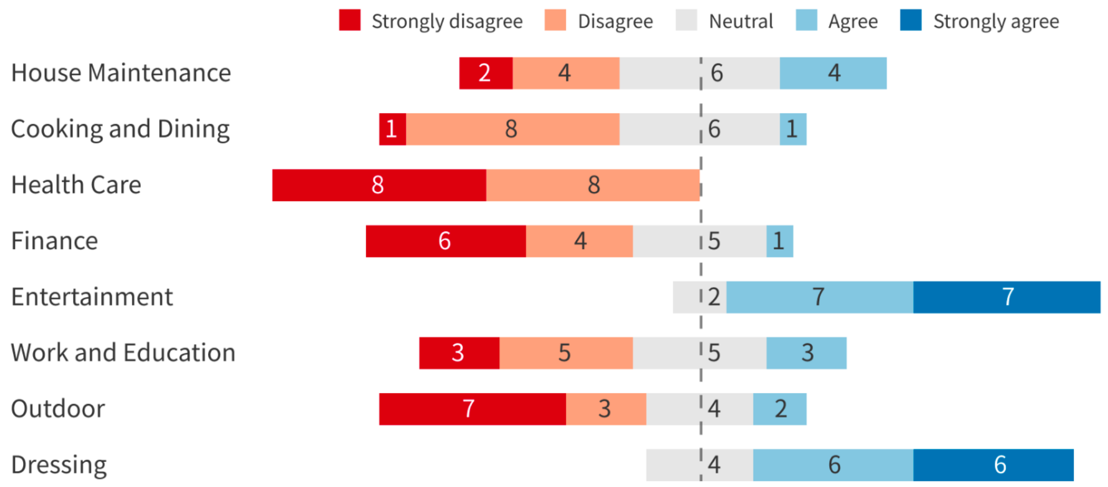


Contextual consistency	observe if the model's response remained consistent in multiple rounds of conversations
Prior knowledge	memories of familiar objects, common sense, or logical reasoning
Multi-sensory approaches	sensitive touch, olfaction, hearing, intuition
External knowledge	share images with family members, sighted volunteers, or other AI models

Participants were also found to flexibly select and combine the four basic detection strategies to deal with different types of hallucinations

Finding 2.2—BLV Individuals' Concerns about Risks

- We identified the potential risks of hallucinations in eight scenarios.
- The tolerance of BLV individuals to these risks varies depending on the usage scenarios.



Finding 2.3— Mental Models of How Hallucinations Occur in LVLMs

Random Lottery	Hallucinations occur purely by chance
Lost in Language	Hallucinations occur when there is a mismatch
Garbage In, Garbage Out	Due to the poor quality of input images
Cognitive Dissonance	Due to imbalances in input and training data

Both four mental models are insufficient

Study 3: Understanding BLV Individuals' Expectations for Mitigating the Impact of Hallucinations

- **Co-design with 12 BLV individuals.**

Warm-up and introduction

Research findings sharing

Fictional inquiry

- Introduced a fictional hallucination expert.
- Envision this expert as an "assistant" and freely express their expectations.
- Explored detailed aspects of the assistant.

Category	Subcategory	Guidelines
Enhancing Information Provision	**Overview Disclosure**	1. In the onboarding tutorial, incorporate voice narration to announce the overall probability of hallucinations (17%). 2. In the help documentation, provide examples of each type of hallucination, along with their characteristics and occurrence probabilities, in both textual and audio formats. 3. In the help documentation, elucidate the generation mechanism of hallucinations.
	Realtime Alert	1. Automatically recognize the current scenario and provide speech prompts on hallucination probability, the most common hallucination type, and potential risks in that scenario. 2. Use auditory cues or vibration to alert scenarios with low risk tolerance levels, high hallucination probability,, especially when the common type of hallucination in that scenario is one that BLV individuals have a low awareness of. 3. Allow users to customize the alerts they find necessary and the forms of alerting (voice, vibration, sound effects, etc.).
Increasing the Trans-parency of Processing	**Source**	1. Provide tactile feedback to let users understand the proportion and location of the image parts referenced in generating the answer. 2. Use different timbres to help users distinguish which parts of the answer come from the image, prior knowledge, external resources, etc. 3. Use speech to explain the thought process when reasoning is involved.
	Confidence Level	1. Assess answer confidence level based on image quality, question clarity, and the model's performance in that specific scenario, and express it through tone or wording. 2. Use speech prompts to explain the primary factors affecting the confidence level of the answer.
Introducing Hallucination Verification Strategies	**Automatic Validation**	1. Verify the answer's consistency with the structure or intent of the user's query. 2. Obtain multiple outputs for the same visual question using various models and check the consistency among the outputs. 3. Use other visual questions from the same time period as context to validate the reliability of the answers. 4. Verify the consistency of the answer with known facts or common sense.
	User-Involved Information Supplement	1. Use speech prompts to elicit additional task information or to request clarification of the intent behind the questions. 2. Prompt users to provide additional images or photos from different angles.
	Third-Party Verification	1. Help users connect with family, friends, or remote volunteers for verification. 2. Utilize small-scale deterministic models for verification of specific tasks.
Incorporating Feedback Mechanism	**Feedback-Driven Improvement**	1. Allow users to quickly tag and report cases they suspect contain hallucinations and the reasons for their suspicion through voice commands or specific gestures. 2. Incorporate user-reported hallucination cases into model training.

- **Analyzing Hallucination-related Challenges through Probability- Hallucination-Risk Multidimensional Lens**

  e.g. Healthcare scenario, the occurrence probability of hallucination is relatively low, the most common type of hallucination in this scenario is Distortion, which is the least detectable by BLV users, and users exhibit the lowest tolerance for risks. Therefore, addressing hallucination-related challenges in the Healthcare scenario remains a critical priority.

- **Charting the Design Space for Hallucination-Mitigating VQA Solutions**

Design / HCI

- introducing fundamental concepts of hallucinations to BLV users.
- adopting transparent designs that proactively disclose unforeseen risks.
- integrating hallucination detection techniques
- …..

Technical

- Including more image samples from scenes with high probabilities in the training dataset.
- adjusting the model architecture to balance different modalities.
- detecting different types of hallucinations through image quality.
- ….

- **BLV Individuals’ Trust in AI-based Visual Assistance Tools**

Existing findings

Placed great trust in computer-generated captions?
Reported low trust in AI descriptions?

Our findings

Lacks trust in AI capabilitie
s

Assign simpler tasks to AI

less informative visual questions

images without quality flaw

don't believe that simple tasks can go wrong

over-reliance on AI’s outputs &blame themselves

“This is My Fault”, Really? Understanding Blind and Low-Vision People’s Perception of Hallucination in Large Vision Language Models

Yilin Tang[1], Yuyang Fang[1], Tianle Wang[2], Liuqing Chen[1]*, Lingyun Sun[1]

[1]Zhejiang University, [2]Beijing University of Posts and Telecommunications

* Corresponding author: chenlq@zju.edu.cn

UIST’25