LaCC Lab

TECHNION
Israel Institute
of Technology

Fine-Grained Prediction of Reading Comprehension from Eye Movements

Omer Shubi

Yoav Meiri

Cfir Hadar

Prof. Yevgeni Berzak

Language, Computation and Cognition (LaCC) Lab
Faculty of Data and Decision Sciences | Technion

Ready… Set… Read

Leading water scientists have issued one of the sternest warnings yet about global food supplies, saying that the world's population may have to switch almost completely to a vegetarian diet by 2050 to avoid catastrophic shortages. Humans derive about 20% of their protein from animal-based products now, but this may need to drop to just 5% to feed the extra two billion people expected to be alive by 2050, according to research by some of the world's leading water scientists. "There will not be enough water available on current croplands to produce food for the expected nine-billion population in 2050 if we follow current trends and changes towards diets common in western nations," the report by Malik Falkenmark and colleagues at the Stockholm International Water Institute (SIWI) said.

Q: According to Malik Falkenmark's report, what will happen if the world adopts the current diet trends of western nations?

A: By 2050, animal-based protein consumption will reduce from 20% to 5%

B: By 2050, nine billion people will not have enough drinking water

C: There will not be sufficient water to grow enough food for everyone

D: Obesity rates around the world will rise

Q: According to Malik Falkenmark's report, what will happen if the world adopts the current diet trends of western nations?

A: By 2050, animal-based protein consumption will reduce from 20% to 5%

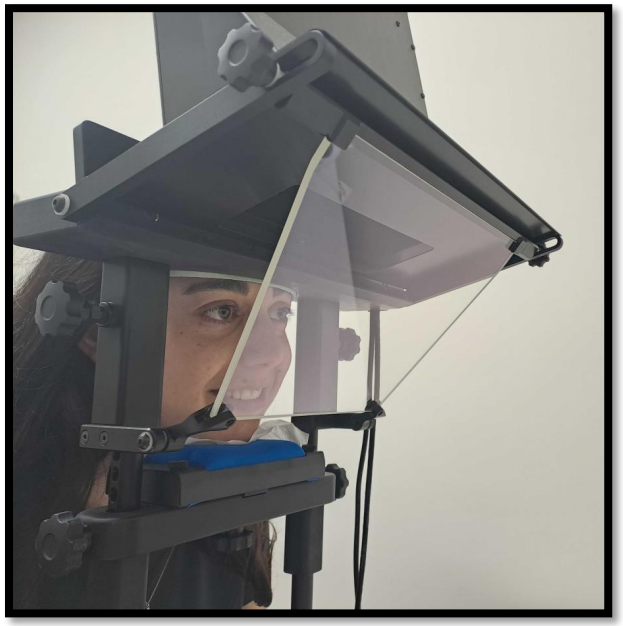
B: By 2050, nine billion people will not have enough drinking water

C: There will not be sufficient water to grow enough food for everyone

D: Obesity rates around the world will rise

What physiological or behavioral cues might reveal comprehension level, bypassing the need for answering traditional comprehension questions?

Eye-Movements Based Comprehension Prediction



Fine-Grained Prediction of Reading Comprehension from Eye Movements

Task Definition

Eye-Tracking Dataset

Text & Eye Movements Modeling

Systematic Evaluation

Fine-Grained Prediction of Reading Comprehension from Eye Movements

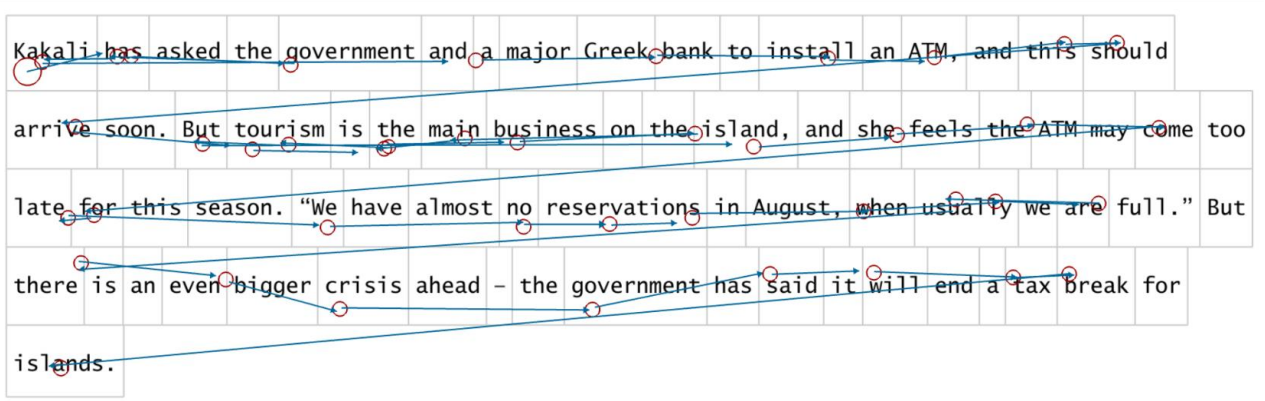
Task Definition

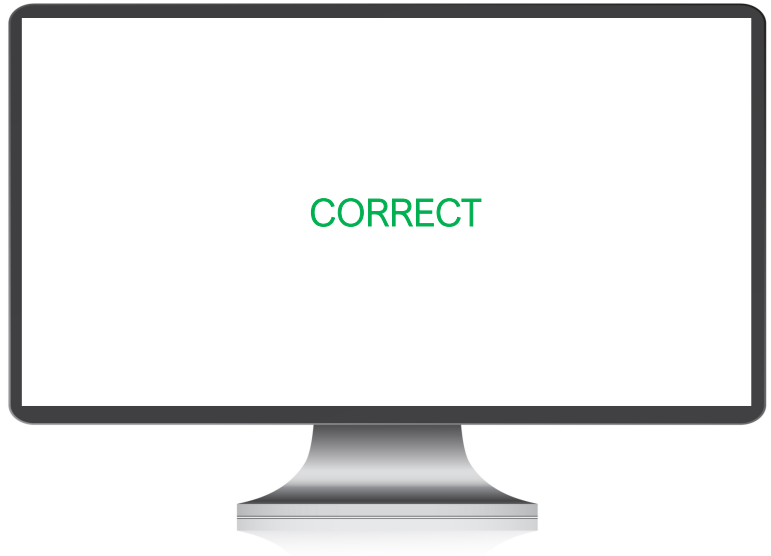
Eye-Tracking Dataset

Text & Eye Movements Modeling

Systematic Evaluation

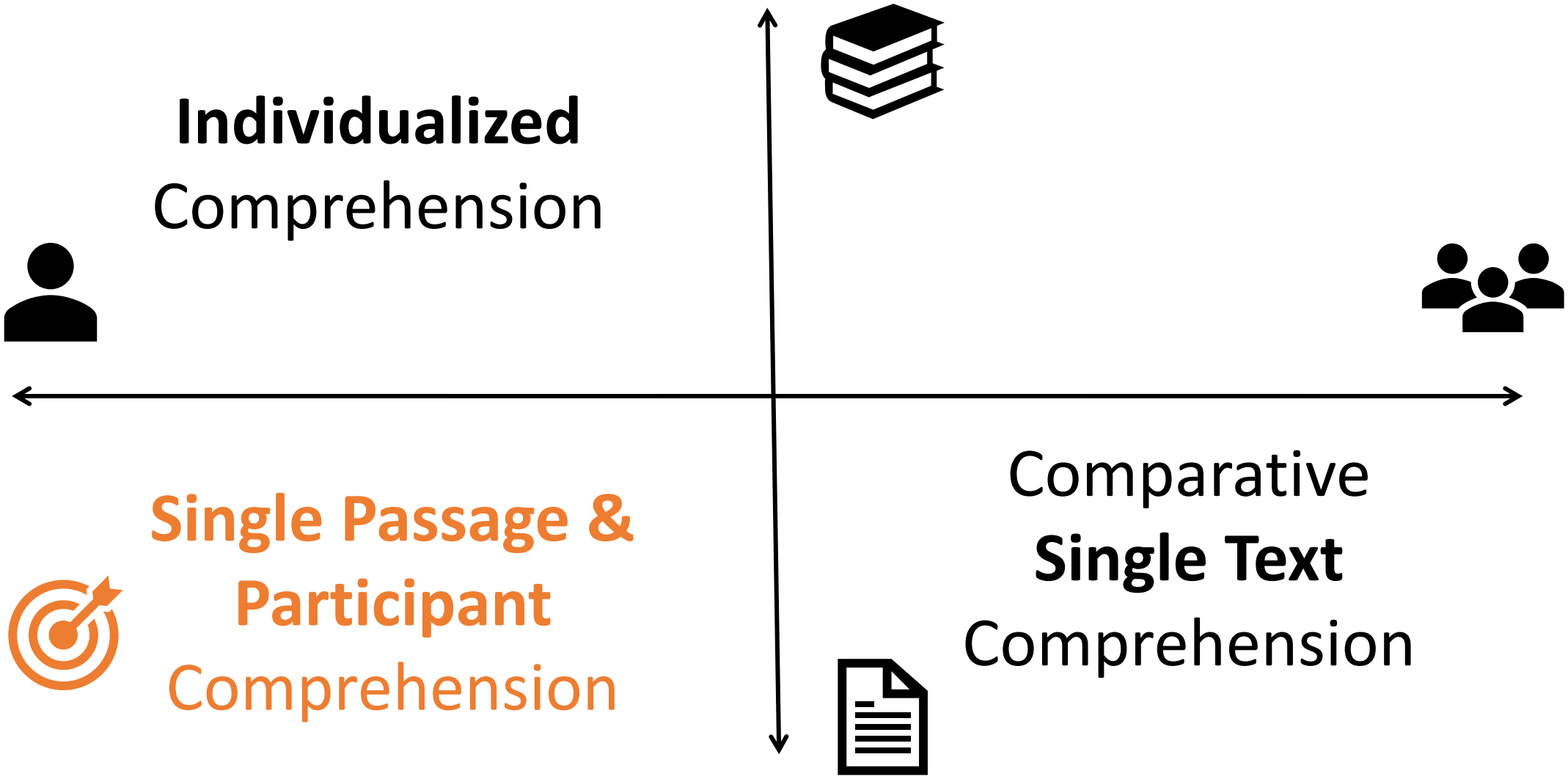
Eye-Movements Based Comprehension Prediction





Q: What will result from an increase in human population in the future?

Eye-Movements Based Comprehension Prediction



Eye-Movements Based Comprehension Prediction

Single Passage & Participant Comprehension

High-Quality Language Models

Large-Scale Eye-Tracking Dataset

Fine-Grained Prediction of Reading Comprehension from Eye Movements

Task Definition

Eye-Tracking Dataset

Text & Eye Movements Modeling

Systematic Evaluation

Data

Textual Materials – OneStopQA

Eye Tracking Experiment – OneStop Eye Movements

Data

Textual Materials – OneStopQA

Eye Tracking Experiment – OneStop Eye Movements

OneStopQA (Berzak 2020)

onestop_qa

- Multiple-choice question-answering dataset

In the next 30 years, the planet's human … billion. Already one billion people do not get enough fo… pressure on agricultural land, water, forests, fisheries and rese… ergy supplies. The cost of meat is increasing – it costs more m… have to destroy a lot of rainforest to make fields or to grow food for cows. Co… methane. The farming of cows, pigs and sheep makes very large amounts of greenhouse gases – 35% of the planet's methane, 65% of its nitrous oxide and 9% of the carbon dioxide.

324 Paragraphs

Q: What will result from an increase in hum… ulation in the future?

a) More pressure on farmi…

b) One billion p… nough food

c) The level o… se gases will increase by 35%

d) Food quality will decrease

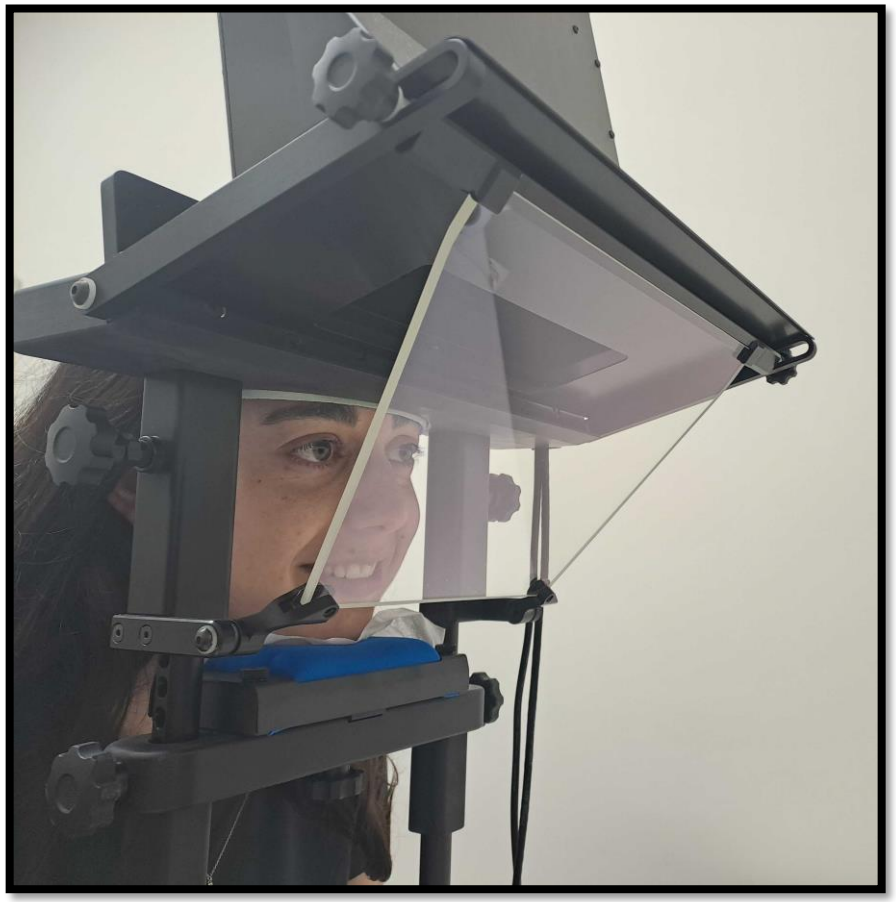
× 3 Questions

Data

Textual Materials – OneStopQA

Eye Tracking Experiment – OneStop Eye Movements

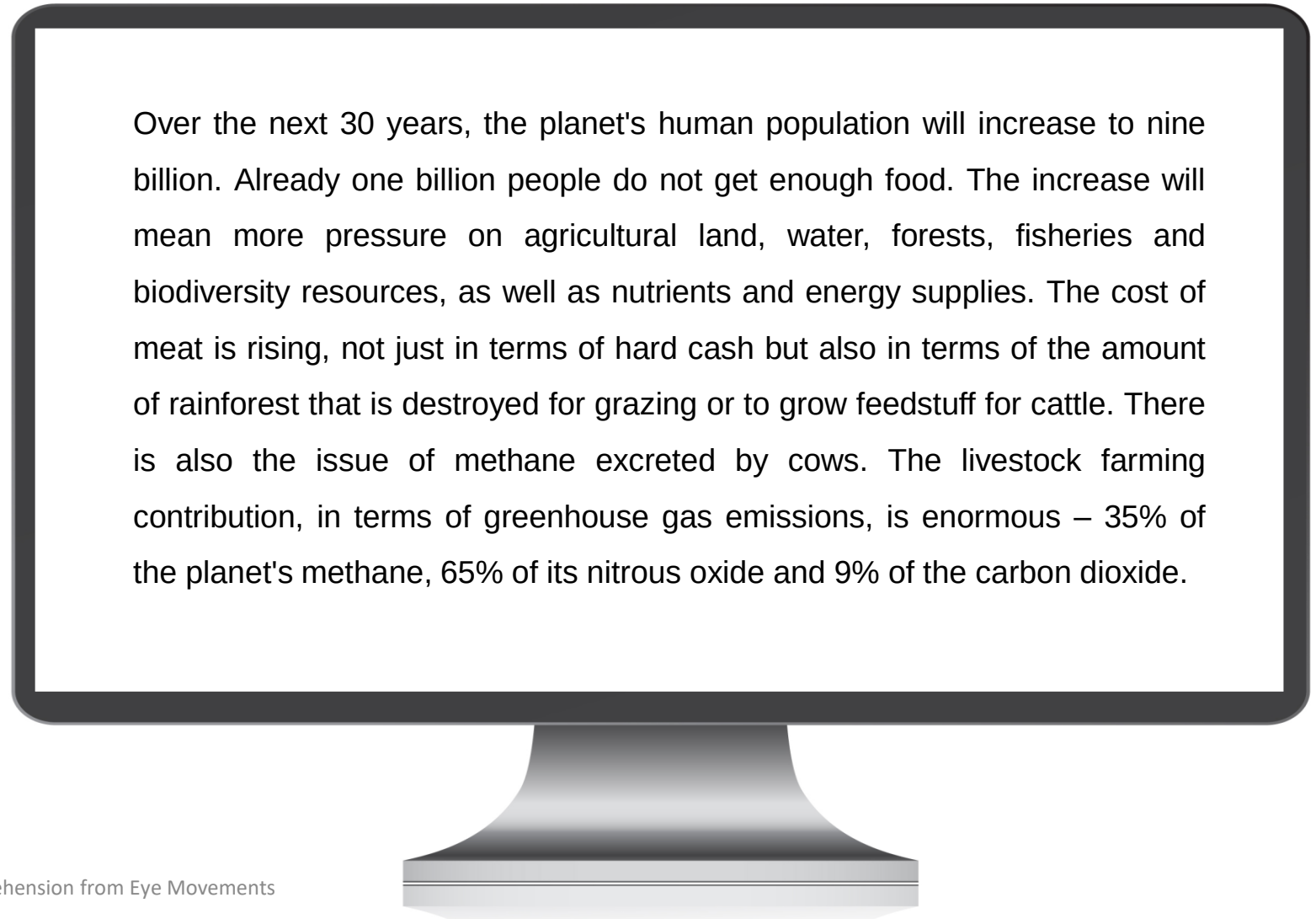
Eye Tracking Experiment



Gathering

Gathering

Gathering

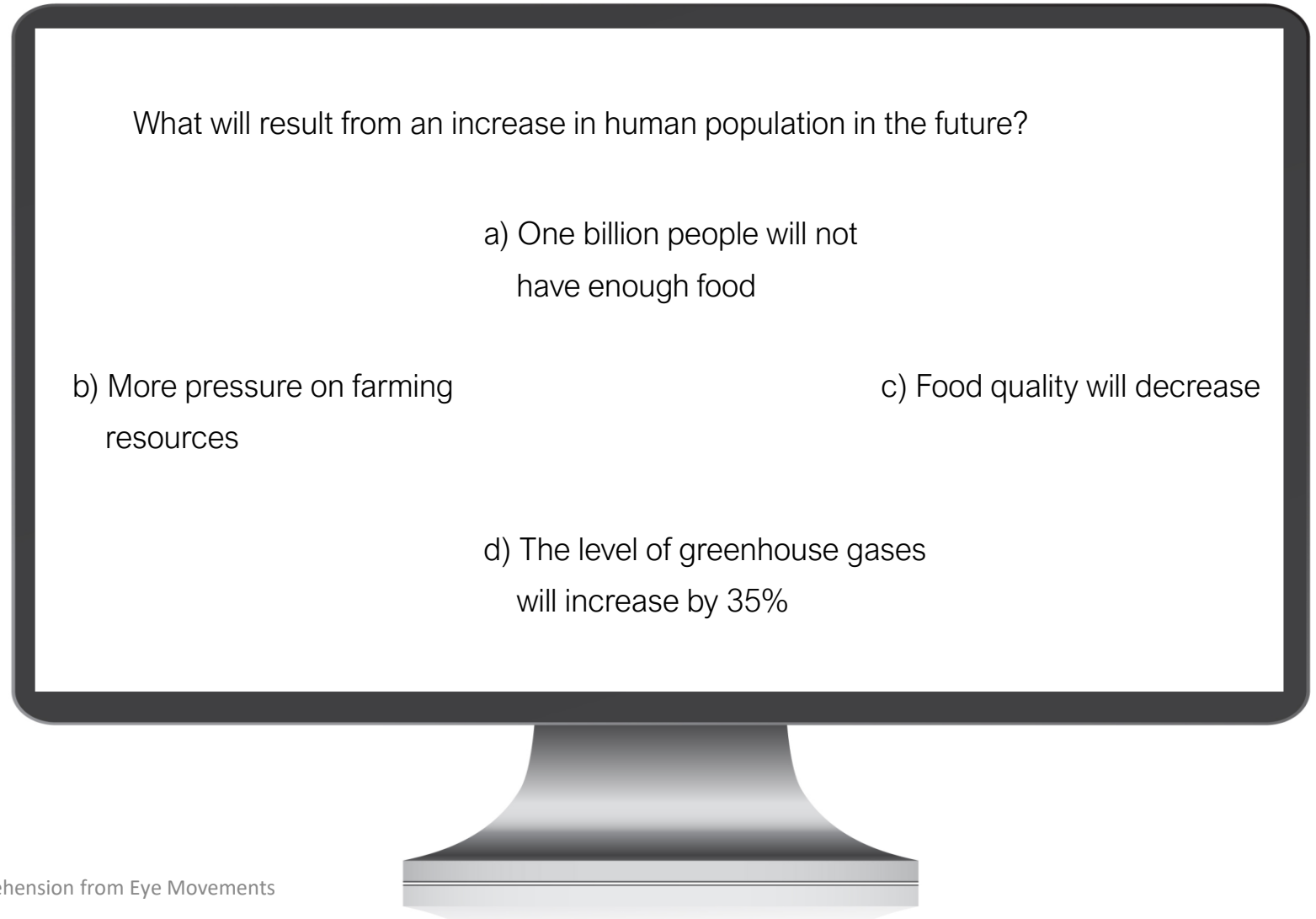

Over the next 30 years, the planet's human population will increase to nine billion. Already one billion people do not get enough food. The increase will mean more pressure on agricultural land, water, forests, fisheries and biodiversity resources, as well as nutrients and energy supplies. The cost of meat is rising, not just in terms of hard cash but also in terms of the amount of rainforest that is destroyed for grazing or to grow feedstuff for cattle. There is also the issue of methane excreted by cows. The livestock farming contribution, in terms of greenhouse gas emissions, is enormous – 35% of the planet's methane, 65% of its nitrous oxide and 9% of the carbon dioxide.

Gathering

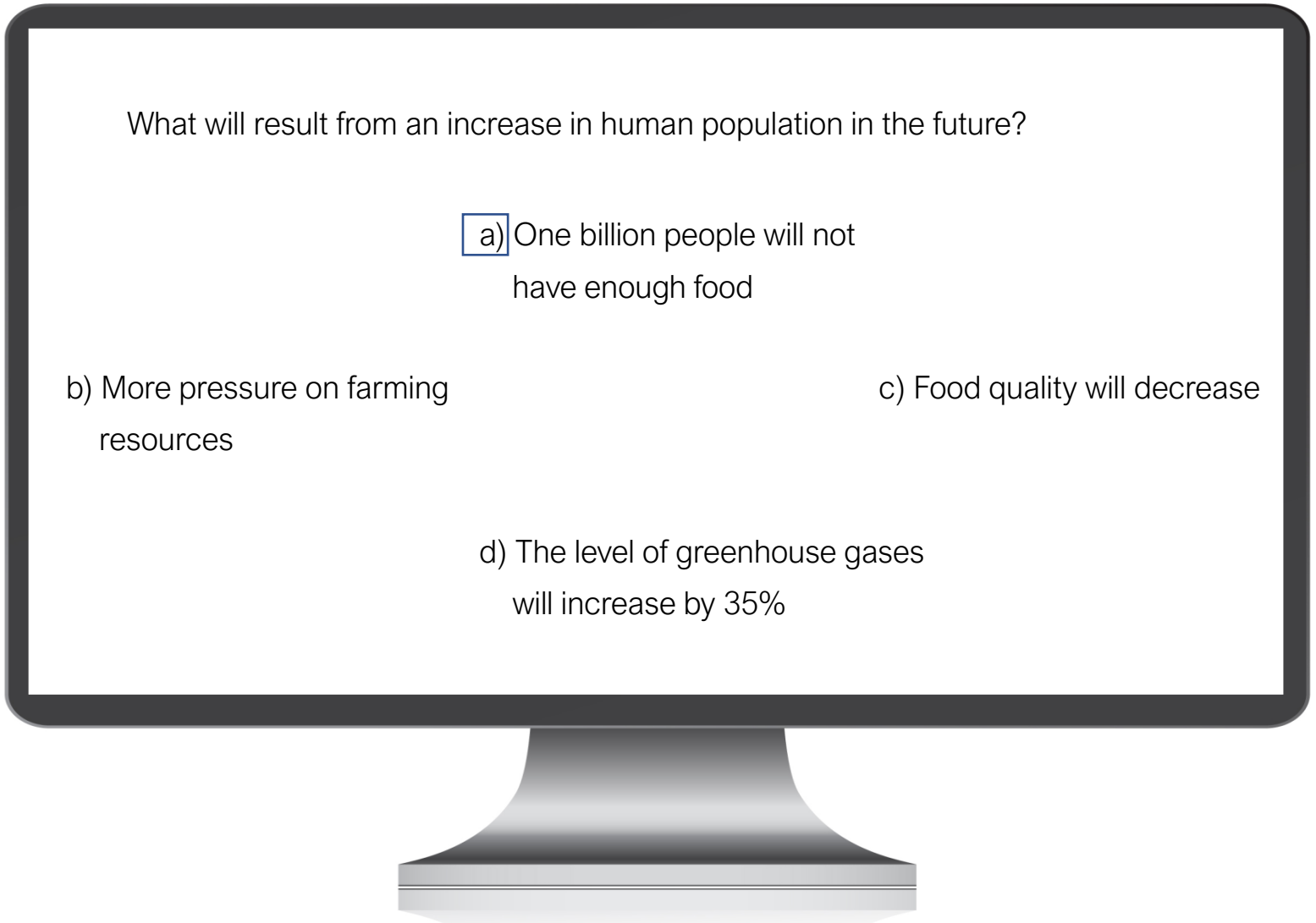
Gathering


What will result from an increase in human population in the future?

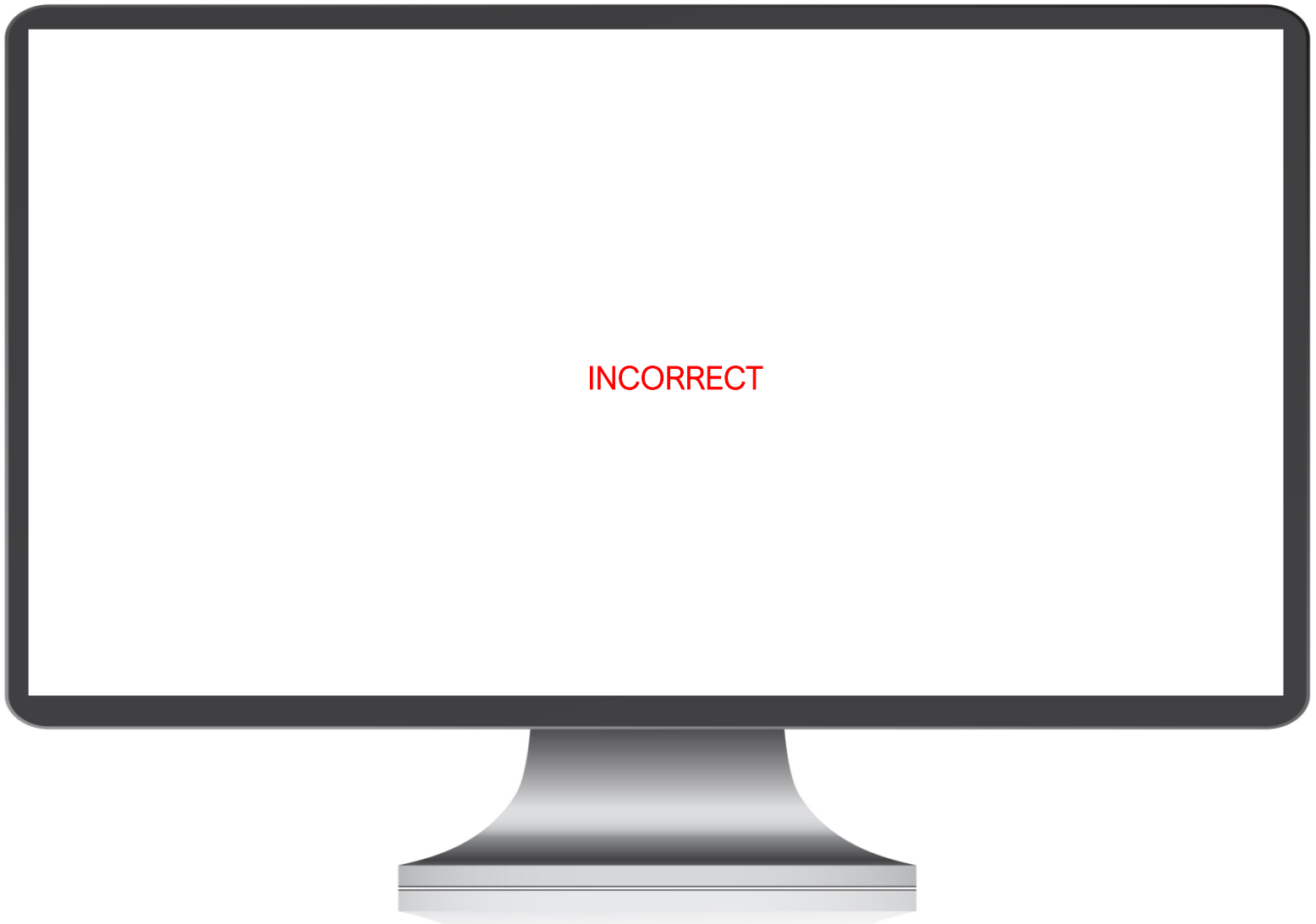
Gathering


What will result from an increase in human population in the future?
a) One billion people will not have enough food
b) More pressure on farming resources
c) Food quality will decrease
d) The level of greenhouse gases will increase by 35%

Gathering


What will result from an increase in human population in the future?
a) One billion people will not have enough food
b) More pressure on farming resources
c) Food quality will decrease
d) The level of greenhouse gases will increase by 35%

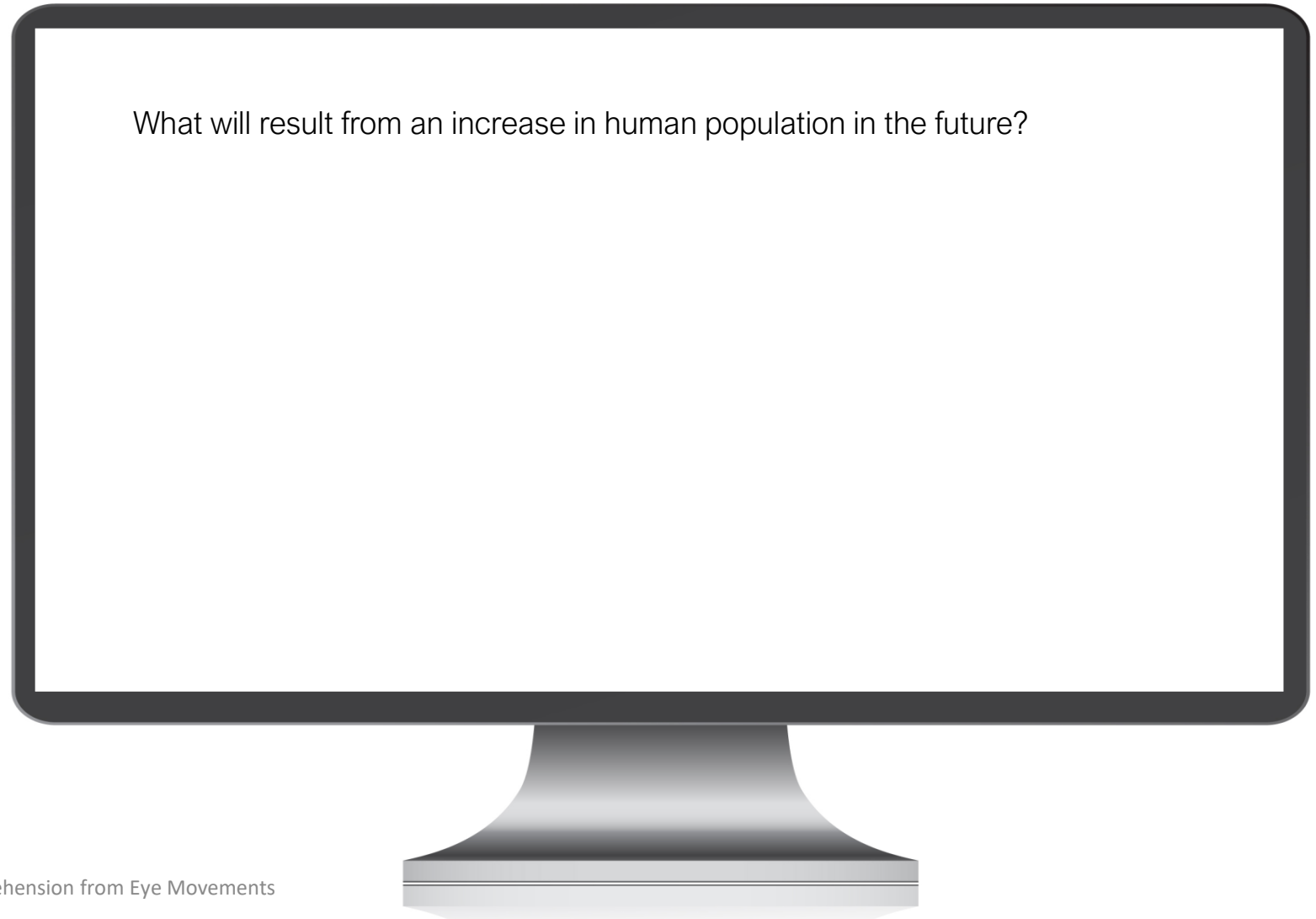
Gathering



Hunting

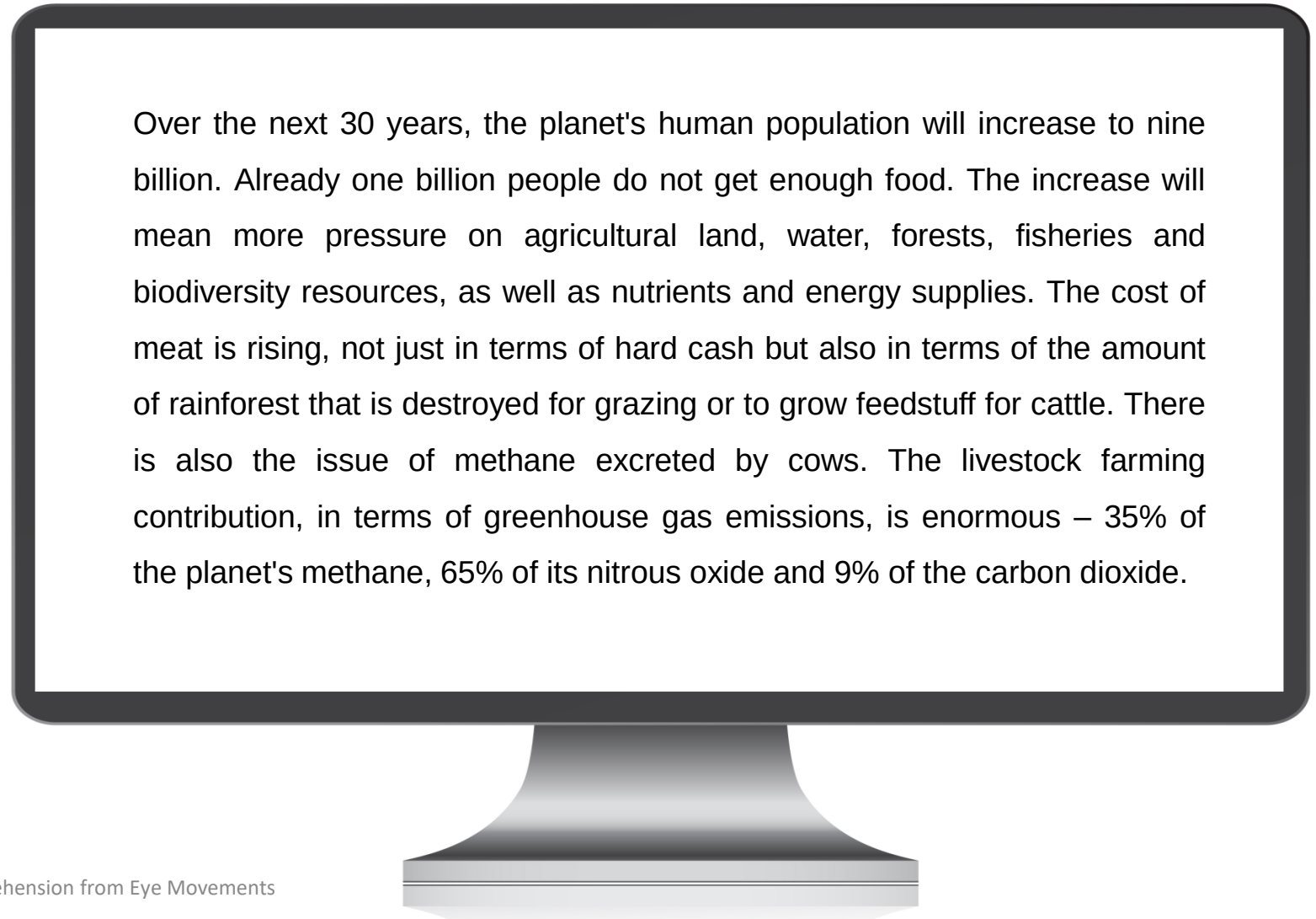
Hunting

Hunting


What will result from an increase in human population in the future?

Hunting

Hunting

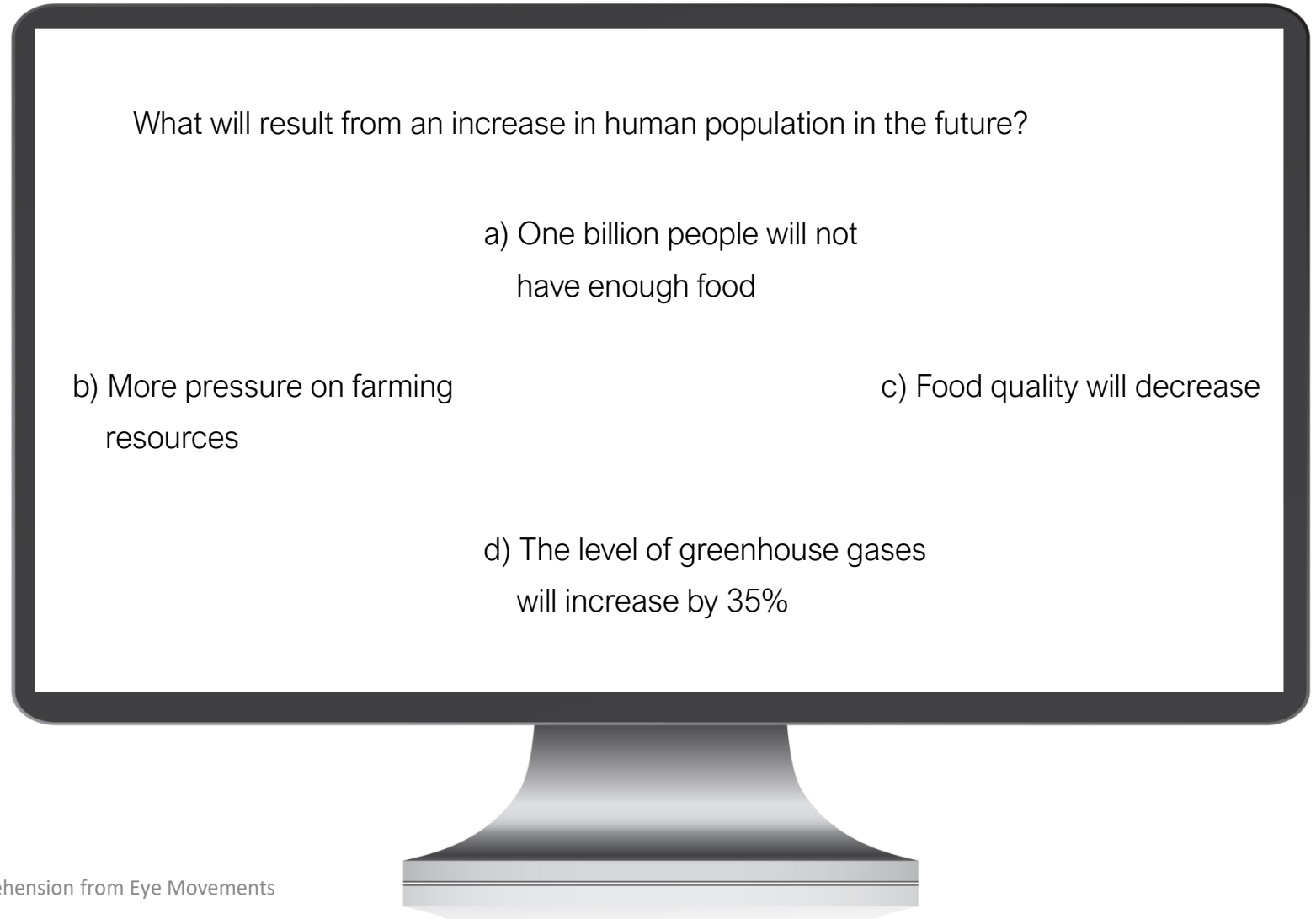

Over the next 30 years, the planet's human population will increase to nine billion. Already one billion people do not get enough food. The increase will mean more pressure on agricultural land, water, forests, fisheries and biodiversity resources, as well as nutrients and energy supplies. The cost of meat is rising, not just in terms of hard cash but also in terms of the amount of rainforest that is destroyed for grazing or to grow feedstuff for cattle. There is also the issue of methane excreted by cows. The livestock farming contribution, in terms of greenhouse gas emissions, is enormous – 35% of the planet's methane, 65% of its nitrous oxide and 9% of the carbon dioxide.

Hunting

Hunting


What will result from an increase in human population in the future?

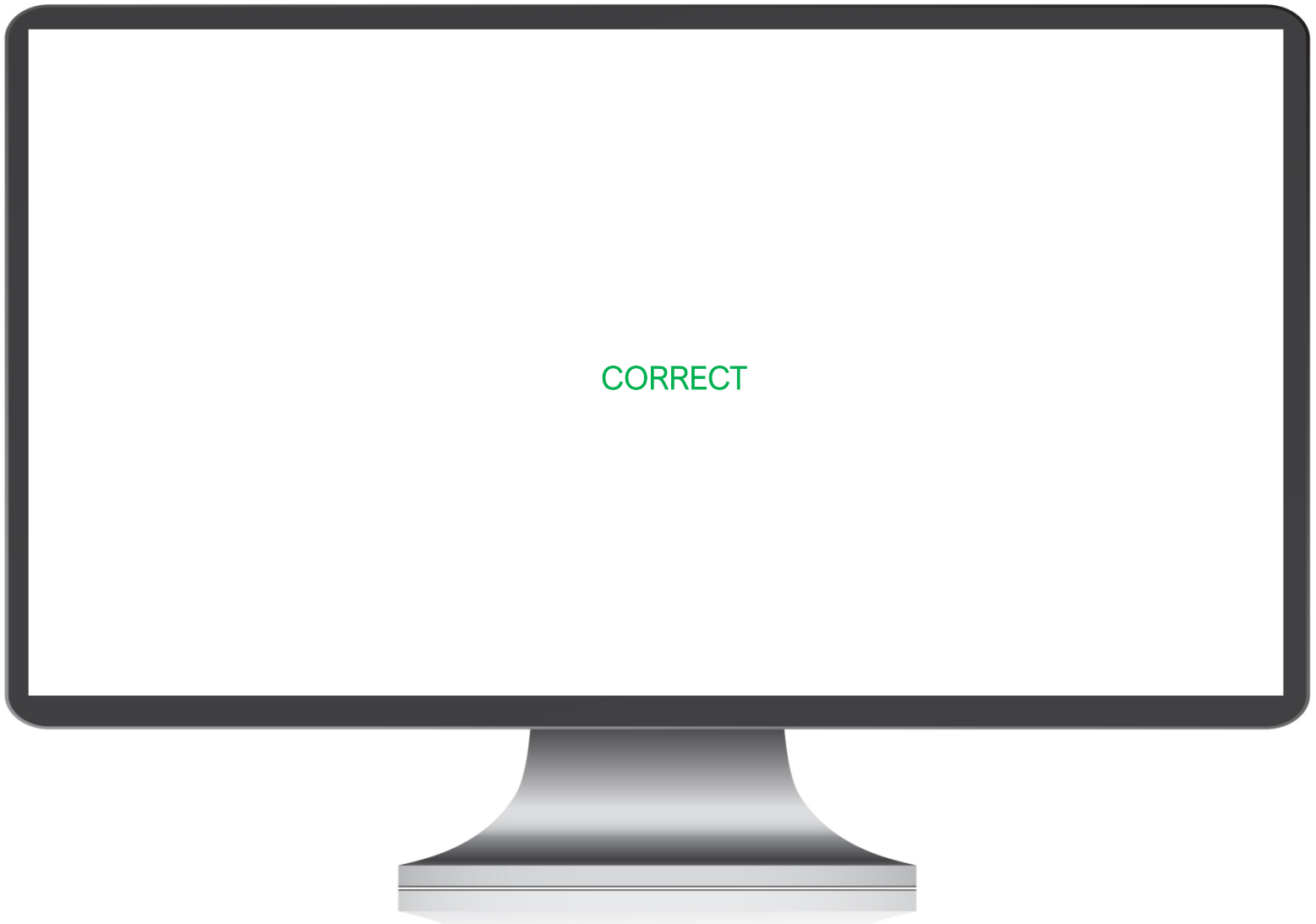
Hunting


What will result from an increase in human population in the future?
a) One billion people will not have enough food
b) More pressure on farming resources
c) Food quality will decrease
d) The level of greenhouse gases will increase by 35%

Hunting


What will result from an increase in human population in the future?
a) One billion people will not have enough food
b) More pressure on farming resources
c) Food quality will decrease
d) The level of greenhouse gases will increase by 35%

Hunting



Experiment Details

× 54 Passages

±20k answers

Fine-Grained Prediction of Reading Comprehension from Eye Movements

Task Definition

Eye-Tracking Dataset

Text & Eye Movements Modeling

Systematic Evaluation

Comprehension Prediction Framework

Task Variations

Eye Movements Representation

Text & Eye Movements Modeling

Comprehension Prediction Framework

Task Variations

Eye Movements Representation

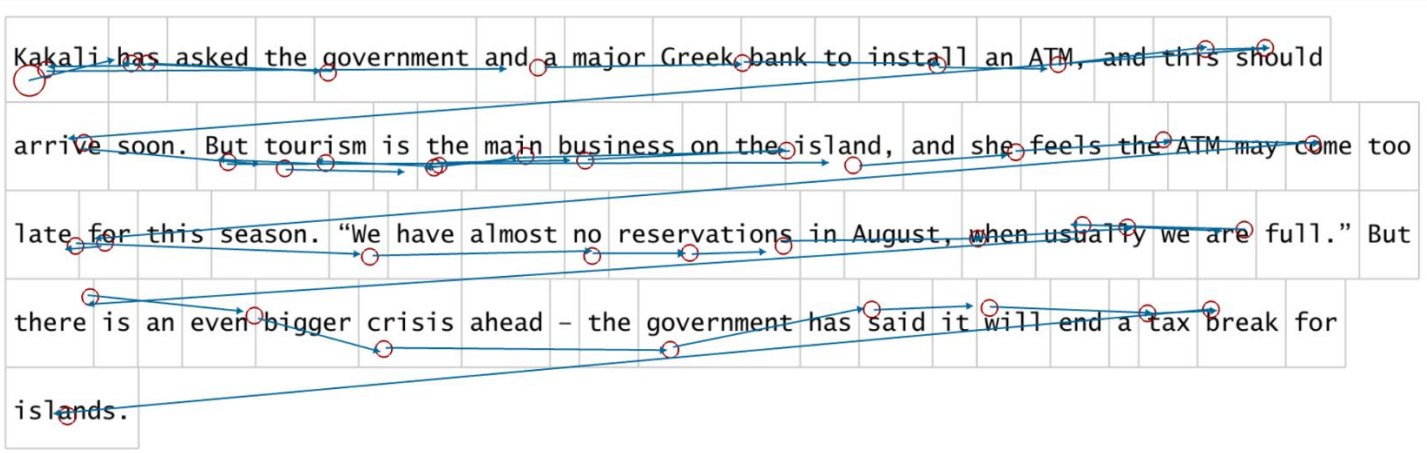
Text & Eye Movements Modeling

Task Variations

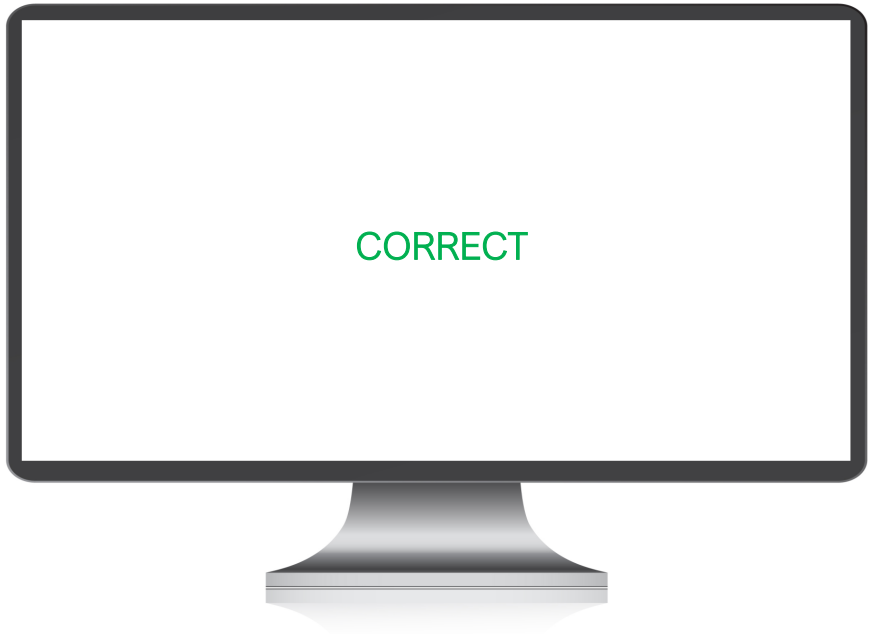
Binary

Multiclass

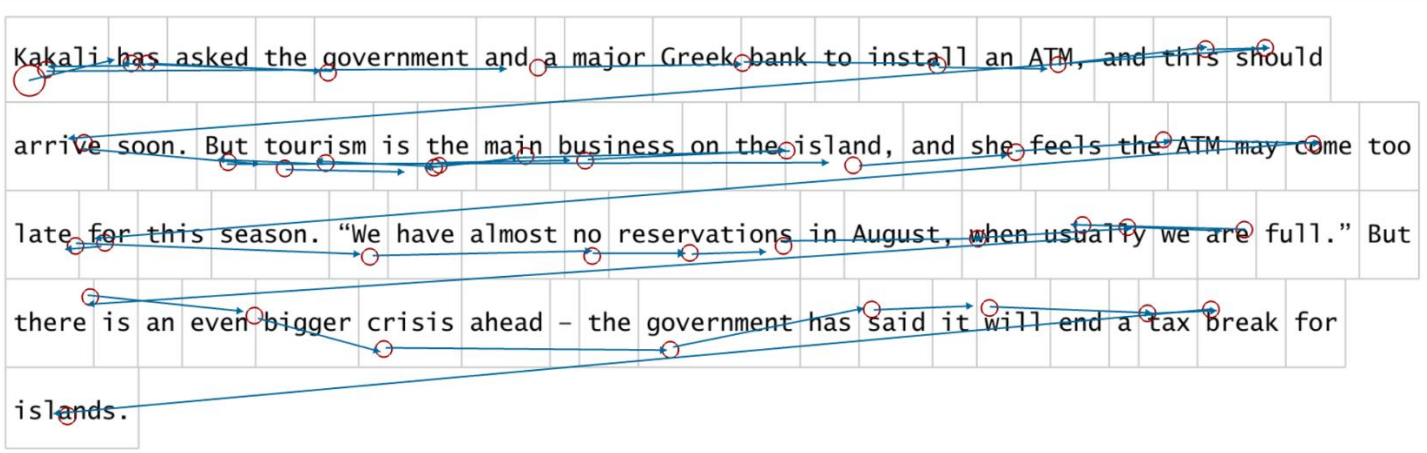
Binary Reading Comprehension

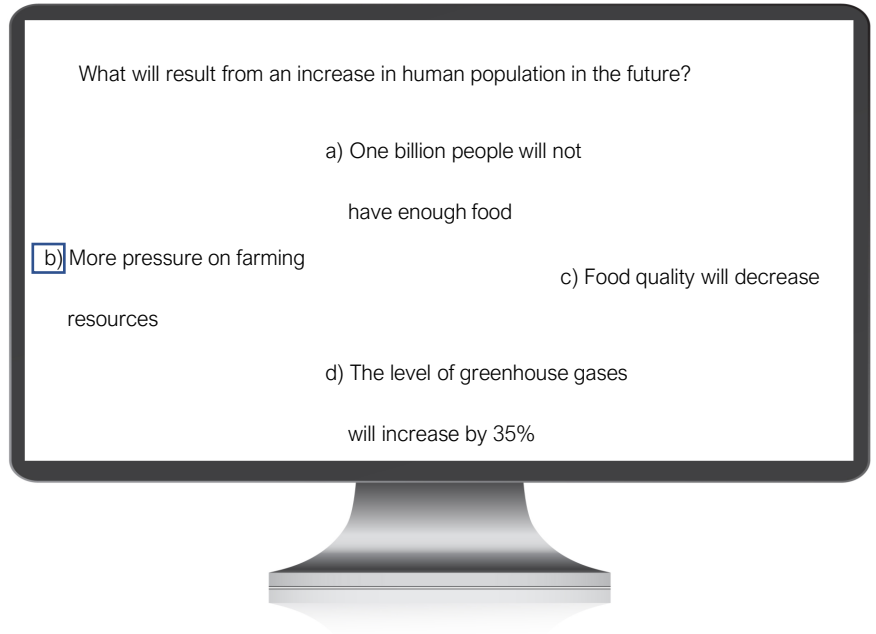


Q: What will result from an increase in human population in the future?



Multiple-choice Question Answering





Q: What will result from an increase in human population in the future?

a) More pressure on farming resources

b) One billion people will not have enough food

c) The level of greenhouse gases will increase by 35%

d) Food quality will decrease

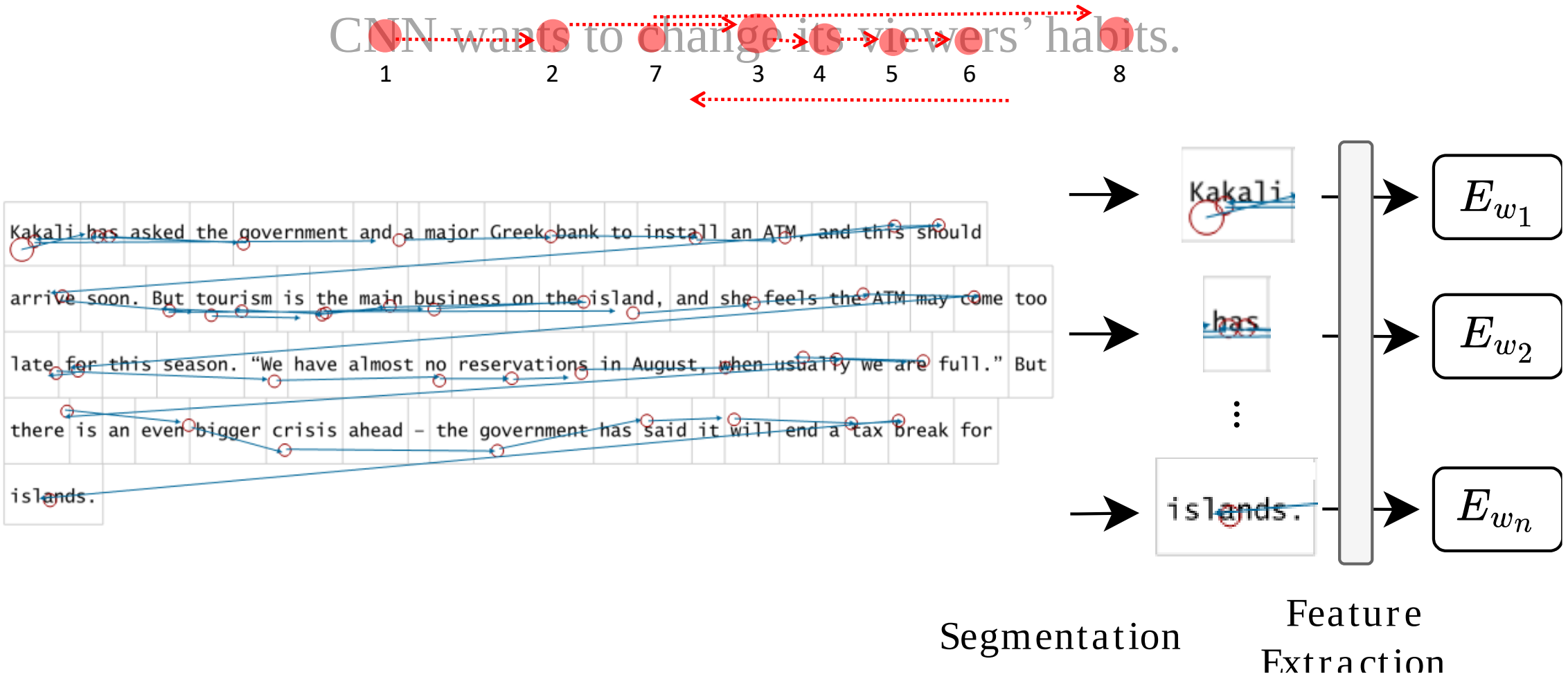
Comprehension Prediction Framework

Task Variations

Eye Movements Representation

Text & Eye Movements Modeling

Representing Eye Movements



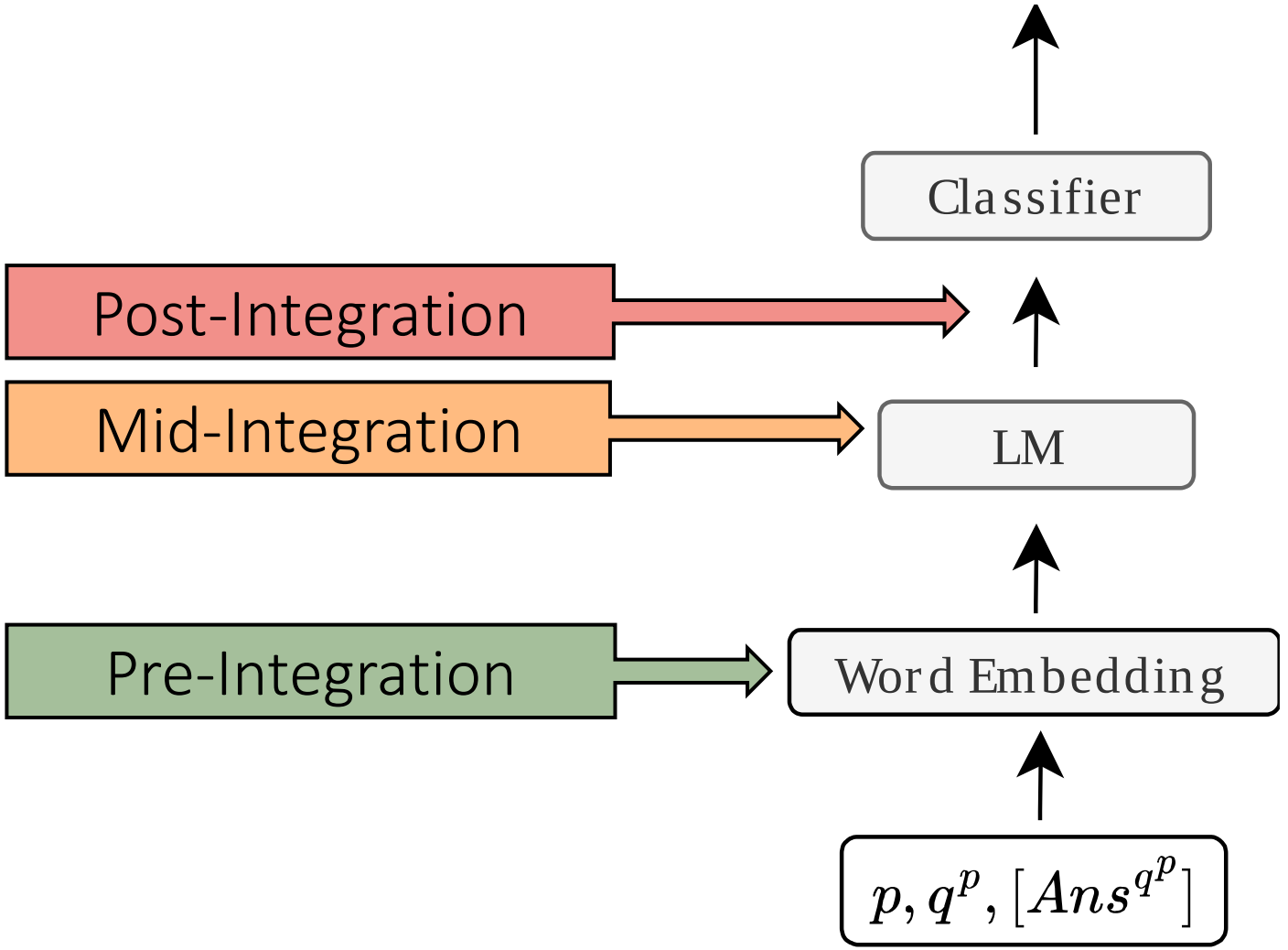
Comprehension Prediction Framework

Task Variations

Eye Movements Representation

Text & Eye Movements Modeling

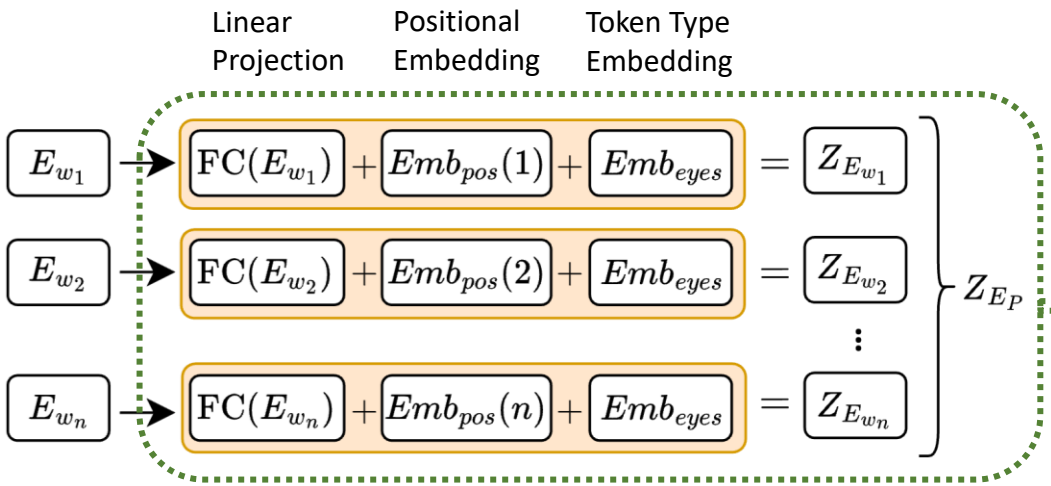
Modeling

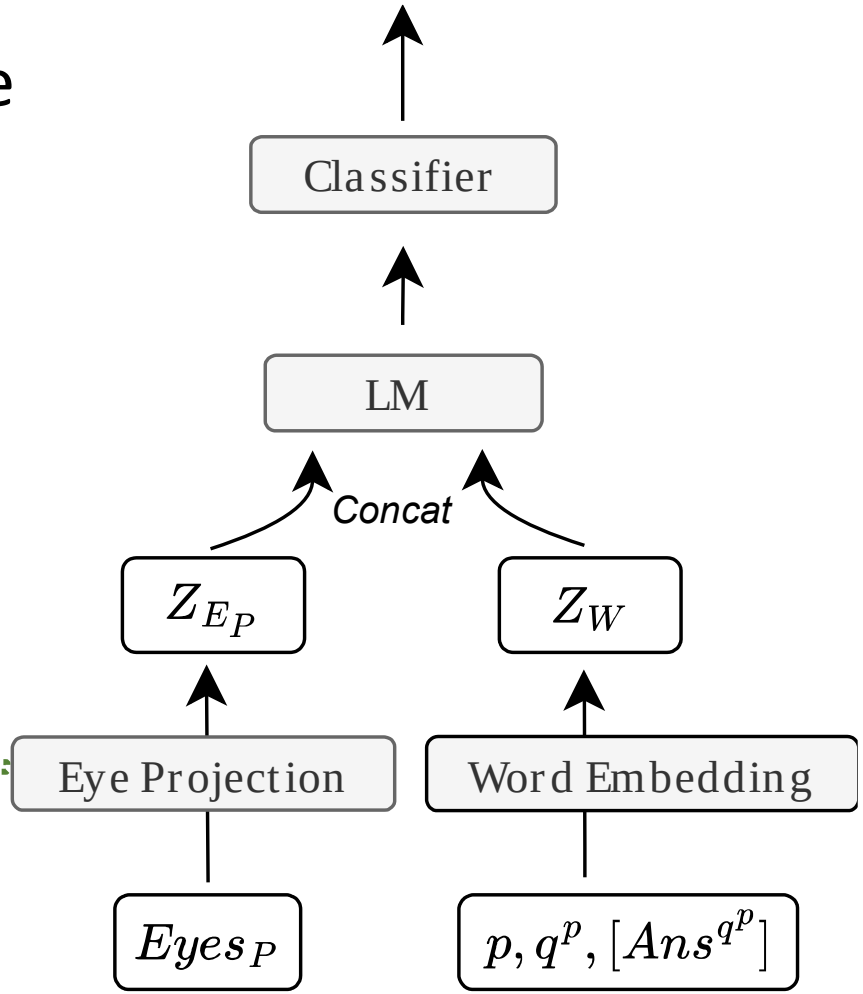

Classifier
Post-Integration
Mid-Integration
LM
Pre-Integration
Word Embedding
$p, q^p, [Ans^{q^p}]$

Pre-integration (RoBERTa-Qeye)

1. Projection to word-representation space
2. Concatenate as embedding-level input

   (soft prompting)

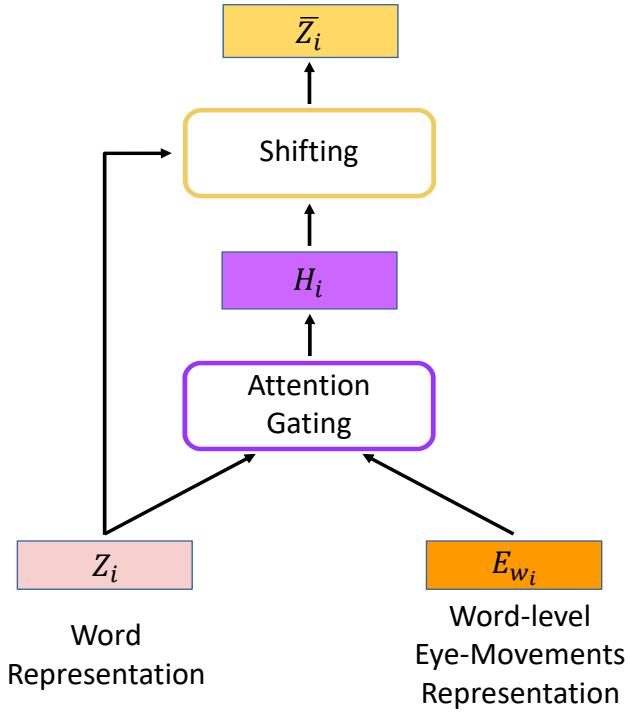


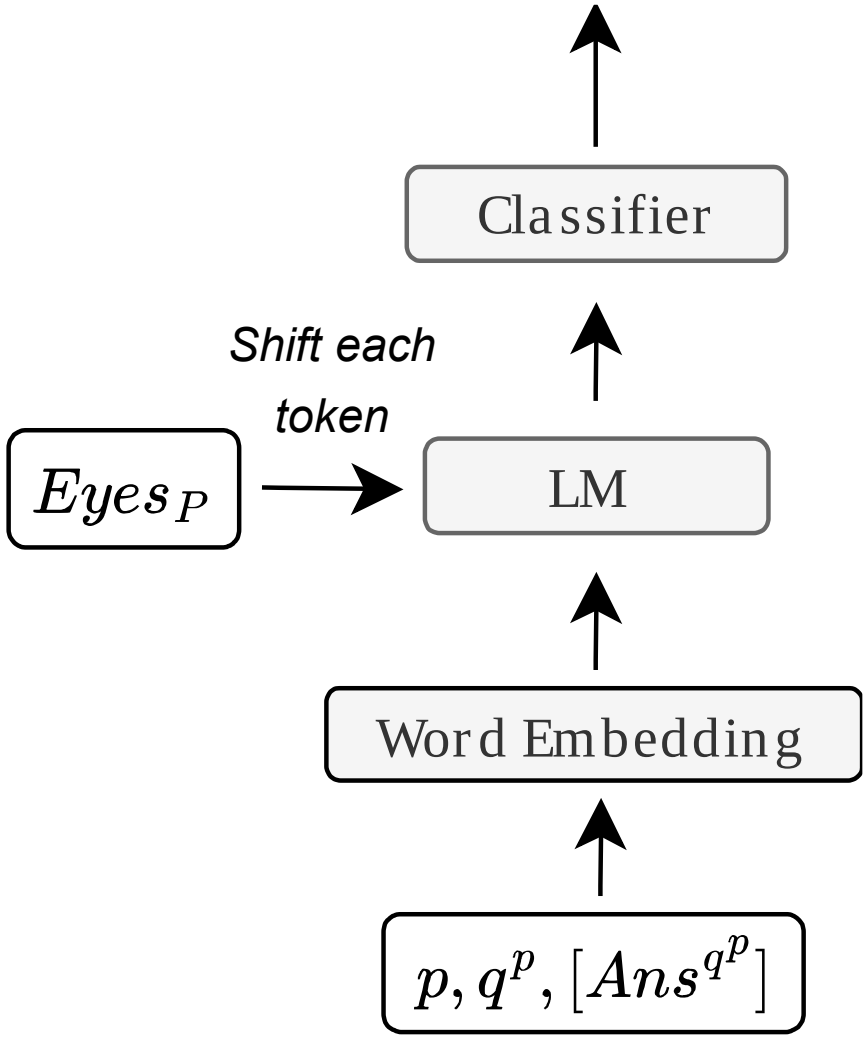


Mid-integration (MAG-Qeye)

Based on Rahman 2020

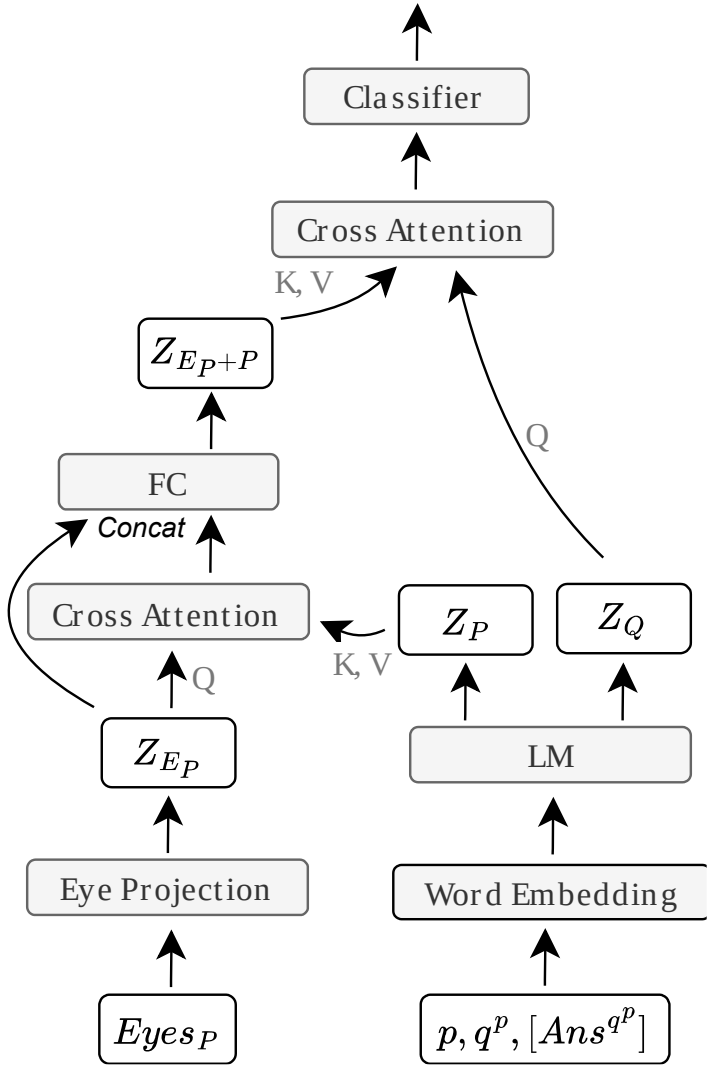
Inject word-level eye-movement representations into mid-layer token representation



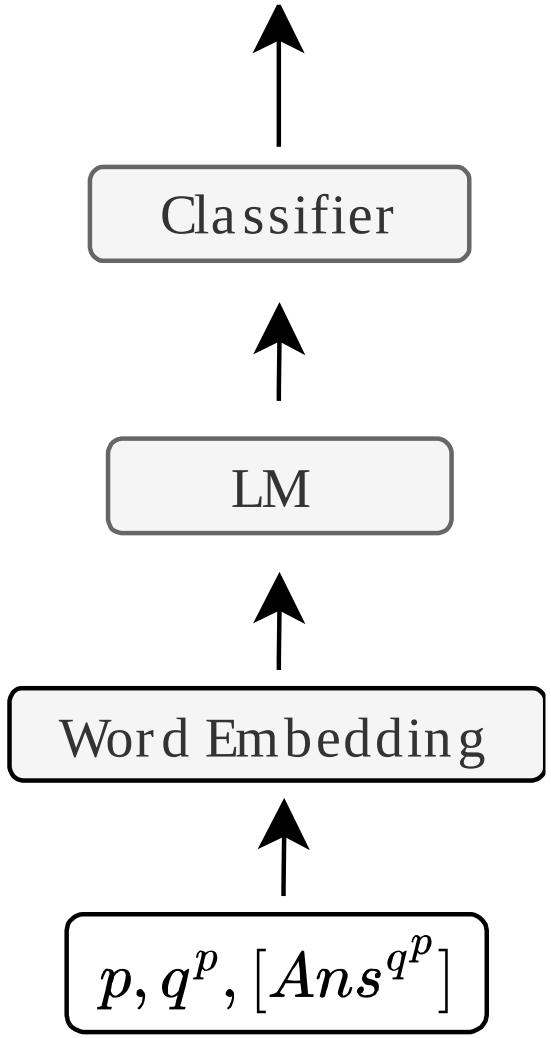


Post-integration (PostFusion-Qeye)

1. Encode Words and eye-movements separately
2. Query Eye-movement representations over text representations
3. Query the result over the question representation



Baseline - Text-only RoBERTa



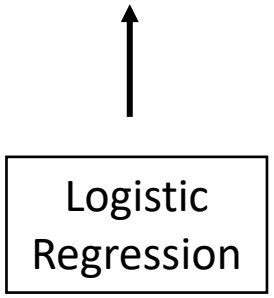
Prior Approaches

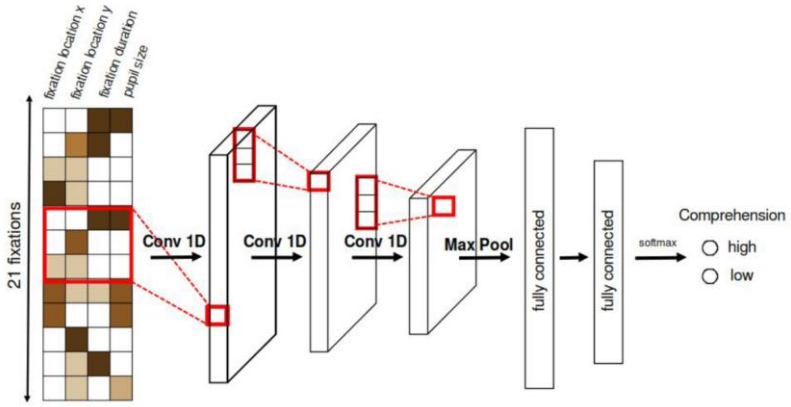
Logistic Regression
(Mézière et al 2023)

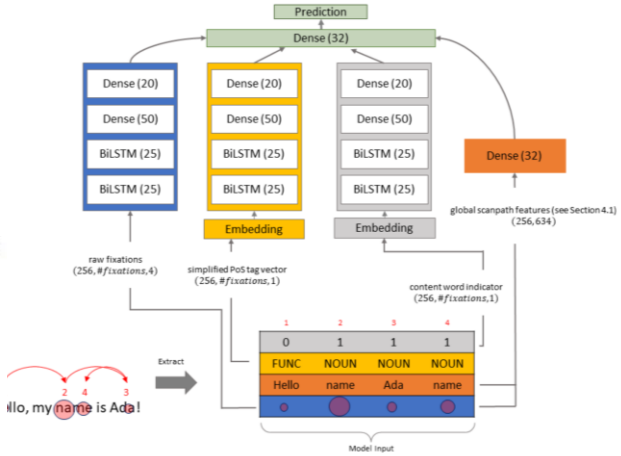
CNN
(Ahn et al 2022)

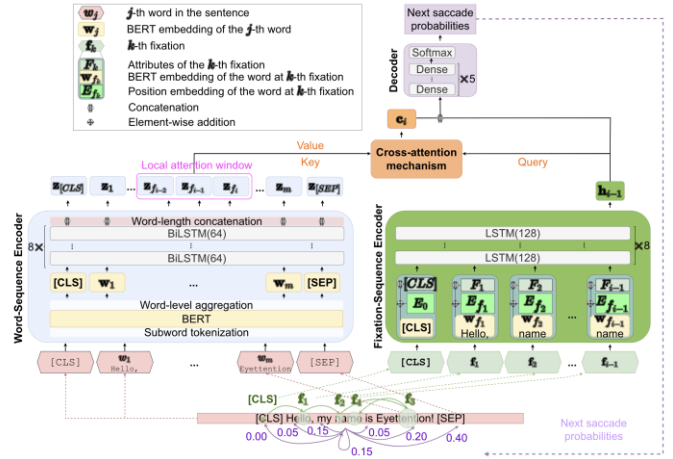
BEyeLSTM
(Reich et al 2022)

Eyettention
(Deng et al 2023)









Individualized Comprehension Prediction

Next Fixation Prediction

Fine-Grained Prediction of Reading Comprehension from Eye Movements

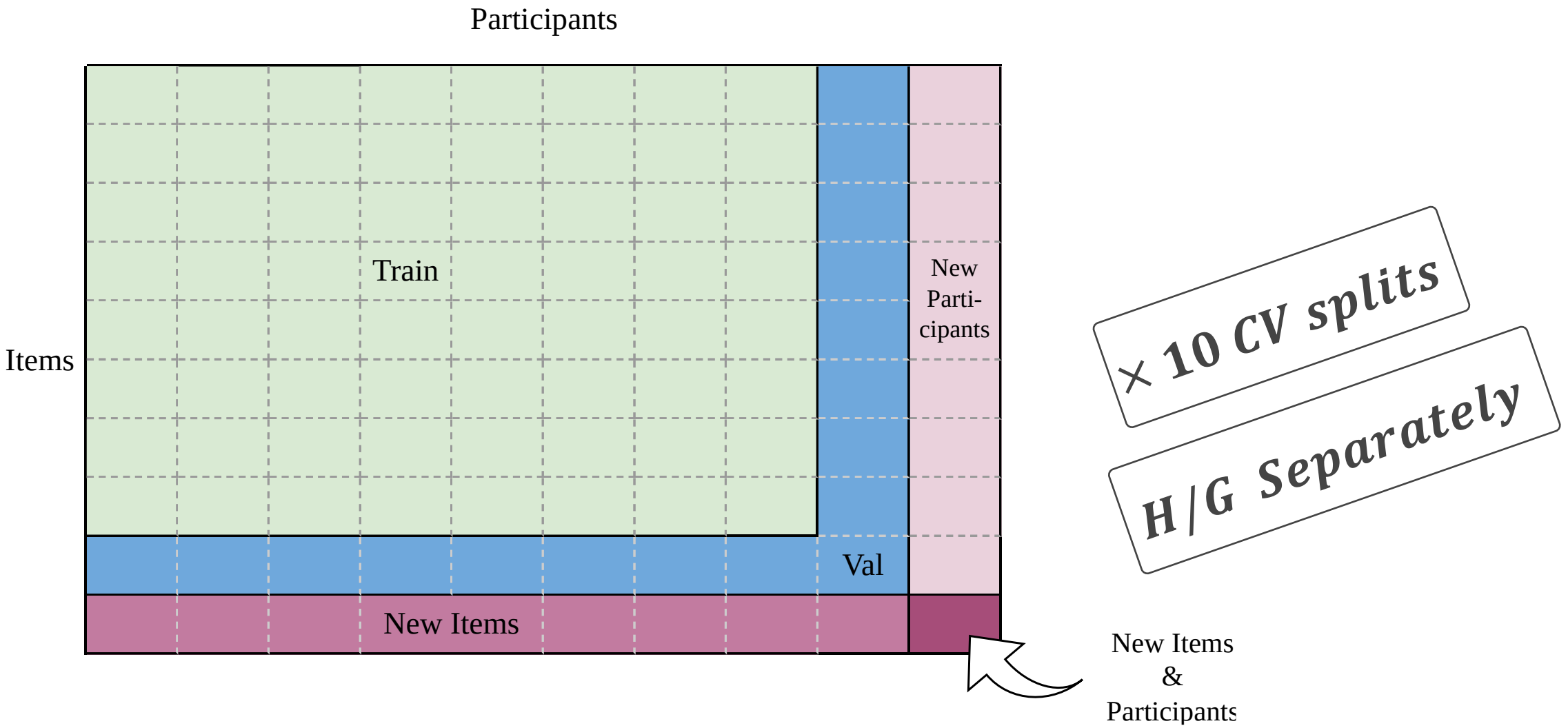
Task Definition

Eye-Tracking Dataset

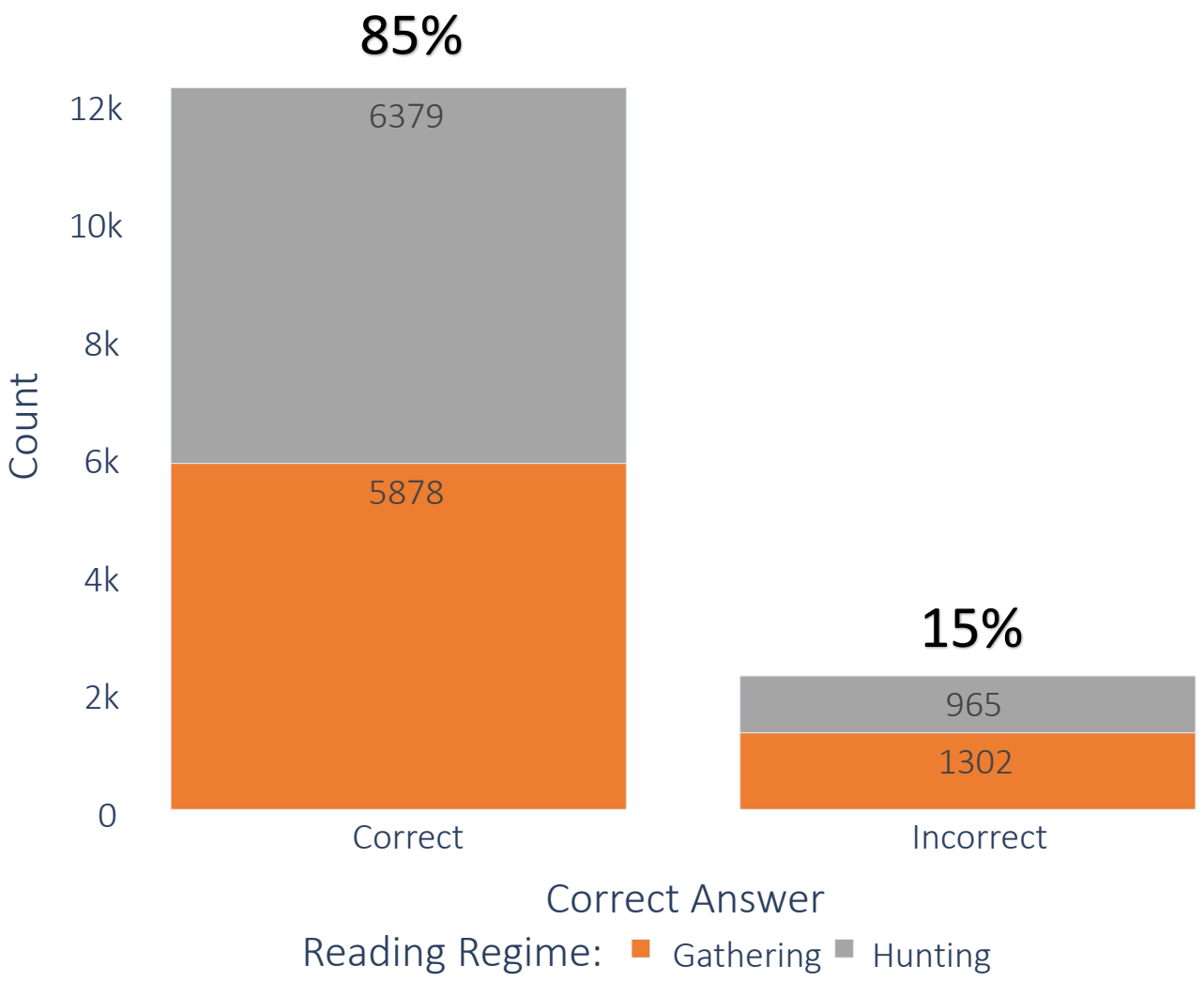
Text & Eye Movements Modeling

Systematic Evaluation

Item-Participant Cross Validation



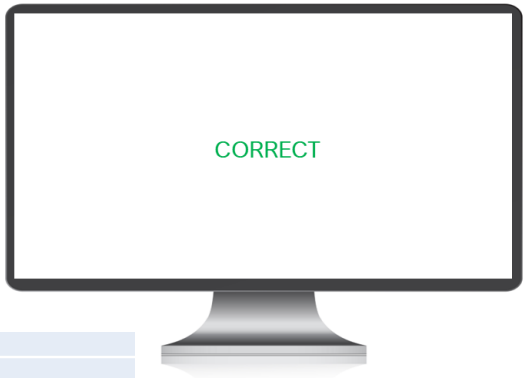
Evaluation Metric

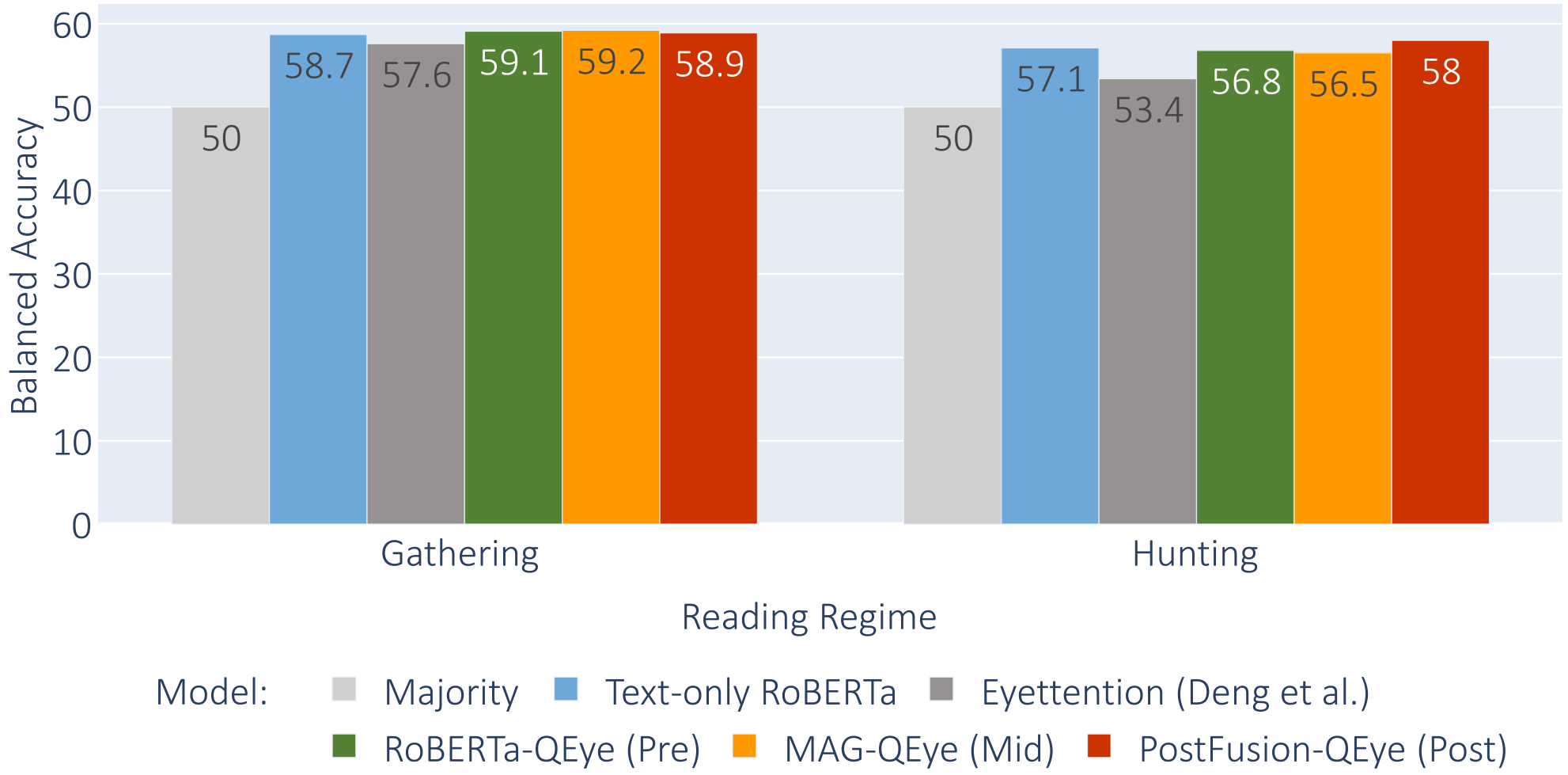


Balanced Accuracy

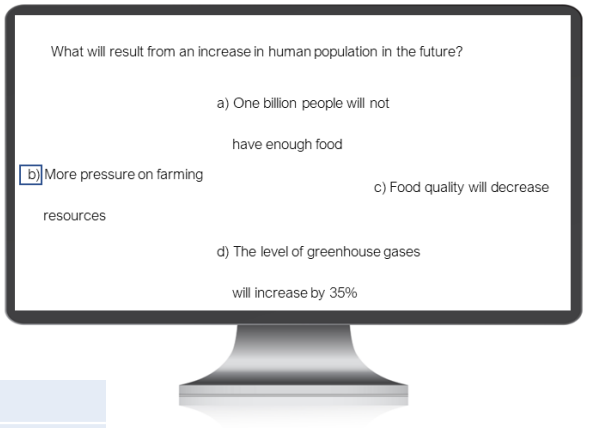
$$\frac{TPR + FPR}{2}$$

Results – Binary Prediction

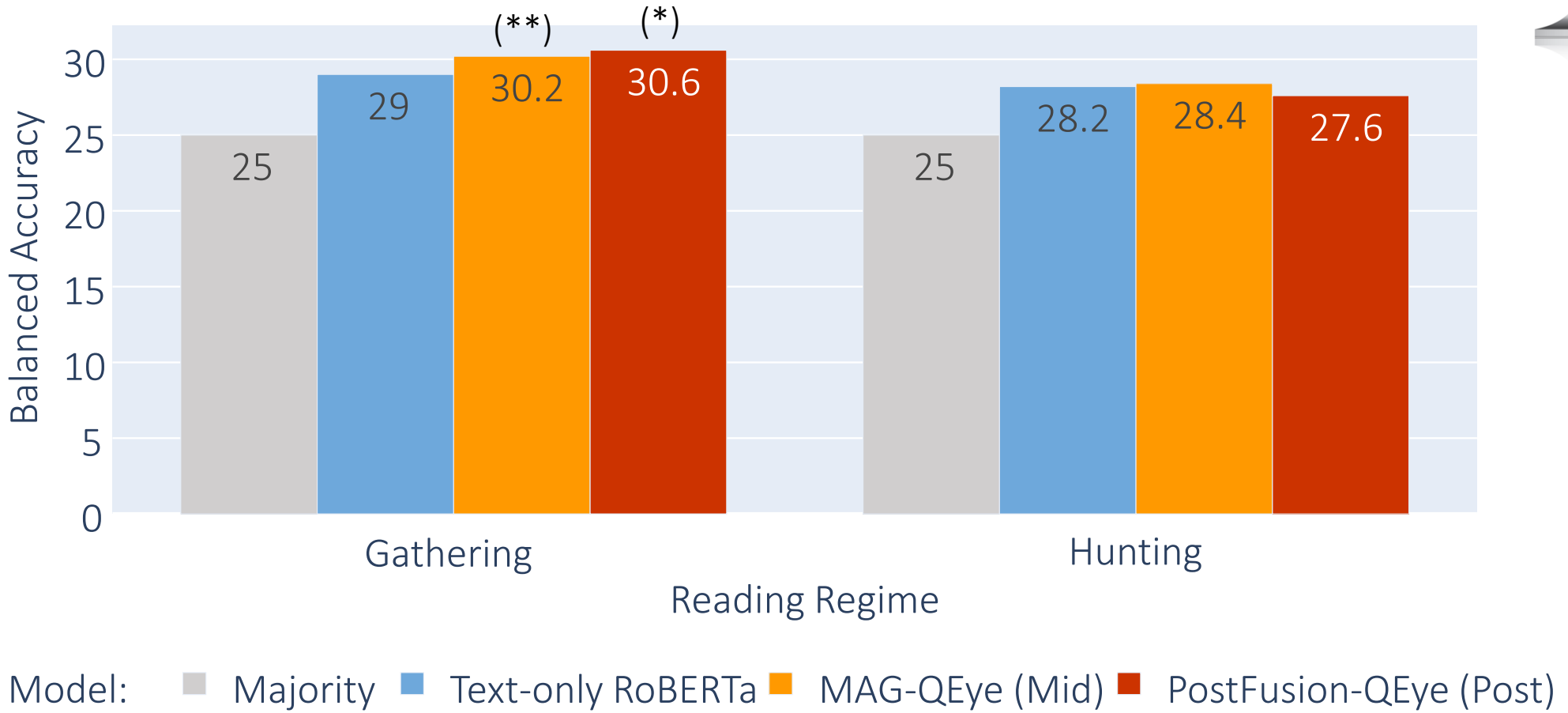




Results – Multiclass Prediction



'(*)' : $p < 0.05$, '(**)' : $p < 0.01$



Takeaways

- Fine-grained reading comprehension prediction is hard
- Moderate improvements over a text-only baseline are achievable.
- Reading goals matter (information-seeking is harder)

LaCC Lab

TECHNION Israel Institute of Technology

Fine-Grained Prediction of Reading Comprehension from Eye Movements

- Fine-grained reading comprehension prediction is hard
- Moderate improvements over a strong text-only baseline are achievable.
- Reading goals matter (information-seeking is harder)

lacclab.github.io

SCAN ME

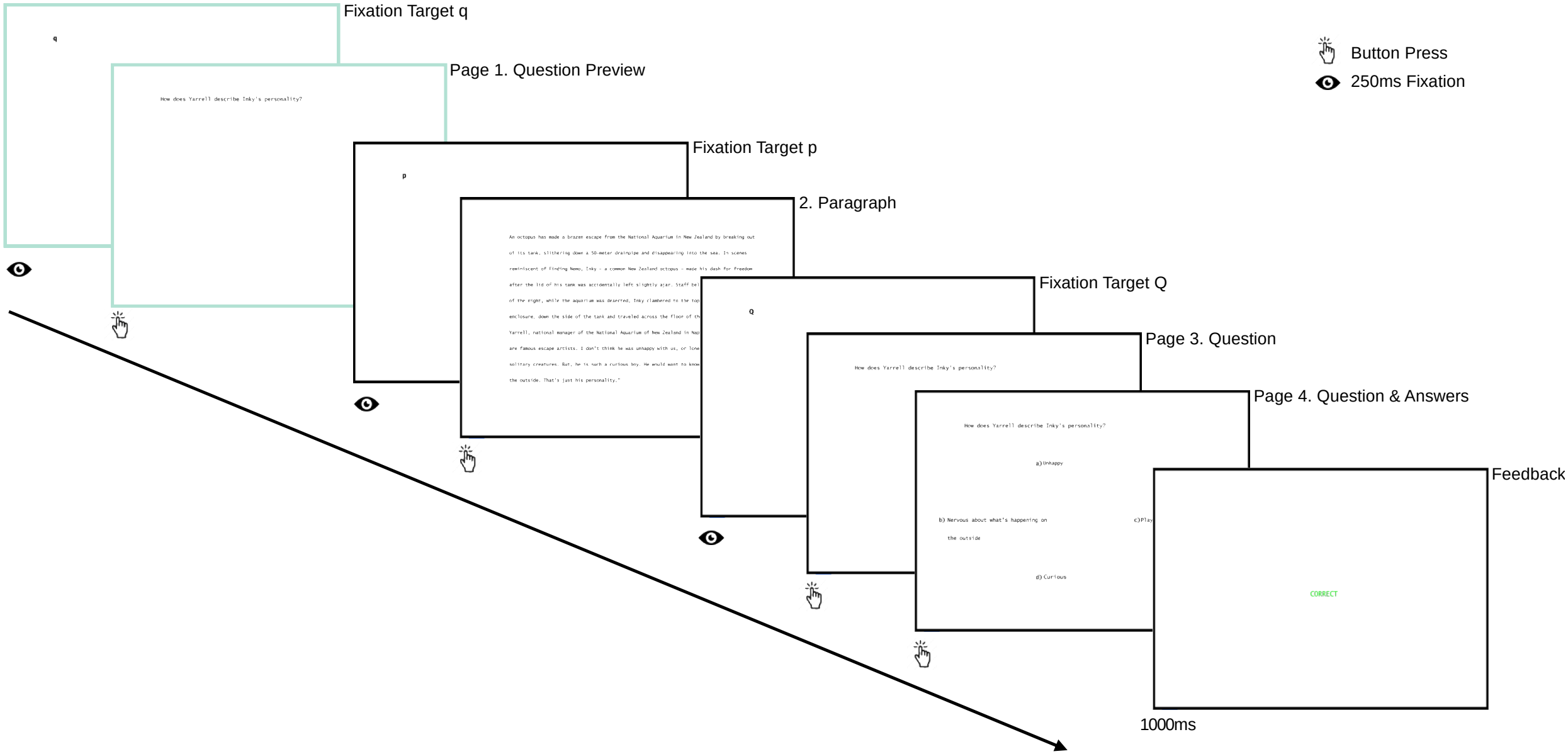
Appendix

Results – Binary Comprehension Prediction

	Binary Reading Comprehension			Ordinary Reading (Gathering)			Information Seeking (Hunting)		
	Model	Gaze Representation	Text Representation	New Item	New Participant	New Item & Participant	New Item	New Participant	New Item & Participant
	Majority	None	None	50.0	50.0	50.0	50.0	50.0	50.0
	Text-only RoBERTa	None	Emb	54.8	63.1	55.2	51.8	63.1	50.5
Prior Work	Log. Reg. (Mézière et al., 2023b)	Global	None	53.3	50.8	53.8	53.2	52.2	52.3
Prior Work	CNN (Ahn et al., 2020a)	Fixations	None	51.0	51.0	51.9	51.4	51.3	49.2
Prior Work	BEyeLSTM (Reich et al., 2022)	Fixations	Ling. Feat.	50.6	55.7	51.1	50.5	55.1	**55.1**
Prior Work	Eyettention (Deng et al., 2023)	Fixations	Emb + Word Len.	54.8	60.4	**57.1**	50.5	56.4	52.3
(Pre)	RoBERTa-QEye	Words	Emb + Ling. Feat.	**55.5**	63.5	52.1	50.5	**63.8**	51.0
(Pre)	RoBERTa-QEye	Fixations	Emb + Ling. Feat.	53.3	61.3	**57.1**	50.3	60.3	50.8
(Mid)	MAG-QEye	Words	Emb + Ling. Feat.	54.8	**64.1***	53.8	52.5	62.3	51.3
(Post)	PostFusion-QEye	Fixations	Emb + Ling. Feat.	54.8	63.5	55.0	**53.8***	62.7	53.8

Results - Multiclass

	Multiple-Choice Reading Comprehension			Ordinary Reading (Gathering)				Information Seeking (Hunting)			
	Model	Gaze Representation	Text Representation	New Item	New Participant	New Item & Participant	All	New Item	New Participant	New Item & Participant	All
	Majority	None	None	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0
	Text-only RoBERTa	None	Emb	25.3	**33.0**	25.2	29.0	25.0	**31.7**	24.8	28.2
(Mid)	MAG-QEye	Words	Emb + Ling. Feat.	27.9***	32.5	30.4***	30.2**	26.8	30.0	**29.0**	**28.4**
(Post)	PostFusion-QEye	Fixations	Emb + Ling. Feat.	**29.4****	31.7	**32.9***	**30.6***	**27.5***	27.9	26.7	27.6


Fixation Target q
Page 1. Question Preview
How does Yarrell describe Inky's personality?
Button Press
250ms Fixation
Fixation Target p
2. Paragraph
An octopus has made a brazen escape from the National Aquarium in New Zealand by breaking out
of its tank, slithering down a 50-meter drainpipe and disappearing into the sea. In scenes
reminiscent of Finding Nemo, Inky - a common New Zealand octopus - made his dash for freedom
after the lid of his tank was accidentally left slightly ajar. Staff bel
of the night, while the aquarium was deserted, Inky clambered to the top
enclosure, down the side of the tank and traveled across the floor of th
Yarrell, national manager of the National Aquarium of New Zealand in Nap
are famous escape artists. I don't think he was unhappy with us, or lone
solitary creatures. But, he is such a curious boy. He would want to know
the outside. That's just his personality."
Fixation Target Q
Page 3. Question
How does Yarrell describe Inky's personality?
Page 4. Question & Answers
How does Yarrell describe Inky's personality?
a) Unhappy
b) Nervous about what's happening on
the outside
c) Play
d) Curious
Feedback
CORRECT
1000ms

OneStop – Eye Tracking Experiment

Two between-subjects conditions:

- **Hunting** (information-seeking)

👀 → **Question** → Passage → Question → Answers

- **Gathering** (ordinary reading)

👀 → Passage → **Question** → Answers

Task Effects in Reading

In	the	next	30	years,	the	planet's	human	population	will
increase	to	nine	billion.	Already	one	billion	people	do	not
get	enough	food.	The	increase	will	put	more	pressure	on
agricultural	land,	water,	forests,	fisheries	and	resources,	and	also	food
and	energy	supplies.	The	cost	of	meat	is	increasing	–
it	costs	more	money	now,	but	also	people	have	to
destroy	a	lot	of	rainforest	to	make	fields	or	to
grow	food	for	cows.	Cows	also	make	methane.	The	farming
of	cows,	pigs	and	sheep	makes	very	large	amounts	of
greenhouse	gases	–	35%	of	the	planet's	methane,	65%	of
its	nitrous	oxide	and	9%	of	the	carbon	dioxide.	

Adapted from Malmaud J, Levy R, Berzak Y., Bridging Information-Seeking Human Gaze and Machine Reading Comprehension, ACL 2020

Task Effects in Reading

Gathering (no question preview)

In	the	next	30	years,	the	planet's	human	population	will
increase	to	nine	billion.	Already	one	billion	people	do	not
get	enough	food.	The	increase	will	put	more	pressure	on
agricultural	land,	water,	forests,	fisheries	and	resources,	and	also	food
and	energy	supplies.	The	cost	of	meat	is	increasing	–
it	costs	more	money	now,	but	also	people	have	to
destroy	a	lot	of	rainforest	to	make	fields	or	to
grow	food	for	cows.	Cows	also	make	methane.	The	farming
of	cows,	pigs	and	sheep	makes	very	large	amounts	of
greenhouse	gases	–	35%	of	the	planet's	methane,	65%	of
its	nitrous	oxide	and	9%	of	the	carbon	dioxide.	

Hunting (with question preview)

Q: What will result from an increase in human population in the future?

In	the	next	30	years,	the	planet's	human	population	will
increase	to	nine	billion.	Already	one	billion	people	do	not
get	enough	food.	The	increase	will	put	more	pressure	on
agricultural	land,	water,	forests,	fisheries	and	resources,	and	also	food
and	energy	supplies.	The	cost	of	meat	is	increasing	–
it	costs	more	money	now,	but	also	people	have	to
destroy	a	lot	of	rainforest	to	make	fields	or	to
grow	food	for	cows.	Cows	also	make	methane.	The	farming
of	cows,	pigs	and	sheep	makes	very	large	amounts	of
greenhouse	gases	–	35%	of	the	planet's	methane,	65%	of
its	nitrous	oxide	and	9%	of	the	carbon	dioxide.	

Task Effects in Reading

Gathering (no question preview)

In	the	next	30	years,	the	planet's	human	population	will
increase	to	nine	billion.	Already	one	billion	people	do	not
get	enough	food.	The	increase	will	put	more	pressure	on
agricultural	land,	water,	forests,	fisheries	and	resources,	and	also	food
and	energy	supplies.	The	cost	of	meat	is	increasing	–
it	costs	more	money	now,	but	also	people	have	to
destroy	a	lot	of	rainforest	to	make	fields	or	to
grow	food	for	cows.	Cows	also	make	methane.	The	farming
of	cows,	pigs	and	sheep	makes	very	large	amounts	of
greenhouse	gases	–	35%	of	the	planet's	methane,	65%	of
its	nitrous	oxide	and	9%	of	the	carbon	dioxide.	

Hunting (with question preview)

Q: Which of these is mentioned as a cost of meat?

In	the	next	30	years,	the	planet's	human	population	will
increase	to	nine	billion.	Already	one	billion	people	do	not
get	enough	food.	The	increase	will	put	more	pressure	on
agricultural	land,	water,	forests,	fisheries	and	resources,	and	also	food
and	energy	supplies.	**The**	**cost**	**of**	**meat**	**is**	**increasing**	**–**
it	**costs**	**more**	**money**	**now,**	**but**	**also**	**people**	**have**	**to**
destroy	**a**	**lot**	**of**	**rainforest**	**to**	**make**	**fields**	**or**	**to**
grow	**food**	**for**	**cows.**	Cows	also	make	methane.	The	farming
of	cows,	pigs	and	sheep	makes	very	large	amounts	of
greenhouse	gases	–	35%	of	the	planet's	methane,	65%	of
its	nitrous	oxide	and	9%	of	the	carbon	dioxide.	

Malmaud et. al. found shorter reading times in Hunting, out of the critical span

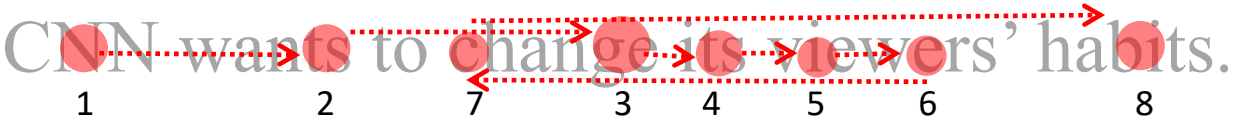
How do people read?

Fixations

How do people read?

Saccades

How do we model reading?



- Raw millisecond-level X Y screen coordinates
- Fixation / Scanpath measures
- Word-level measures
- Global

OneStopQA – STARC Annotation Framework (Berzak 2020)

- Multiple-choice question-answering dataset
- Structured distractors

Critical span:

Information essential for answering the question.

In the next 30 years, the planet's human population will increase to nine billion. Already one billion people do not get enough food. The increase will put more pressure on agricultural land, water, forests, fisheries and resources, and also food and energy supplies. The cost of meat is increasing – it costs more money now, but also people have to destroy a lot of rainforest to make fields or to grow food for cows. Cows also make methane. The farming of cows, pigs and sheep makes very large amounts of greenhouse gases – 35% of the planet's methane, 65% of its nitrous oxide and 9% of the carbon dioxide.

Q: What will result from an increase in human population in the future?

A) More pressure on farming resources ✓

B) One billion people will not have enough food

C) The level of greenhouse gases will increase by 35%

D) Food quality will decrease

Distractor span

Degree of comprehension

What physiological or behavioral cues might reveal comprehension level in real time?