



# The Data Miners

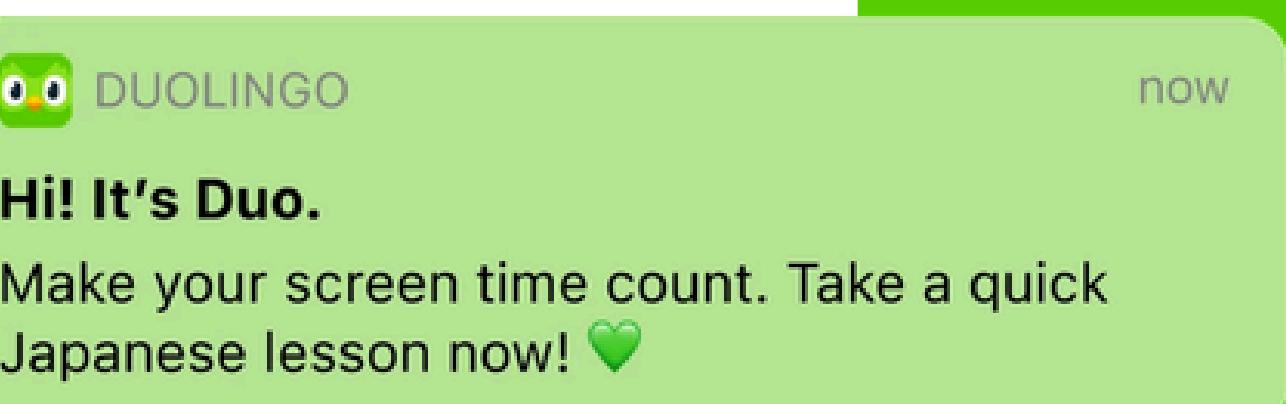
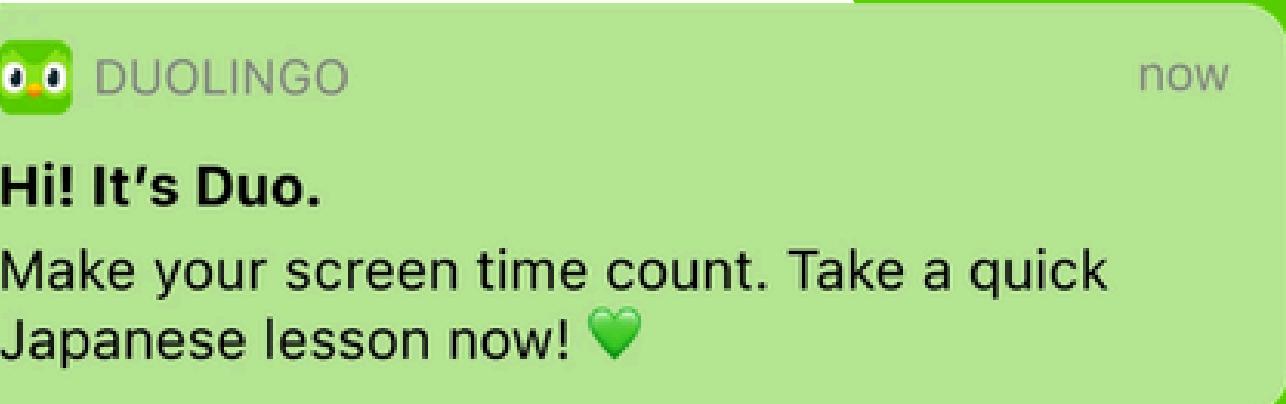
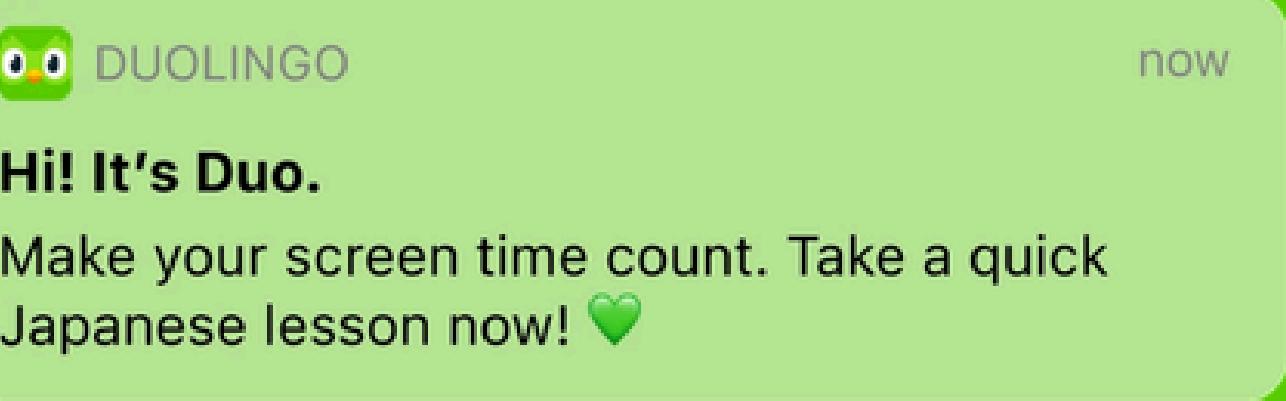
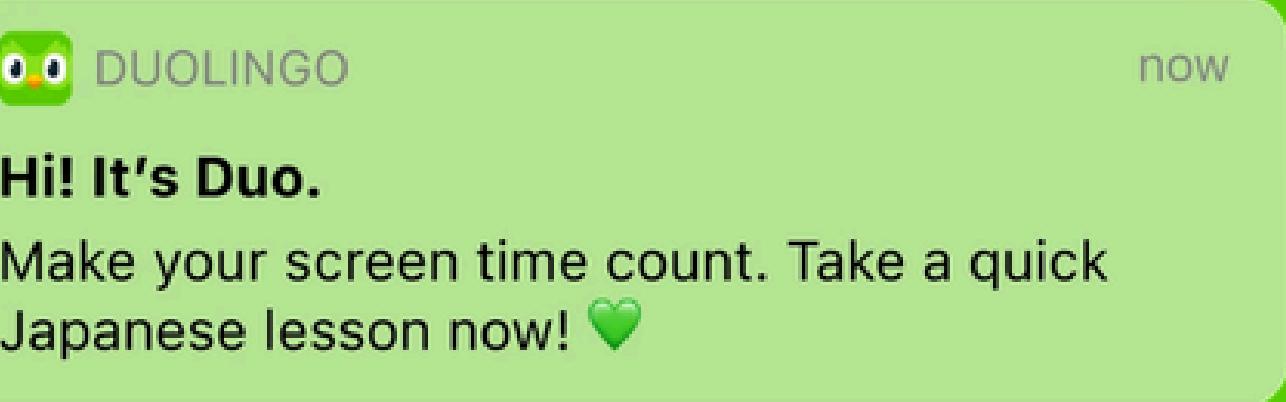
## Members:

Alexander Thomas Savvides  
Hafsa Azeb  
Ha Van Pham  
Garima Sharda  
Timon Knol

# The Da

## Members:

Alexander T  
Hafsa Azeb  
Ha Van Phan  
Garima Shar  
Timon Knol



# The Data Miners



DUOLINGO

**Hi! It's Duo.**

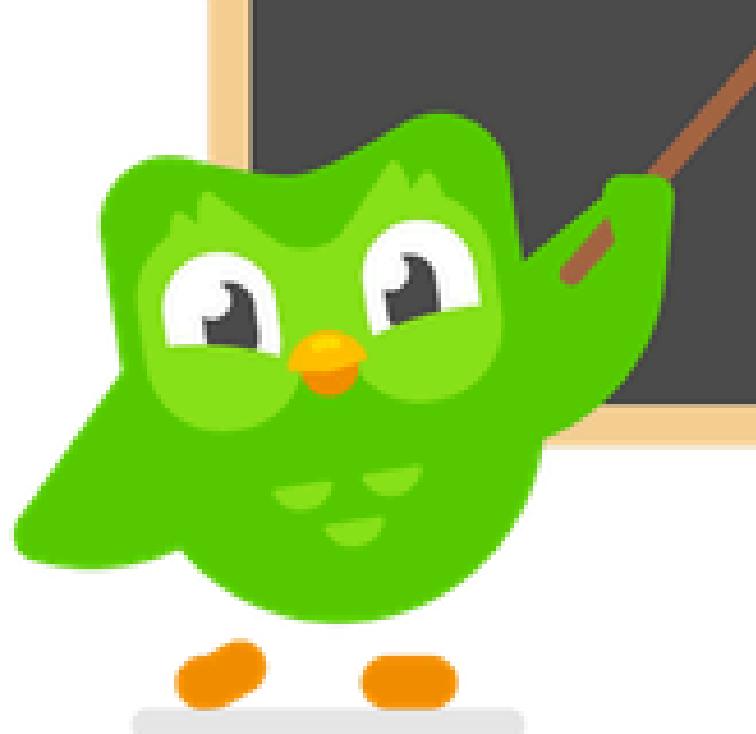
Good luck with you presentation!

now

...Khalid ...nas Savvides  
Hafsa Azeb  
Ha Van Pham  
Garima Sharda  
Timon Knol

# TABLE OF CONTENTS

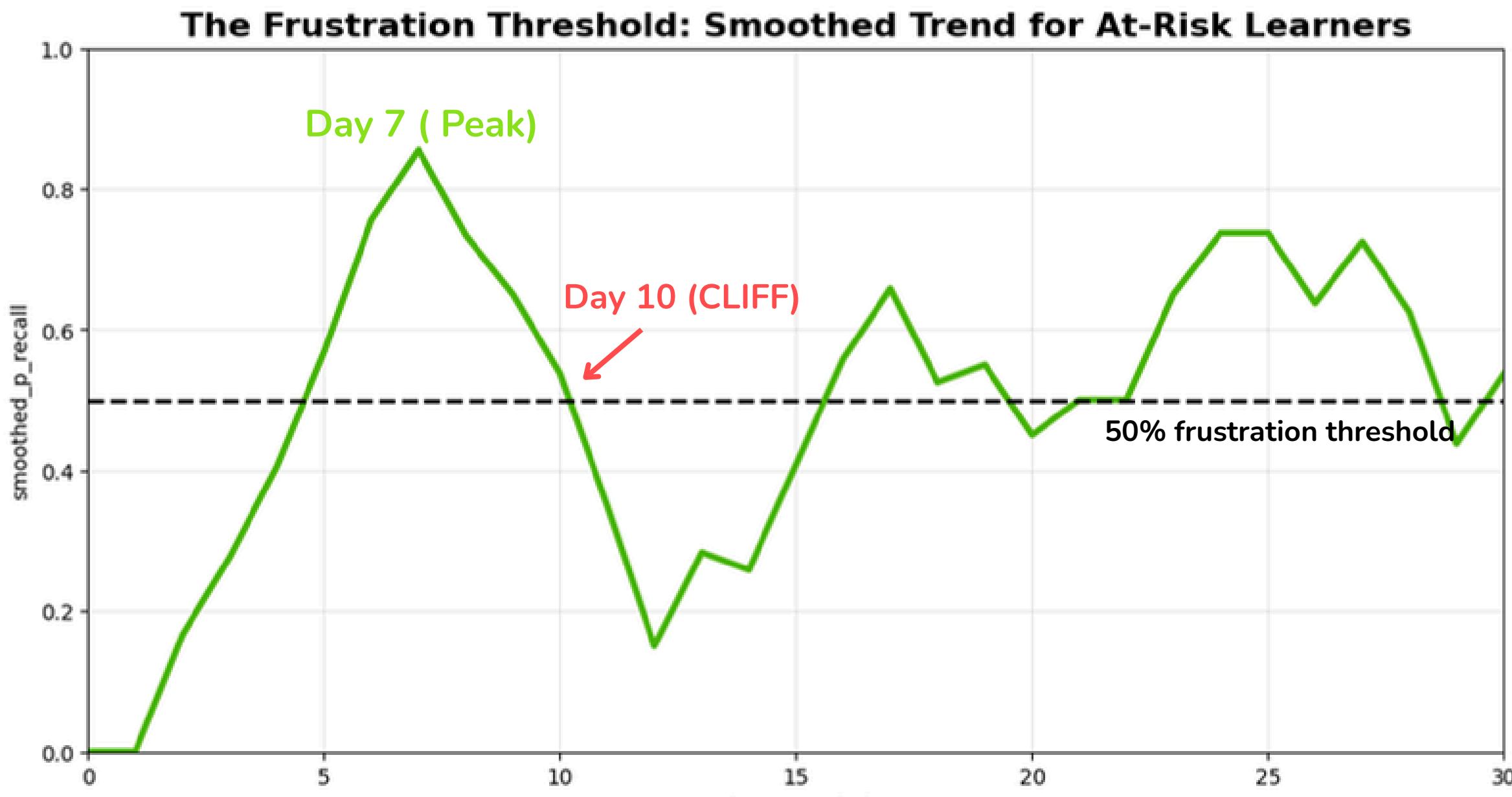
UNIT	Phase	Status
1	User Engagement	UNLOCKED
2	Quality of Education	LOADING
3	Quality of Content	LOADING



# Unit 1: User Engagement



# The "Cognitive Cliff"



## Key Takeaways

Day 1-7: False sense of security.  
Recall peaks.

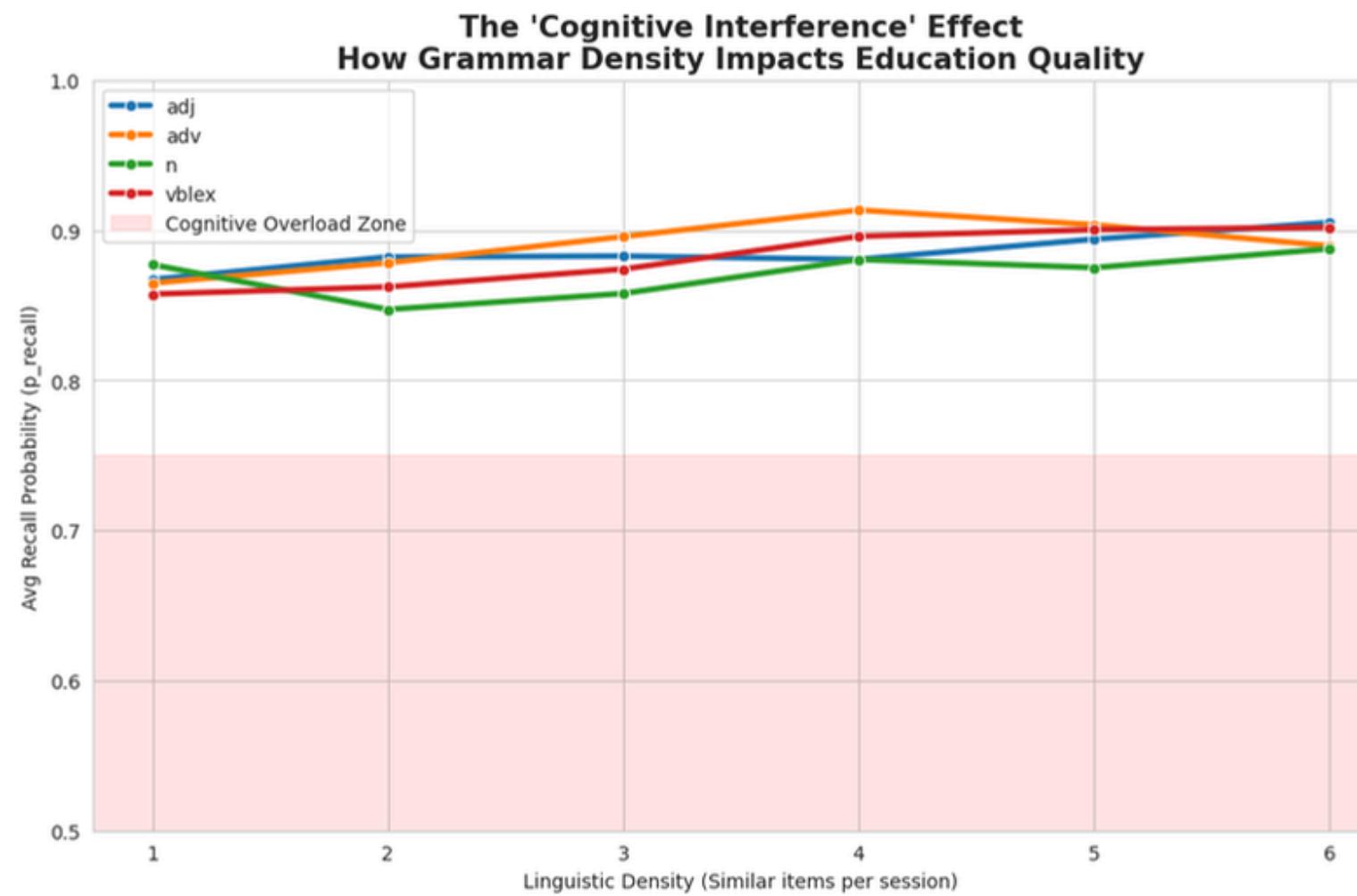
Day 10: The Cognitive Cliff.  
Recall plummets to 20% by Day 12.

## Solution

**DON'T Wait** until Day 12 to prompt review.  
The user loses the information and the motivation to continue.

**Intervention Required:**  
Day 9 Engagement Buffer.

## Users are Resilient



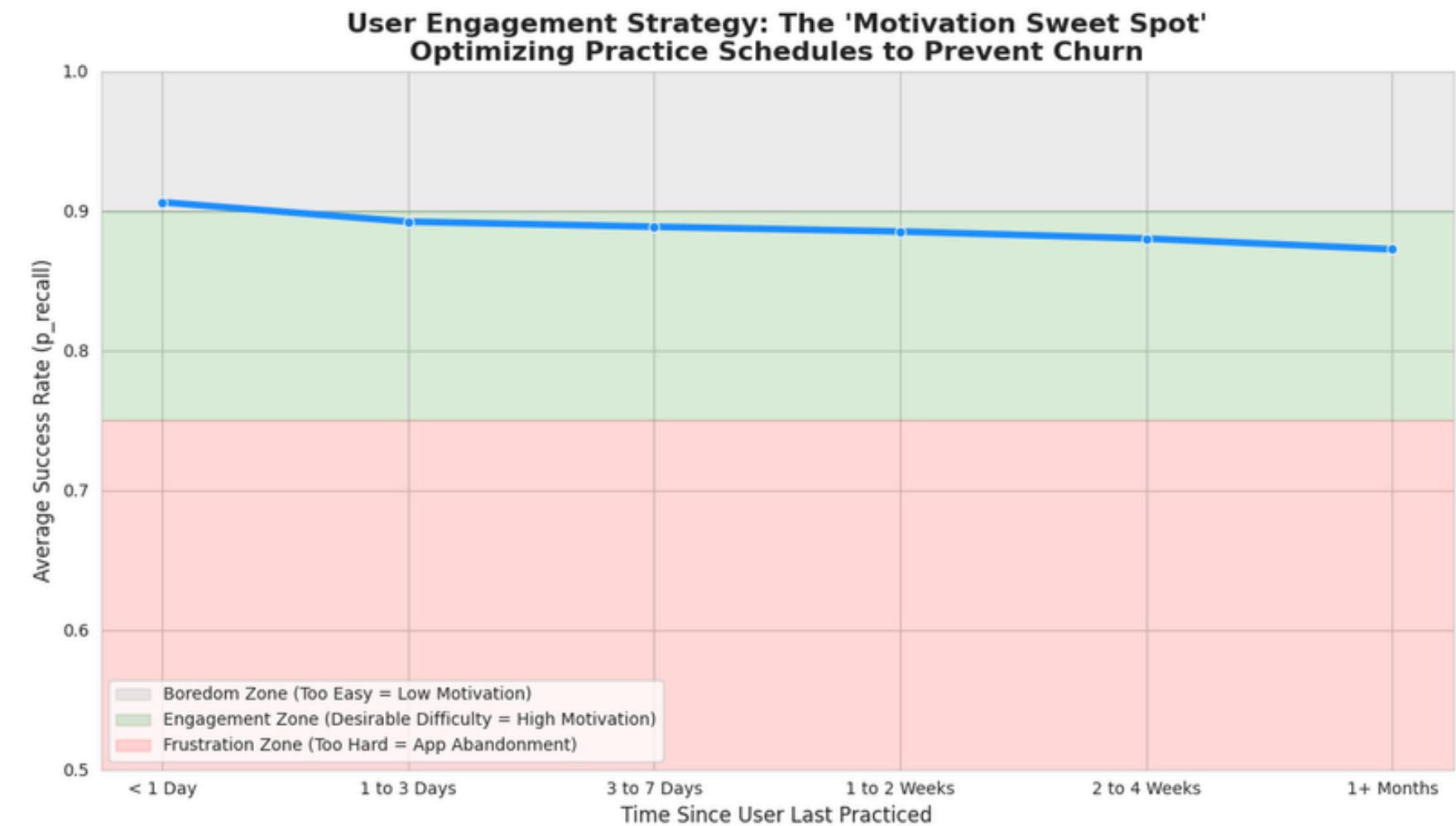
**'Cognitive Interference'**: Traditional education warns against introducing too many similar concepts at once.

**Our data completely disagrees.**

- Recall stays rock-solid (**85-90%**) whether a user sees 1 or 6 verbs.
- Nouns, adjectives, and verbs all show identical stability.

## Mythbuster:

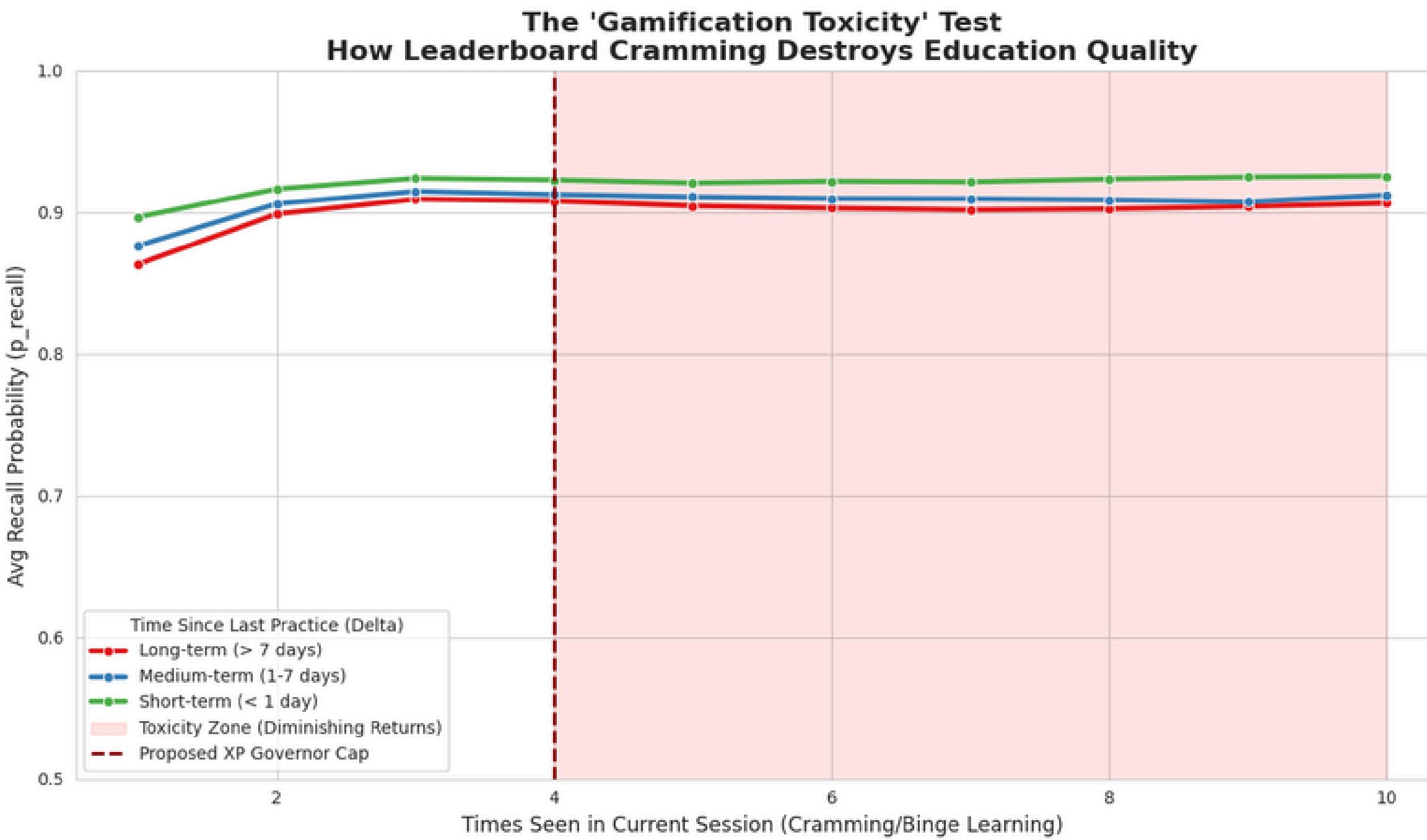
## Users get Bored



**Assumption:** Stepping away for 1 month, leads to forgetting everything. **Result:** crash into the red "Frustration Zone," and quit.

- After 1 month away from a word, **users still get it right 87%** of the time.
- The app is keeping users hovering near the '**Boredom Zone**'.
- Data proves **Duolingo is playing it too safe**.

# Zero ROI after the 4th Repetition



## Key Insights

- Reps 1-3: Memory benefit improves.
- Reps 4+: Benefit completely flatlines.

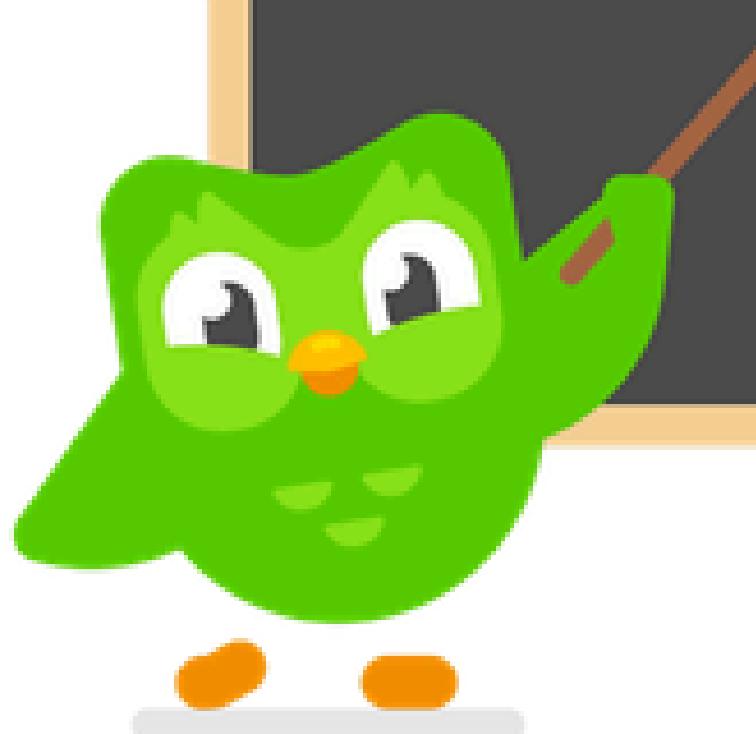
## Solution: Target lower amounts of repetition

Users stops earning Leaderboard points after repeating a word 4 times.



# TABLE OF CONTENTS

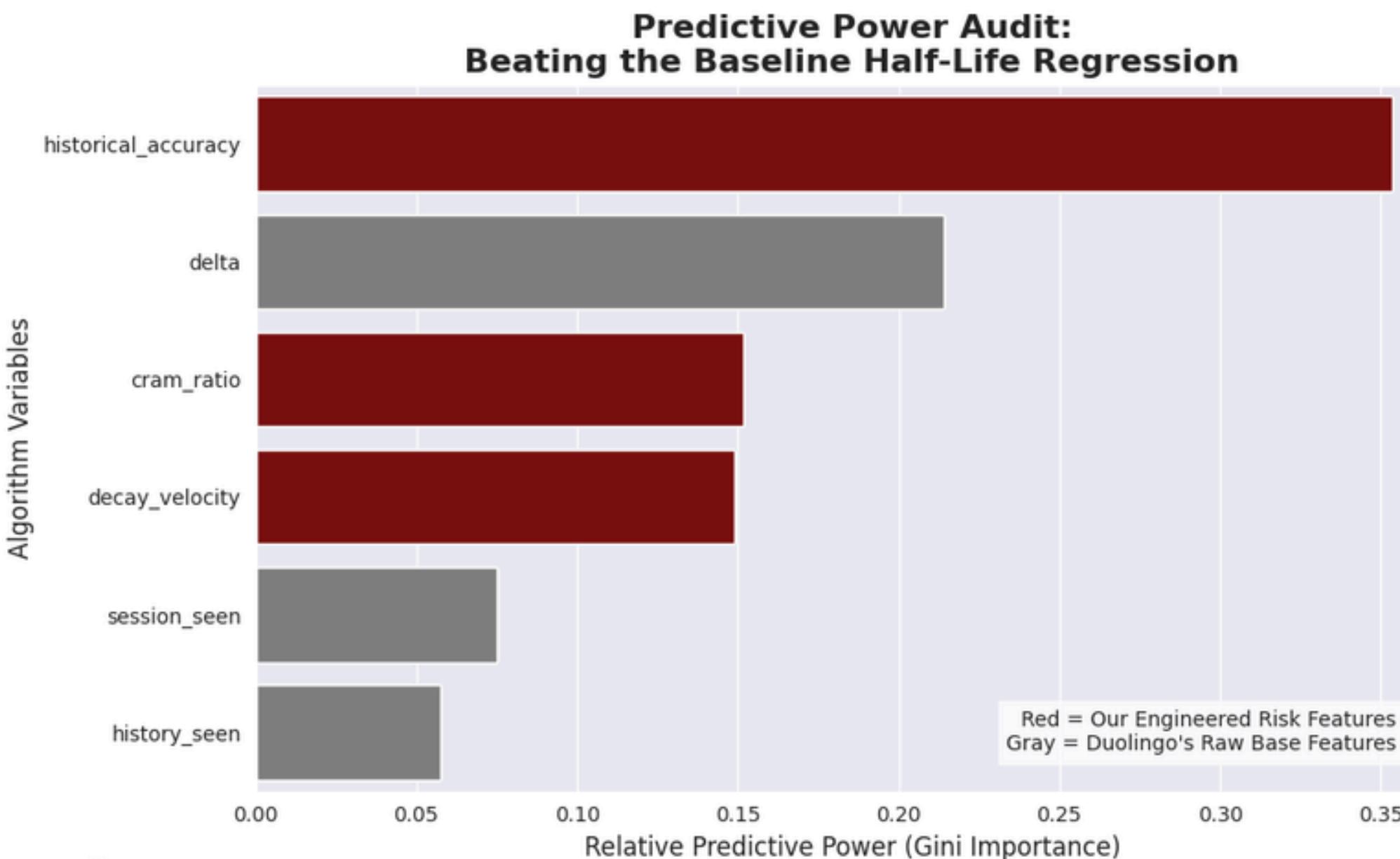
UNIT	Phase	Status
1	User Engagement	COMPLETE
2	Quality of Education	UNLOCKED
3	Quality of Content	LOADING



# Unit 2: Quality of Education



# What Really Predicts Memory: Accuracy, Time, or Behavior?



## Key Insights

### Confidence > Recency

→ Prioritize performance history over time decay

### Time is Secondary

→ Avoid over-weighting Half-Life scheduling

### Volume ≠ Learning

→ Raw repetition counts add little predictive value

### Behavior > Exposure

→ Optimize how users learn, not how often they click





## Duolingo Behavioral Routing Engine

Adjust the metrics below to diagnose the learner's state, forecast recall via Random Forest, and prescribe an automated product intervention.

Live Routing Engine

Global Analytics (13M Rows)

Methodology

### User Context & Metrics

Total Times Word Seen



140

0

Times Seen Today (Session)

5

0

Time Elapsed Since Last Review (Hours)

24

0

### ML Engine Output

Detected Persona

Recommended Product Action

RF Predicted Recall Probability

Metric Visualization

ANALYZE USER STATE & PREDICT RECALL

# TABLE OF CONTENTS

UNIT	Phase	Status
1	User Engagement	COMPLETE
2	Quality of Education	COMPLETE
3	Quality of Content	UNLOCKED



# Unit 3: Quality of Content



# Duolingo Grading Simulator

## The Problem

- Duolingo learners can submit many correct translations for one English sentence
- Traditional grading systems only accept one "correct" answer
- How do you grade fairly when there are 10+ valid translations?

## The Solution:

- Create a weighted grading system:
- Track how often real learners use each translation
  - Assign **confidence scores** based on frequency
  - Grade translations by how **common/natural** they are, not just correctness

Confidence Level	Frequency	Stars	Points
High	>10% of learners	★★★	10
Medium	1-10% of learners	★★	8
Low	<1% of learners	★	6



# OUR GRADING SIMULATOR



## Choose your language

pt Portuguese

ja Japanese

vn Vietnamese

hu Hungarian

kr Korean



## Choose your exercise

can i have a bottle of water?

i do not want to see you suffer.

my brother had earned more than my father.

the book shows brazil.

he came home yesterday.

i finally got an answer.

she did not follow the writer.

they will be able to exercise.

he is afraid of the dog.

i had arrived at school.

she thinks he's a nice guy.

we made pasta with fish last week.

he is older than her.

i touch my plate.

thank you and you are welcome.

we will see a big castle.

i am yours.

it will rain this saturday.

the attack came from the right.

what do you want to drink?

## Explore Learner Translations

Translation	Confidence %	Ranking	Total points
a férfi a lánynál idősebb.	0.33	★	6
a férfi a lánynál öregebb.	0.33	★	6
a férfi a nőnél idősebb.	0.33	★	6
a férfi a nőnél öregebb.	0.33	★	6
a férfi idősebb a lánynál.	0.33	★	6
a férfi idősebb a nőnél.	1.60	★★	8
a férfi idősebb nála.	0.33	★	6
a férfi idősebb önála.	0.33	★	6
a férfi idősebb töle.	0.33	★	6
a férfi idősebb, mint a lány.	0.33	★	6
a férfi idősebb, mint a nő.	2.86	★★	8
a férfi idősebb, mint ő.	0.33	★	6
a férfi nála idősebb.	0.33	★	6
a férfi nála öregebb.	0.33	★	6
a férfi önála idősebb.	0.33	★	6
a férfi önála öregebb.	0.33	★	6
a férfi öregebb a lánynál.	0.33	★	6
a férfi öregebb a nőnél.	0.79	★	6
a férfi öregebb nála.	0.33	★	6
a férfi öregebb önála.	0.33	★	6
a férfi öregebb töle.	0.33	★	6
a férfi öregebb, mint a lány.	0.33	★	6
a férfi öregebb, mint a nő.	1.56	★★	8
a férfi öregebb, mint ő.	0.33	★	6
a fiú a lánynál idősebb.	0.33	★	6
a fiú a lánynál öregebb.	0.33	★	6
a fiú a nőnél idősebb.	0.33	★	6
a fiú a nőnél öregebb.	0.33	★	6
a fiú idősebb a lánynál.	1.14	★★	8
a fiú idősebb a nőnél.	0.33	★	6
<b>Total</b>	1,994.89		21406

# Key Insights

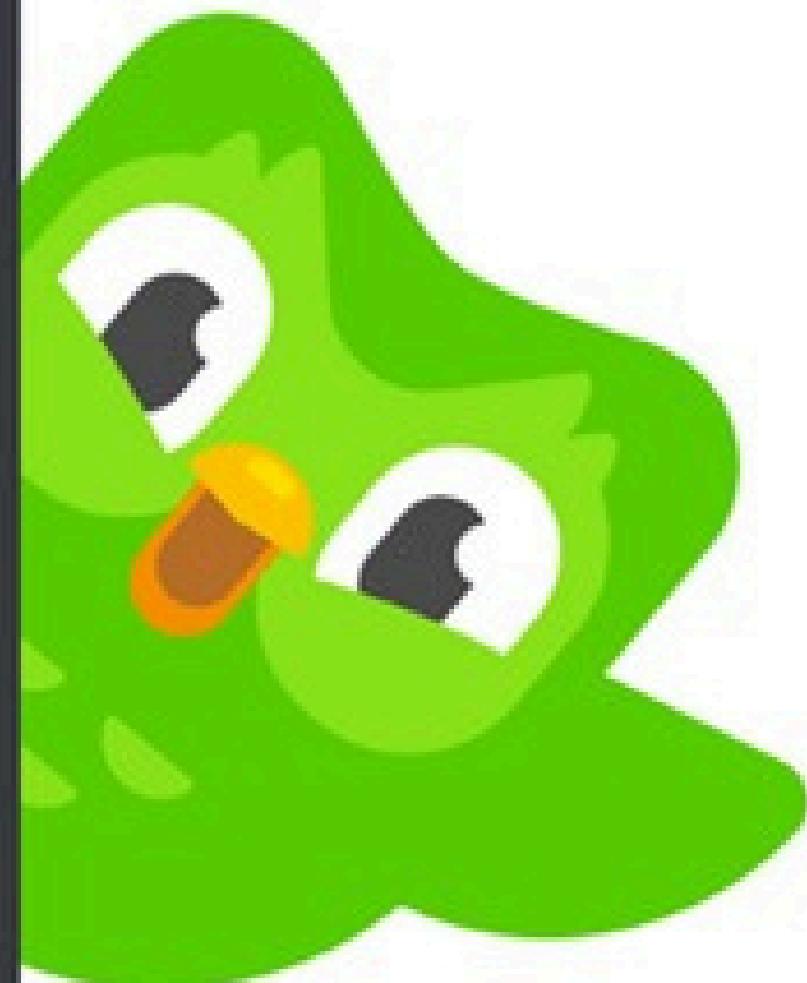
- **Not all correct answers are equal** - frequency matters
- **Language patterns emerge** - Vietnamese uses/omits "con", Hungarian has "férfi" vs "fiú" variations
- **Real learner data beats theory** - most common translations aren't always the most "textbook" correct
- **Grading can be nuanced** - reward natural language use, not just grammatical correctness

## ◆ Why This Matters for Duolingo

- **Better grading** - matches how real humans speak
- **Improved learning** - students learn which translations are most natural
- **Data-driven insights** - understand language patterns across 5 languages
- **Scalable** - could work for any language pair



Thank you



# Any Questions?

