

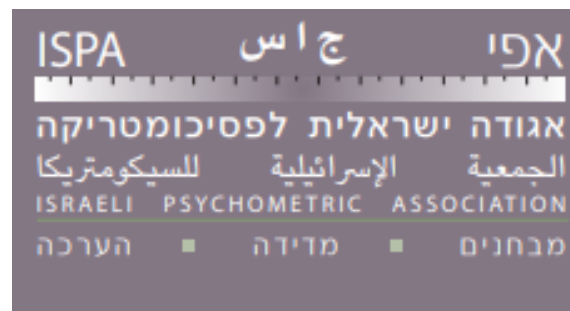
# First Steps towards Formative Assessment of Open Questions in Biology using HNLP\*

Moriah Ariely, Tanya Nazaretsky, Giora Alexandron

\* Hebrew Natural Language Processing

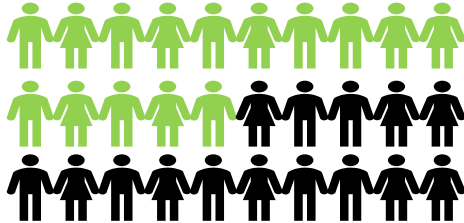


מכון ויצמן למדע  
WEIZMANN INSTITUTE OF SCIENCE

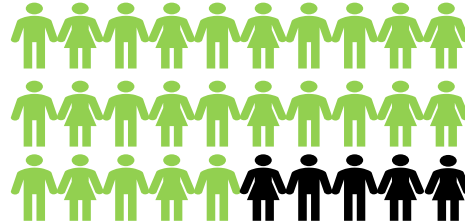


# The 2 Sigma Problem: Personalized feedback provides better results\*

Conventional  
(1-30)



Feedback-corrective  
(1-30)



Tutorial instruction  
(1-1)



\*Benjamin Bloom (1984)



## **Reality in Schools**

# Reality in Computerized Systems

Q. 8 9.09/	Q. 7 18.18/	Q. 6 9.09/	Q. 5 9.09/	Q. 4 9.09/	Q. 3 9.09/	Q. 2 9.09/	Q. 1 9.09/	ציון/100.00	משך הזמן שלקח לענות
נדרש מתן ציון	18.18 ✓	נדרש מתן ציון	6.82 ✗	9.09 ✓	9.09 ✓	נדרש מתן ציון	9.09 ✓	0	1 יום
נדרש מתן ציון	18.18 ✓	נדרש מתן ציון	9.09 ✓	9.09 ✓	9.09 ✓	נדרש מתן ציון	9.09 ✓	0	9 שעות 6 דקות
נדרש מתן ציון	18.18 ✓	נדרש מתן ציון	9.09 ✓	9.09 ✓	6.82 ✗	נדרש מתן ציון	9.09 ✓	0	7 שעות 15 דקות
נדרש מתן ציון	9.09 ✗	נדרש מתן ציון	9.09 ✓	9.09 ✓	9.09 ✓	נדרש מתן ציון	9.09 ✓	0	4 שעות 53 דקות
נדרש מתן ציון	18.18 ✓	נדרש מתן ציון	6.82 ✗	9.09 ✓	9.09 ✓	נדרש מתן ציון	9.09 ✓	0	1 שעה 22 דקות
נדרש מתן ציון	18.18 ✓	נדרש מתן ציון	9.09 ✓	9.09 ✓	9.09 ✓	נדרש מתן ציון	9.09 ✓	0	3 ימים 18 שעות
נדרש מתן ציון	18.18 ✓	נדרש מתן ציון	6.06 ✗	9.09 ✓	9.09 ✓	נדרש מתן ציון	9.09 ✓	0	1 שעה 14 דקות
נדרש מתן ציון	12.12 ✗	נדרש מתן ציון	4.55 ✗	9.09 ✓	9.09 ✓	נדרש מתן ציון	6.06 ✗	0	1 שעה 8 דקות

## שאלון לתלמידי ביולוגיה

\* Required

\* התאמה בין מבנה לתפקוד



(התמונה מאתר Shutterstock)

מבנה תאי-הדם האדומים מותאם לתפקודם בגוף האדם. דוגמא אחת להתאמה בין מבנה לתפקוד של תאי-הדם האדומים, היא היותם קטנים ופחוסים (ראה איור).  
**הסבר באמצעות הדוגמא הנ"ל כיצד מבנה תאי הדם האדומים מותאם לתפקודם.**

Your answer

More than 90% of students' written responses to open-ended questions are **not graded** and the students get **no feedback!**



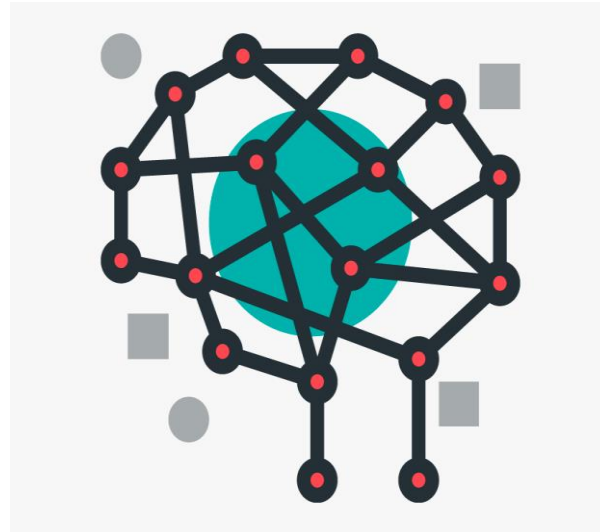
Can  
Artificial Intelligence (AI)  
and Natural Language  
Processing (NLP)  
assist in providing  
**Personalized Feedback?**



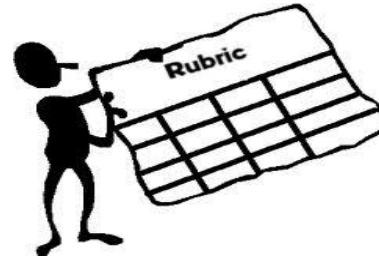
# **Leveraging NLP Technologies for Scientific Writing Instruction - Our contribution**

- Focus on formative assessment
  - Automatically grade written responses according to a grading rubric
  - Done in Hebrew
  - Our research can provide the scientific basis for developing computerized systems that will support science educators in teaching scientific writing
-

Students provide  
scientific explanations



Automated  
formative feedback



## Rationale for NLP-based feedback

# Research Design

## Research population

669 students from 10th to 12th grade

23 geographic locations of varied demographics

## The questions

- Students were presented with different phenomena resulting from smoking, anemia and high altitude
- All phenomena are connected to blood oxygen levels, and its relation to people's ability to exercise

אנשים מעשנים סובלים מקושי בביצוע פעילות גופנית. בסיגריה מספר חומרים מזיקים ביניהם הגז-CO (פחמן חד-חמצני). הגז-CO, המשתחרר בעת עישון הסיגריה, הוא בעל נטייה חזקה (גבוהה פי 250 מזו של החמצן) להיקשר להמוגלובין. **הסבר כיצד** נוכחות גבוהה של הגז-CO מסבירה את הקושי של אנשים מעשנים לבצע פעילות גופנית.

Your answer

\* אנמיה ופעילות גופנית

אצל אדם נמצאה בבדיקת דם כמות נמוכה של תאי דם אדומים (אנמיה). אדם זה הגיע לרופא עם תלונות של עייפות וקושי בביצוע פעילות גופנית. **הסבר כיצד** כמות נמוכה של תאי דם אדומים מסבירה את העייפות והקושי בביצוע פעילות גופנית.

Your answer

\* שהייה בגובה

כאשר עולים לגובה רב (כמו בעת טיפוס הרים), האוויר הוא דליל (לחץ אטמוספרי נמוך) וכמות החמצן בו מעטה.



# Research Design – the Grading Rubric

- The expected causal chain for the explanation
- Relevant concepts

Categories	
A	השפעת נוכחות CO על רמות חמצן בדם
B	תיאור השינוי ברמות החמצן המגיעות לתאים
C	תיאור השינוי על קצב/ כמות נשימה תאית
D	שימוש במושג נשימה תאית
E	תיאור השינוי בכמות/ בקצב הפקת אנרגיה
F	שימוש במושג אנרגיה / ATP
G	אנרגיה דרושה לפעילות גופנית / לפעילות תאי שריר

## \* עישון ופעילות גופנית

אנשים מעשנים סובלים מקושי בביצוע פעילות גופנית. בסיגריה מספר חומרים מזיקים ביניהם הגז-CO (פחמן חד-חמצני). הגז-CO, המשתחרר בעת עישון הסיגריה, הוא בעל נטייה חזקה (גבוהה פי 250 מזו של החמצן) להיקשר להמוגלובין. הסבר כיצד נוכחות גבוהה של הגז-CO מסבירה את הקושי של אנשים מעשנים לבצע פעילות גופנית.

Your answer

---

# Research Design – the Tagging Process

## An example of a students' explanation

הגז CO מתחבר אל ההמוגלובין שתפקידו לקלוט חמצן ולהעבירו לתאים ומכיוון שהגז מתחבר אליו החמצן לא יכול להתרכב איתו וההמוגלובין לא מעביר חמצן לתאים כאשר יש פחות חמצן בתאים יש פחות אנרגיה מכיוון שהתאים מייצרים את האנרגיה בעזרת חמצן ועם פחות אנרגיה קשה יותר לעשות פעילות גופנית

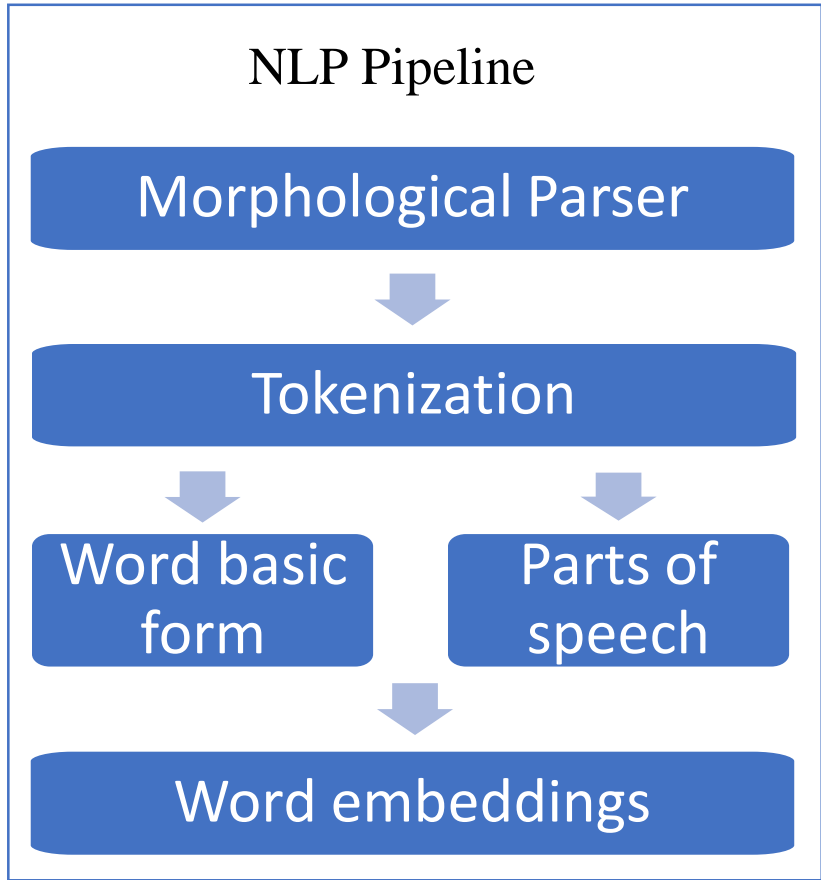
A	B	C	D	E	F	G
1	1	0	0	1	1	1

The answer lacks a reference to changes in cellular respiration

# Machine Learning Process

- NLP: Convert raw text of student responses into a numeric representation that can be processed by machine learning algorithms.
- Use expert grading to ‘learn’ a function that models the relation between student responses and their gradings
- Apply the function to unseen responses to predict the grades
- Compare machine gradings to expert grading

Raw answer  
"הגז פחמן חד חמצני  
מתחבר אל  
ההמוגלובין שתפקידו"  
לקלוט חמצן ולהעבירו  
לתאים.

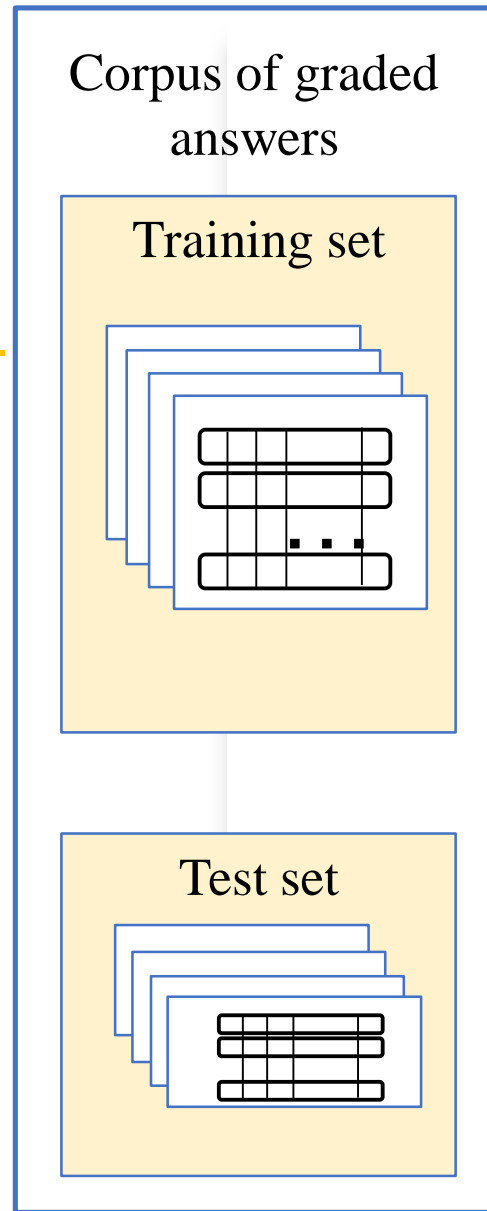
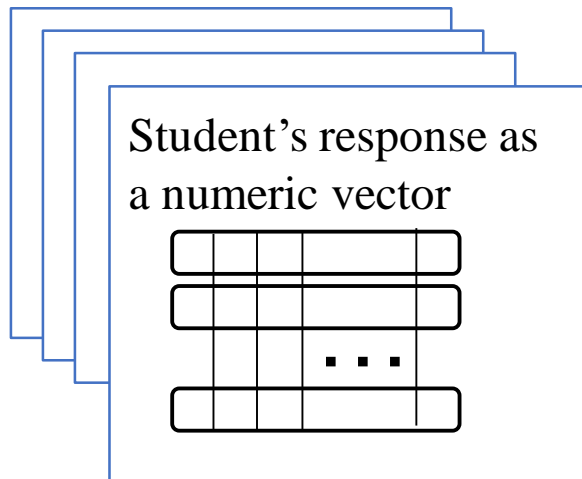


Numeric representation for AI:

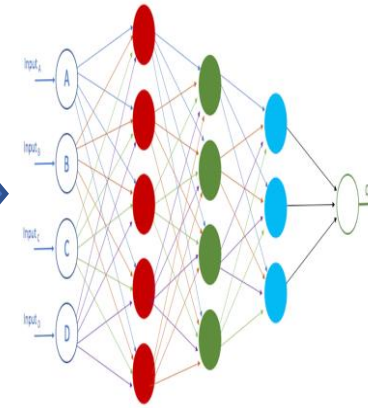
גז				
פחמן				
			...	
תא				

# NLP analysis

# Machine Learning Process



(per grading category)



Compare predicted grades to human grades:

- Accuracy
- Cohen's Kappa

# Results\* - Automatic Grading

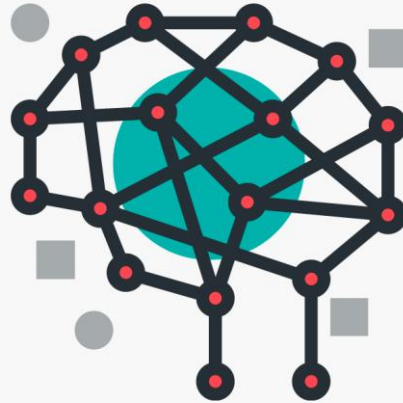
Category		Training set – Smoking Test set - Smoking	
		Accuracy	Cohen's Kappa
A	השפעת נוכחות CO על רמות חמצן בדם	0.88	0.53
B	תיאור השינוי ברמות החמצן המגיעות לתאים	0.85	0.61
C	תיאור השינוי על קצב/ כמות נשימה תאית	0.88	0.71
D	שימוש במושג נשימה תאית	0.98	0.95
E	תיאור השינוי בכמות/ בקצב הפקת אנרגיה	0.91	0.83
F	שימוש במושג אנרגיה / ATP	0.96	0.91
G	אנרגיה דרושה לפעילות גופנית / לפעילות תאי שריר	0.87	0.68

\*Ariely, Nazaretsky & Alexandron, submitted

# Next steps: NLP-based feedback



AI system



Auto' grading

Grade	Weight	Category
1	2	השפעת נוכחות CO על רמות חמצן בדם
1	3	תיאור השינוי ברמות החמצן המגיעות לתאים
1	2	תיאור השינוי על קצב/ כמות נשימה תאית
1	1	שימוש במושג נשימה תאית
0	1	תיאור השינוי בכמות/ בקצב הפקת אנרגיה
0	1	שימוש במושג אנרגיה / ATP
0	2	אנרגיה דרושה לפעילות גופנית / לפעילות תאי שריר



Computerized Feedback

Final grade:

Weighed sum: 8  
Grade:  $8/12 = 67$

Formative Feedback:

How are low cellular oxygen levels related to low levels of ATP/ energy?



Moriah



Tanya



Giora

**Thank you for listening!**

We wish to thank Dr. Cipy Hofman  
for her contribution