Evaluating Tooth Brushing Performance With Smartphone Sound Data

JOSEPH KORPELA[1] · RYOSUKE MIYAJI[1] · TAKUYA MAEKAWA[1]

KAZUNORI NOZAKI[2] · HIROO TAMAGAWA[2]

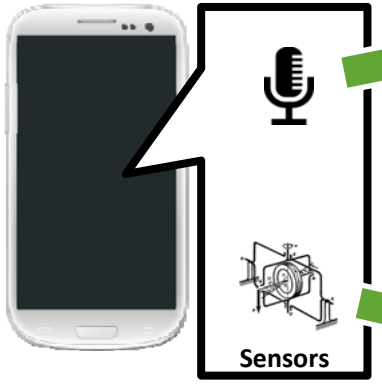
[1]OSAKA UNIVERSITY, GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY

[2]OSAKA UNIVERSITY DENTAL HOSPITAL

大阪大学
OSAKA UNIVERSITY

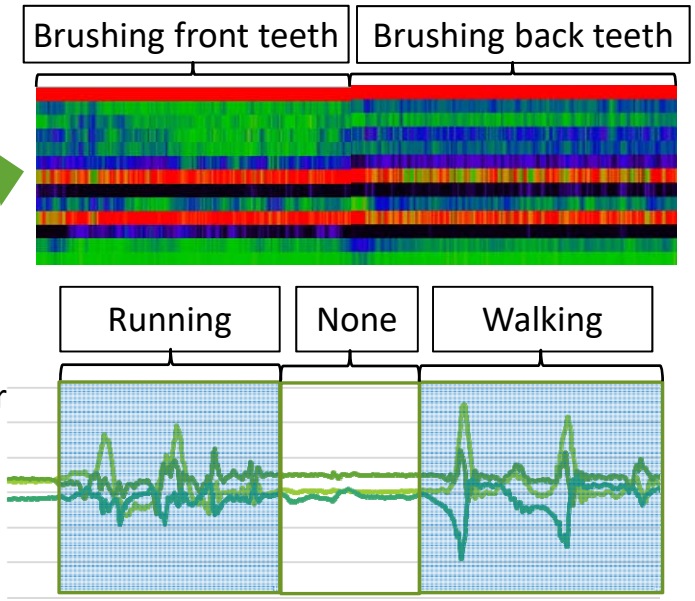
Activity Recognition

Daily Activities



Audio Data

Accelerometer Data



Sensor Data Collected Using a Smartphone

Machine Learning Used to Recognize Activities Based on Sensor Data

Activity Recognition in Health Care

Tracking sleep quality/quantity

Tracking exercise

Tracking medication intake

Tracking food intake

Dental Health

Teeth are important to our health
- Need to last a lifetime
- Tooth loss leads to loss of appetite and decreased nutrition

Brushing is important for our teeth
- Proper brushing improves dental health
- Improper brushing can damage teeth and gums

Yet, most people don't brush well enough

Activity Recognition for Dental Health

Significant improvement in brushing habits when provided feedback via activity recognition techniques[1]

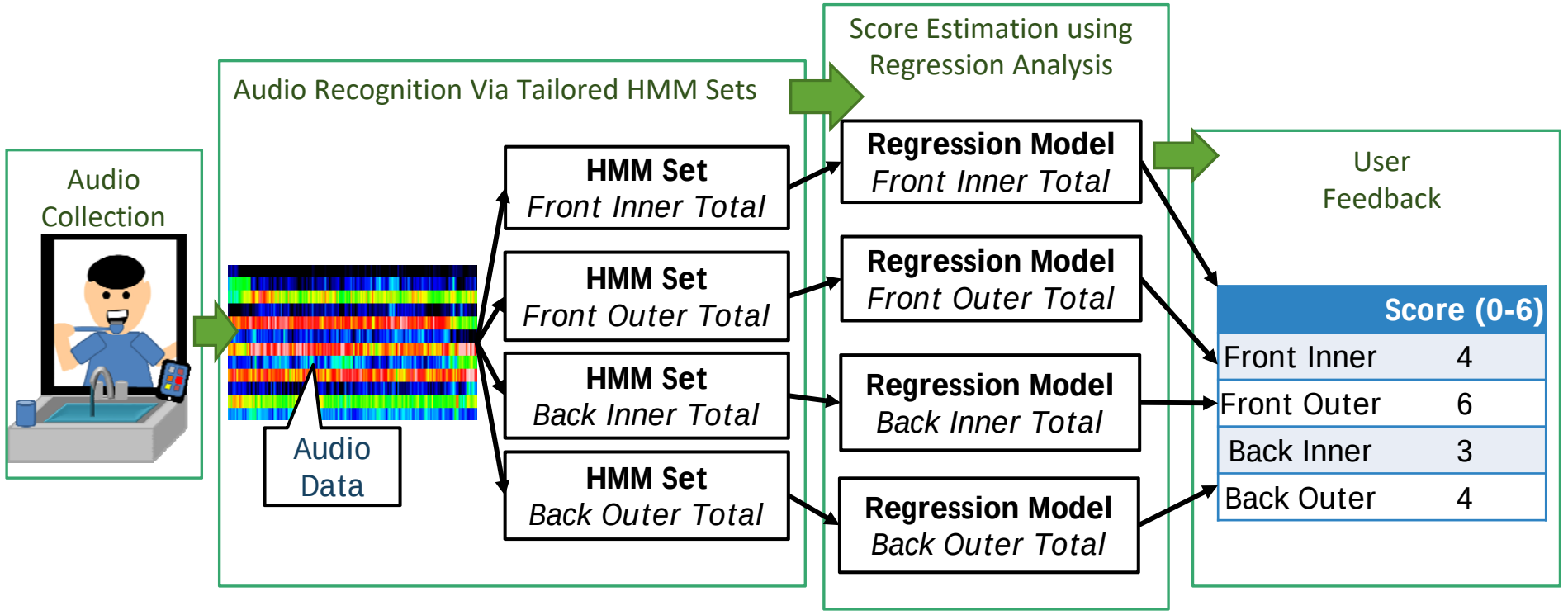
Previous methods have required specialized equipment

- LED extension for toothbrush[1]
- Accelerometer extension toothbrush[2]

Our method uses only *audio* data:
Allows users to evaluate brushing using an off-the-shelf smartphone

1. Chang, Y.-C., Lo, J.-L., Huang, C.-J., Hsu, N.-Y., Chu, H.-H., Wang, H.-Y., Chi, P.-Y., and Hsieh, Y.-L. Playful toothbrush: ubicomp technology for teaching tooth brushing to kindergarten children. In *CHI 2008* (2008), 363–372.
2. Graetz, C., Bielfeldt, J., Wolff, L., Springer, C., Fawzy El-Sayed, K. M., Salzer, S., Badri-Hoher, S., and Dorfer, C. E. Toothbrushing education via a smart software visualization system. *Journal of Periodontology 84*, 2 (2013), 186–195.

Proposed Method



Evaluation Scores: Plaque Tests

Evaluation Scores
- Regression models need scores to use as training data

Plaque Test

Typical method of evaluating tooth brushing effectiveness

1. Apply plaque indicator liquid to teeth
2. Liquid makes plaque easily visible
3. Dentist evaluates based on plaque left remaining

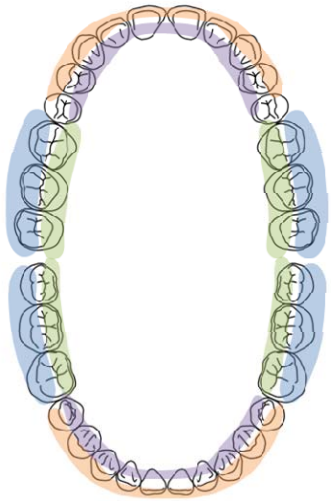
Issues with using plaque test
- Influenced by all tooth brushing performed over last few days
- Influenced by foods/drinks recently consumed
- Costly to gather a large number of scores

Evaluation Scores: Video-based

Video Capture Using Smartphone

Dentist Evaluates Video

Dentist Assigns 12 Scores



Four areas of the mouth:

Front teeth, outer surface

Back teeth, outer surface

Front teeth, inner surface

Back teeth, inner surface

Three scores per area:

- **Coverage** *(2 pts)*
- **Stroke** *(2 pts)*
- **Duration** *(2 pts)*

Example: *Front teeth, inner surface coverage* = [0,2]

Video-based Scores vs. Plaque Test Scores

14 Subjects

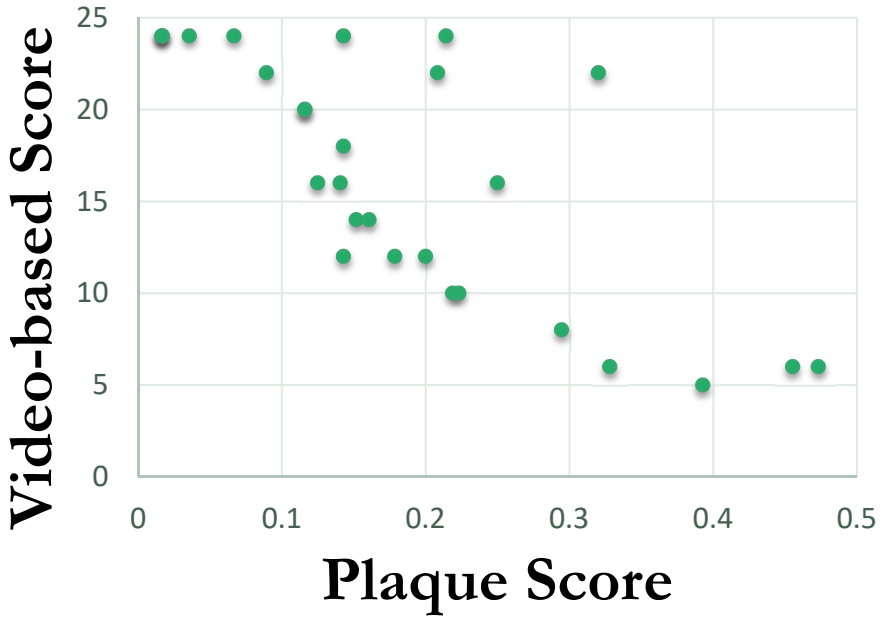
Day 1:

- Brushed teeth with video recorded
- Received plaque test

Day 2:

- Received instruction on proper brushing technique
- Brushed teeth with video recorded
- Received plaque test

Video data was then used to generate scores for each session



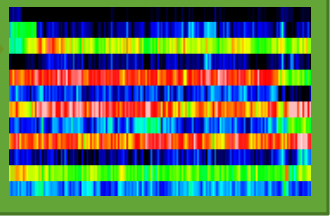
Audio Recognition

Raw Audio Collected by Smartphone Microphone

Feature Extraction

12-order MFCC* + Delta + Acceleration

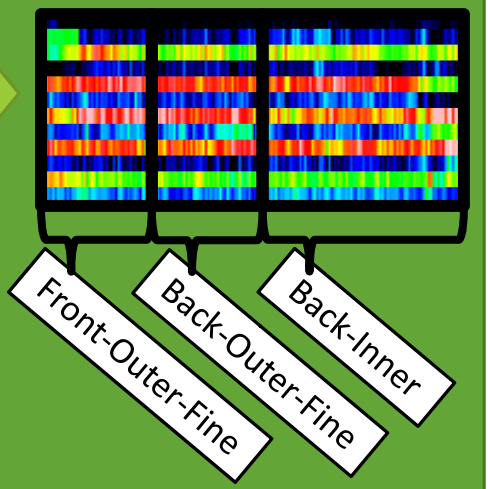
50 ms windows



Audio Recognition
GMM-based HMM

HMM Classes (7 total)	
None: No tooth brushing activity	
Outer front teeth, fine	Outer front teeth, rough
Outer back teeth, fine	Outer back teeth, rough
Inner front teeth**	Inner back teeth**

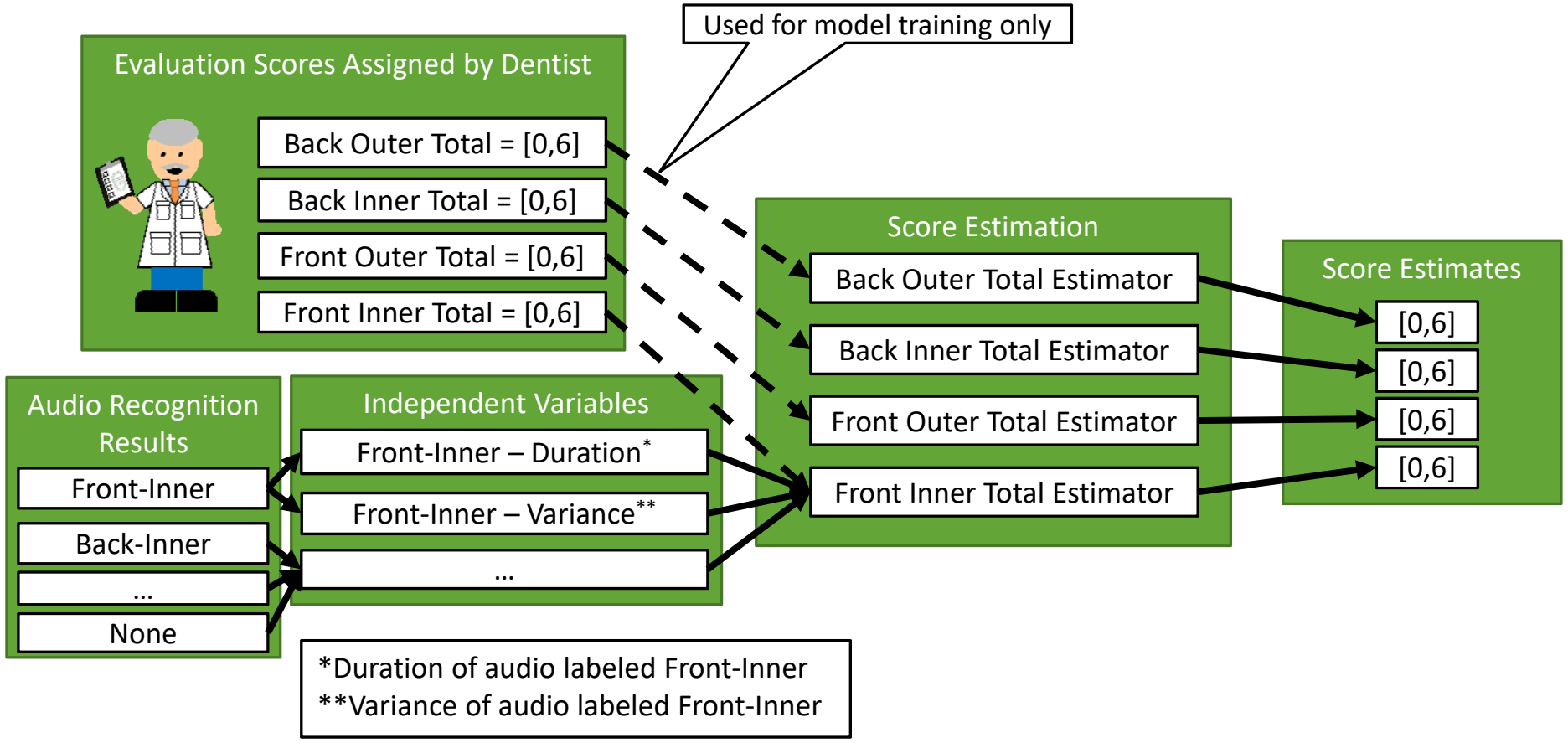
Results used as input for Regression Analysis



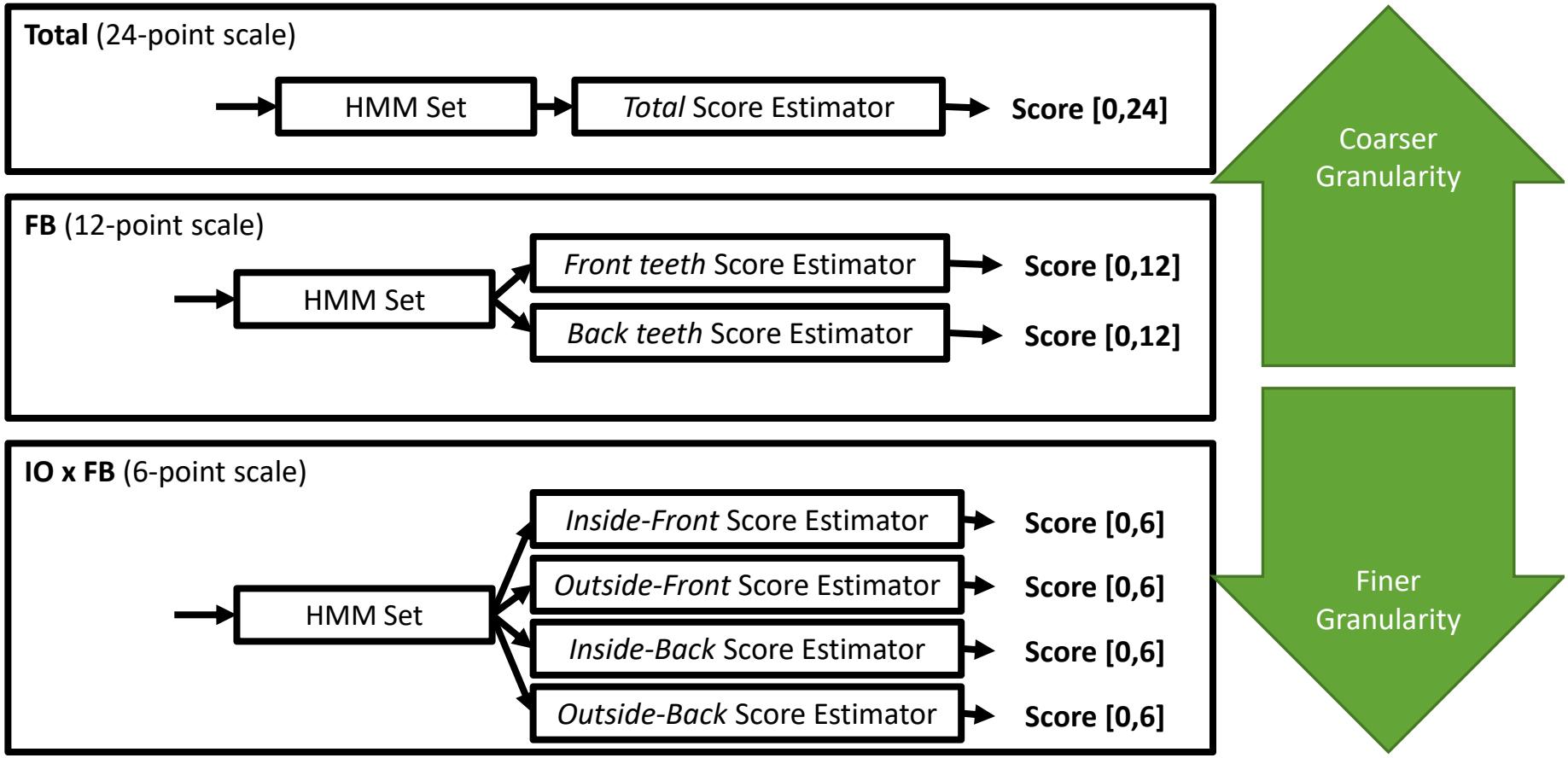
*MFCC: Represent audio as a series of logarithmically-spaced coefficients (Commonly used in speech recognition and environmental sound recognition studies.)

**No fine/rough stroke distinction (due to an insufficient amount of data)

Score Estimation



Score Architectures

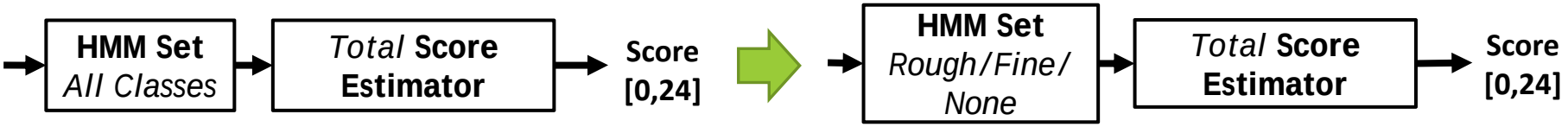


Improving HMM Performance

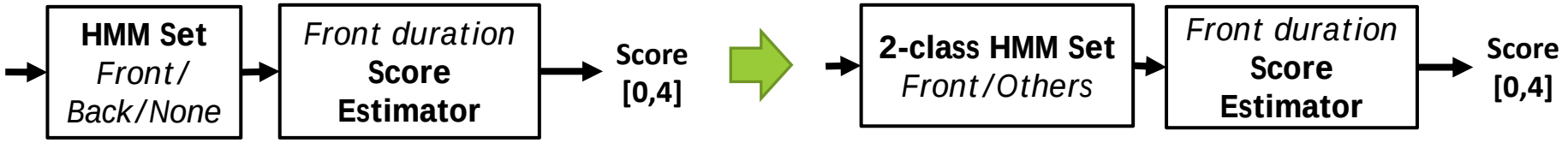
Audio recognition performance is better at coarser granularities

(Accuracy when using all classes: 45.1% → when using only 3-classes: 68.4%)

1. HMM granularity required depends on the score granularity



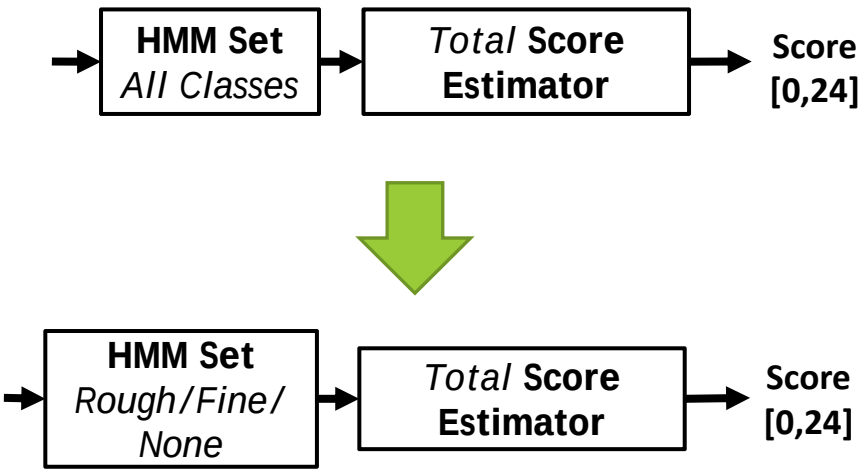
2. Individual scores require different sets of HMMs



Improving Performance:

1. Create HMM sets with varying granularity
2. Create HMM sets that are tailored to each score

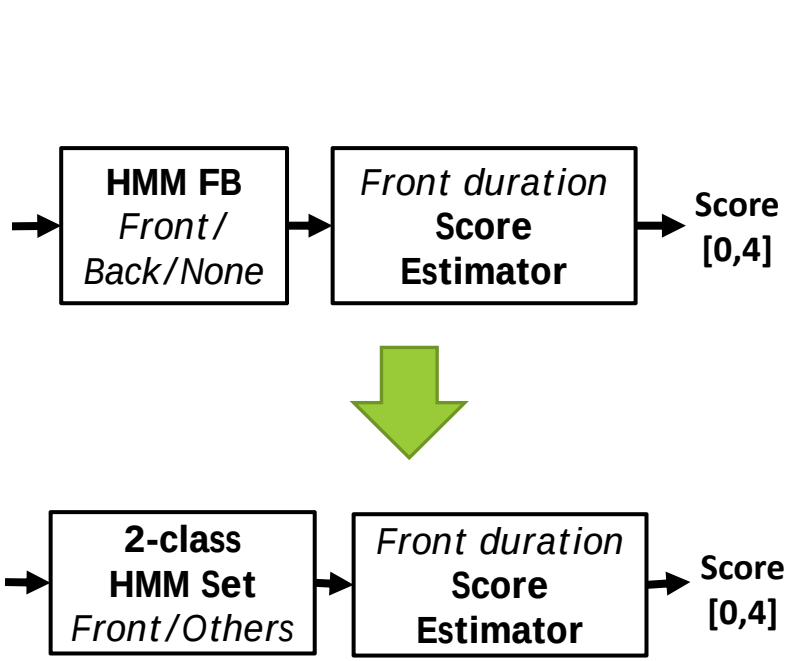
Varying HMM Granularity

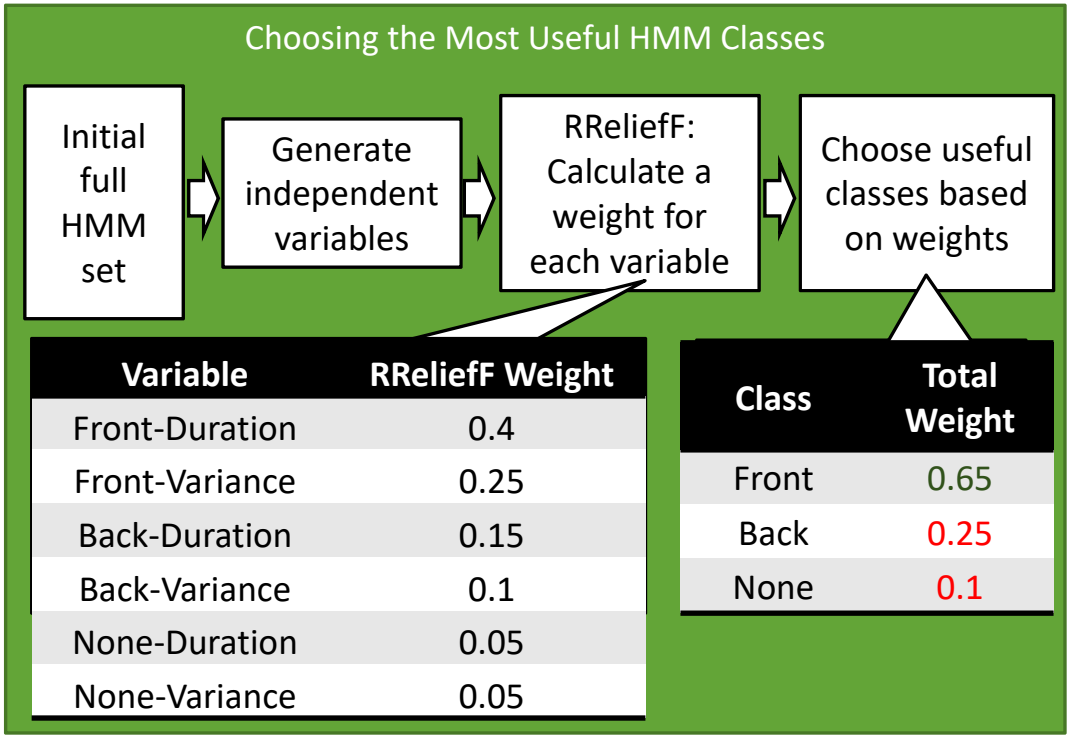


Four sets with varying granularity

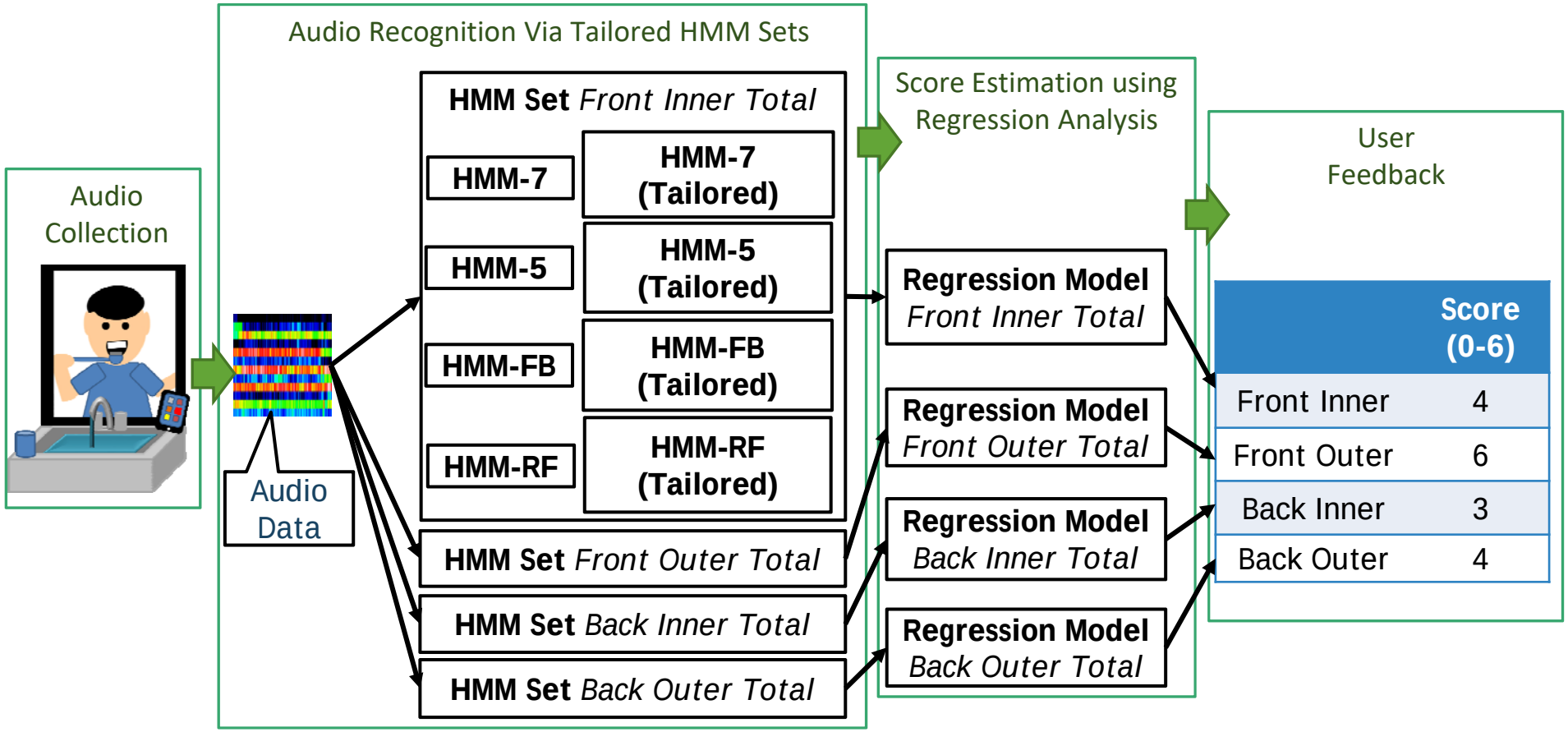
- HMM-7 (All classes)
- HMM-5 (No Rough/Fine Distinction)
- HMM-FB (Front/Back/None)
- HMM-RF (Rough/Fine/None)

Tailoring HMM Sets to Regression Scores





Proposed Method

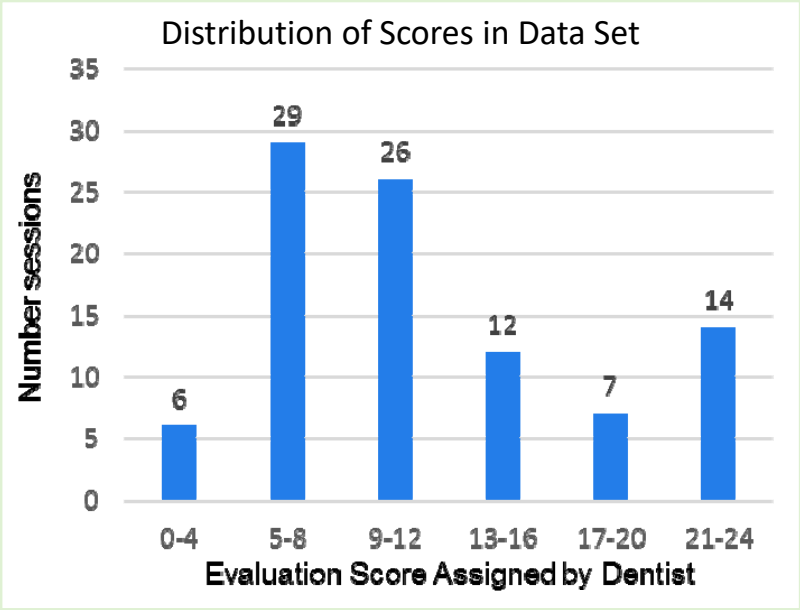


Evaluation Methodology

Data Set
- 94 sessions total
- 14 participants
- Average length of each session: 94 seconds

Environment
- Collected either in our lab or in the participant's own home
- Users allowed to use own toothbrush or one provided by us

Evaluated using *leave-one-user-out* cross validation

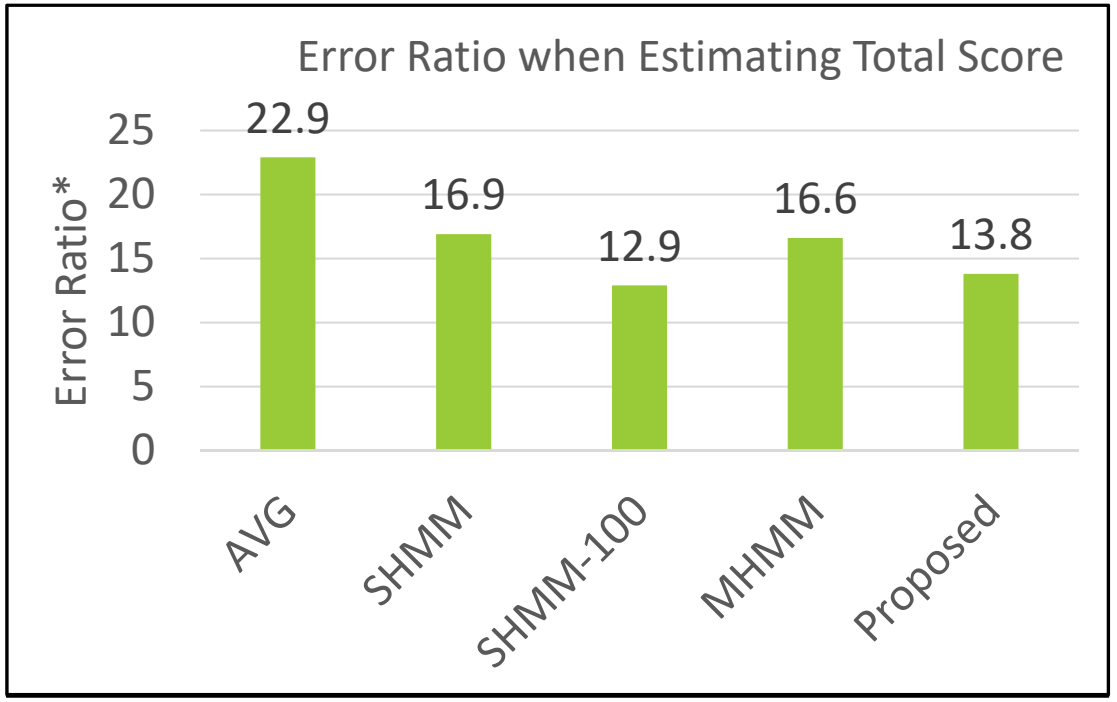


Score Estimation Evaluation: Methods

1. **Avg**: Each user's scores are estimated using the average scores for other users.
2. **SHMM**: The same HMM set (*HMM set 7*) is used to generate independent variables for estimating all scores.
3. **SHMM-100**: A variation of the *SHMM* method in which we built the regression models using corrected labels, i.e., this method assumed 100% recognition accuracy for *HMM set 7*.
4. **MHMM**: Four basic HMM sets: *HMM set 7*, *HMM set 5*, *HMM set FB*, and *HMM set RF*, are used to generate independent variables for estimating the scores.
5. **Proposed**: The proposed method, in which we prepared a tailored group of HMM sets for each of the scores.

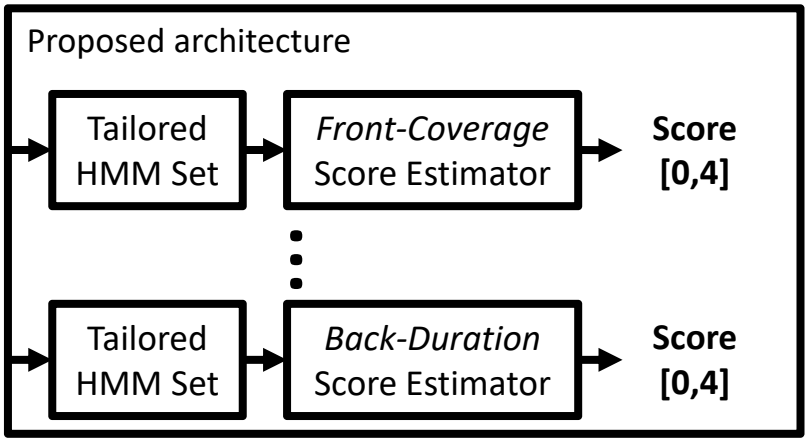
Score Estimation Evaluation: Total Architecture

Estimated a single score (24-point scale) that represents the total score for all tooth brushing activity in the session.

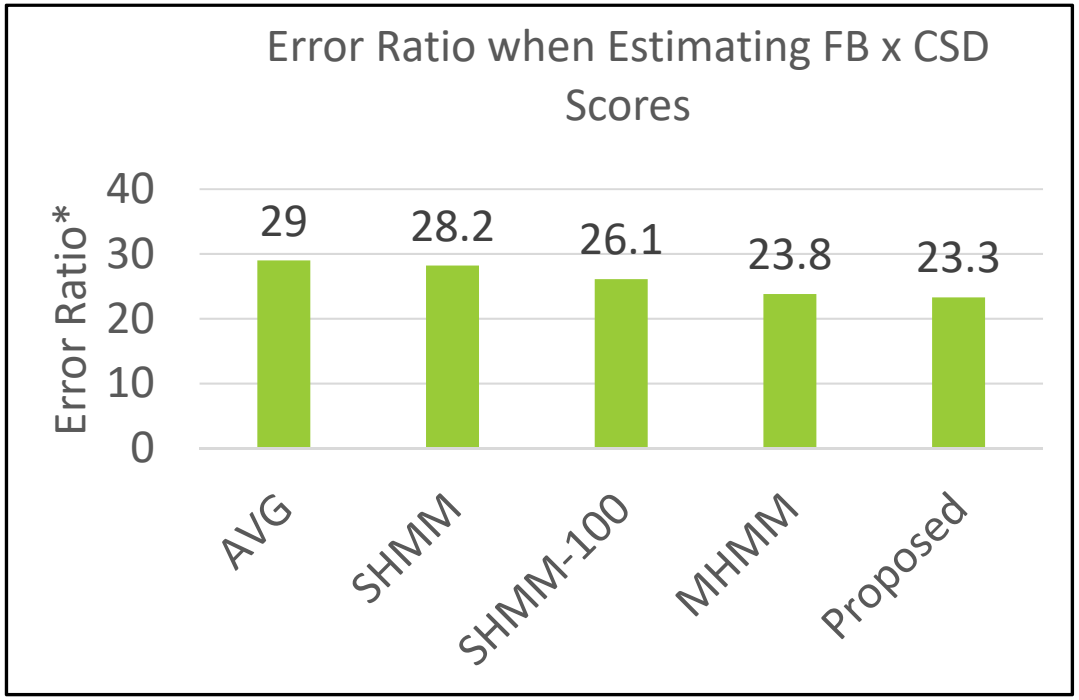


***Error Ratio = MAE / Total score**

Score Estimation Evaluation: FB x CSD Architecture



Estimated six scores (4-point scale), corresponding to each of the three evaluation criteria for both the front teeth and back teeth.

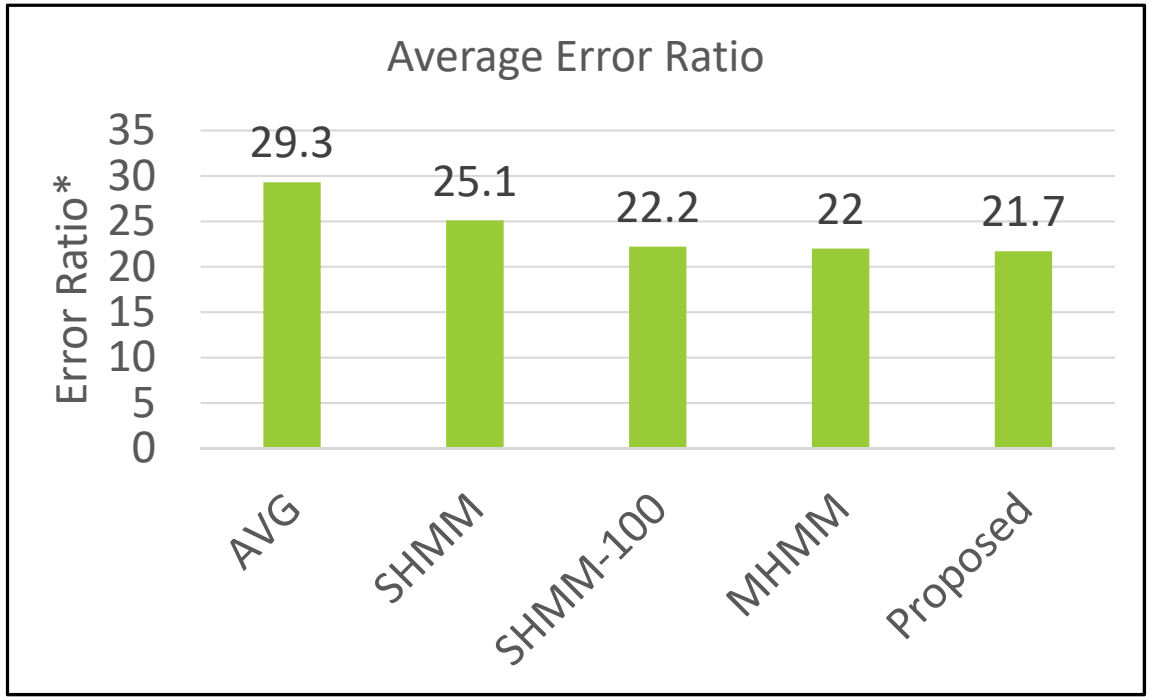


***Error Ratio = MAE / Total score**

Score Estimation Evaluation: Average Results

Average results for all architectures

1. **Total** (24-pt scale): Estimates one score
2. **FB** (12-pt scale): Two scores: **front** and **back** teeth
3. **CSD** (8-pt scale): Three scores: **coverage**, **stroke**, and **duration**
4. **IO x FB** (6-pt scale): Four scores: One for each area of the mouth **outer front**, **inner front**, **outer back**, and **inner back**
5. **FB x CSD** (4-pt scale): Six scores: **CSD** scores for front teeth and back teeth
6. **IO x FB x CSD** (2-pt scale): Twelve scores: **CSD** for four areas of mouth



***Error Ratio = MAE / Total score**

Conclusion

Proposed a method for evaluating tooth brushing based on audio data

- Create training data using video-based evaluation
  - Enables creation of large amounts of training data
- Perform evaluation on test data using audio-based evaluation
  - Makes method easily accessible to average user
- Tailor HMM sets to score being evaluated
  - Improves performance by avoiding unnecessary distinctions