A LOW-COST EEG SYSTEM FOR EARLY DETECTION OF LEARNING DIFFICULTIES IN STUDENTS BASED ON COGNITION STATES

Project Reference No.: *49S_BE_2724*

College : Bapuji Institute of Engineering and Technology, Davanagere
Branch : Department of Computer Science and Engineering (Data Science)
Guide : Prof. Ranjana B Jadekar
Student(s) : Ms. Sejal M
Ms. M Sindhu
Ms. Srushti R Gaonkar
Ms. Vijayalaxmi B

Keywords:
Electroencephalography (EEG), Epistemic Cognition, Cognitive State Classification, Transformer Model

Introduction:

1. Understanding how students learn and process knowledge is essential for improving educational outcomes. One important concept in this area is epistemic cognition, which refers to how learners understand, evaluate, and apply knowledge during learning activities. Traditional methods such as surveys and tests are often subjective and fail to capture real-time cognitive states.
2. Electroencephalography (EEG) provides a non-invasive way to measure brain activity and offers continuous insights into a student's cognitive engagement, attention, and mental workload. EEG signals consist of different frequency bands such as alpha, beta, and gamma, which are associated with various cognitive processes like relaxation, focus, and higher-order thinking.

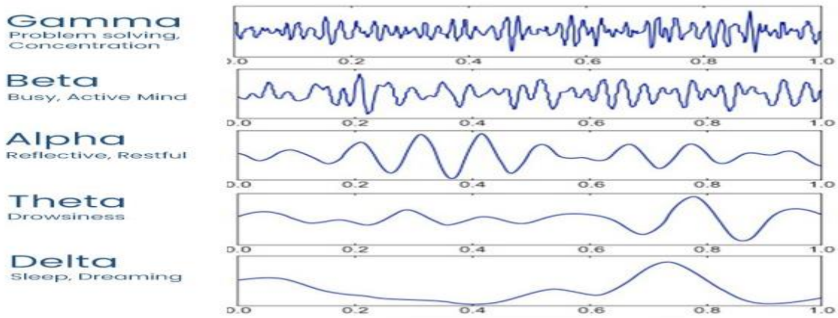


Figure 1: EEG Brainwave Frequency Bands

Frequency Band	Range(Hz)	Cognitive State
Delta (δ)	0.5–4 Hz	Deep sleep, unconsciousness.
Theta (θ)	4–8 Hz	Drowsiness,meditation,early learning,creativity.
Alpha (α)	8–13 Hz	Relaxed wakefulness,calm but alert,eyes closed.
Beta (β)	13–30 Hz	Active thinking, focus, problem-solving.
Gamma (γ)	30–100 Hz	High-level cognition,memory binding,consciousness

Table 1: EEG Frequency Ranges and Their Cognitive States

Cognitive State	EEG Characteristics	Description
Naive	Higher theta (4-8Hz), lower alpha(8-10Hz)	Passive learning or early-stage cognitive tasks.
Developing	Mid alpha and beta(10-15Hz)	Improving attention and more structured thinking.
Sophisticated	Higher beta and gamma activity(18-40Hz)	Shows expert cognition and strong working memory.

Table 2: Distinguishing Cognitive States

1. With advancements in artificial intelligence, deep learning models can effectively analyse complex EEG signals. In this project, a Transformer-based deep learning model is used to classify students' cognitive states into categories such as naïve, developing, and sophisticated. These categories represent different levels of understanding, ranging from basic memorization to deeper analytical thinking.
2. Before classification, EEG signals undergo preprocessing steps such as filtering and segmentation to improve data quality. The processed data is then used to train the model and evaluate performance using metrics like accuracy and precision.
3. The proposed system aims to enable monitoring of students' cognitive states and support personalized learning. By identifying students who need assistance, this approach can help educators provide timely interventions and improve overall learning effectiveness.

Objectives:

- To preprocess EEG signals from students during various learning activities in order to extract meaningful cognitive features linked to attention, engagement, confusion.
- To apply deep learning model for accurate classification of different cognitive states based on EEG signal patterns.
- To evaluate the impact of different educational tasks and environments on student cognition using EEG-based analytics.
- To validate the effectiveness of EEG-based cognitive classification in improving personalized learning experiences and overall academic outcomes.

Methodology:

Step 1: Data Acquisition

Multichannel EEG data is collected from students while they perform educational tasks. Initial verification is carried out to remove missing values, detect abnormal signals, and ensure proper distribution of cognitive states such as naïve, developing, and sophisticated.

Step 2: Feature Extraction

EEG signals are divided into frequency bands including theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–45 Hz). From these bands, cognitive features such as attention, engagement, confusion, cognitive load, and arousal are computed to represent brain activity patterns.

Step 3: Data Preprocessing and Normalization

Preprocessing techniques are applied to improve signal quality. Z-score normalization is used to standardize the data, ensuring consistent scaling and better model performance. To ensure consistent scaling across features, z-score normalization is applied:

$$z = \frac{x - \mu}{\sigma}$$

This standardization improves model convergence and prevents feature dominance.

Step 4: Dataset Balancing and Segmentation

To handle class imbalance among cognitive states, Mini-Batch K-Means clustering is used. This ensures balanced representation of all classes during training.

EEG signals are inherently sequential; therefore, sliding windows are created:

- Sequence length: 64
- Stride: 32

Each window forms a temporal segment of 64 consecutive samples, converted into a format suitable for Transformer input.

Step 5: Transformer-Based Model Training

A Transformer-based deep learning model is used to analyse EEG sequences. It uses self-attention mechanisms to capture relationships across time and features, enabling effective classification of cognitive states into naïve, developing, and sophisticated categories

Classification Layer:

A Softmax output with three classes:

$$\{0 = \text{naïve},\ 1 = \text{developing},\ 2 = \text{sophisticated}\}$$

Training uses:

- Cross-entropy loss
- Adam optimizer
- Class-weights to address residual imbalance
- Dropout regularization to avoid overfitting

Validation accuracy is monitored across epochs to ensure optimal model generalization.

Step 6: Model Evaluation and Performance Analysis

The trained model is evaluated using metrics such as accuracy, precision, recall, and F1-score. A confusion matrix is also used to analyse classification performance across different cognitive states.

Step 7: Cognitive State Prediction and Interpretation

The final model predicts students' cognitive states based on EEG inputs. These predictions can be used to support adaptive learning systems and provide personalized educational feedback.

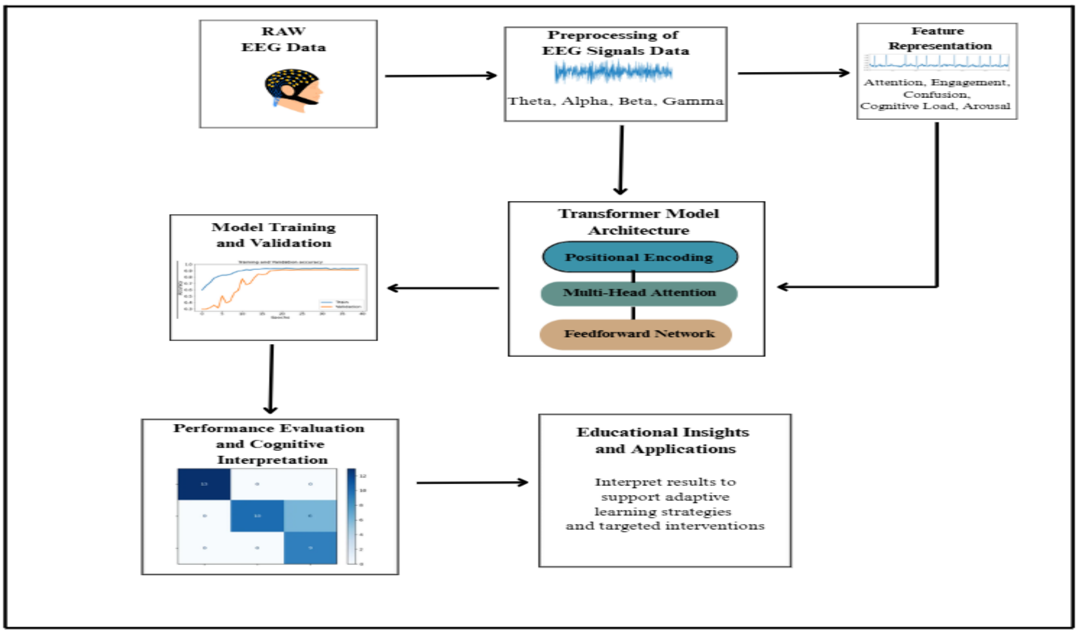


Figure 2: Methodology for Classifying Students Cognitive States in Educational Context Using EEG Signals

Results and Conclusion:

The proposed Transformer-based EEG cognition classification system was evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score. The model achieved an overall classification accuracy of approximately **90%**, indicating strong performance in identifying students' cognitive states as naïve, developing, and sophisticated.

The training and validation curves show a steady increase in accuracy and a decrease in loss across epochs, indicating stable learning and effective convergence of the model. The small gap between training and validation performance suggests minimal overfitting and good generalization capability.

```
Using device: cpu
⚠ Stratify disabled: some videos appear only once
Scaler saved to 'eeg_transformer_scaler.pkl'
Windows shapes -> train: (4340, 64, 5), val: (240, 64, 5), test: (256, 64, 5)
Epoch 01 | Train Acc: 0.5440, Val Acc: 0.5583 | Train Loss: 0.8975, Val Loss: 0.9350
Epoch 02 | Train Acc: 0.6654, Val Acc: 0.7125 | Train Loss: 0.7031, Val Loss: 0.6336
Epoch 03 | Train Acc: 0.7378, Val Acc: 0.7750 | Train Loss: 0.5905, Val Loss: 0.5374
Epoch 04 | Train Acc: 0.7783, Val Acc: 0.7167 | Train Loss: 0.5208, Val Loss: 0.5899
Epoch 05 | Train Acc: 0.7816, Val Acc: 0.7750 | Train Loss: 0.5213, Val Loss: 0.5668
Epoch 06 | Train Acc: 0.8035, Val Acc: 0.7042 | Train Loss: 0.4609, Val Loss: 0.7265
Epoch 07 | Train Acc: 0.8145, Val Acc: 0.8500 | Train Loss: 0.4303, Val Loss: 0.4003
Epoch 08 | Train Acc: 0.8320, Val Acc: 0.8417 | Train Loss: 0.4043, Val Loss: 0.3948
Epoch 09 | Train Acc: 0.8281, Val Acc: 0.8667 | Train Loss: 0.4104, Val Loss: 0.3472
Epoch 10 | Train Acc: 0.8495, Val Acc: 0.8625 | Train Loss: 0.3625, Val Loss: 0.3599
Epoch 11 | Train Acc: 0.8622, Val Acc: 0.8500 | Train Loss: 0.3350, Val Loss: 0.3511
Epoch 12 | Train Acc: 0.8719, Val Acc: 0.8167 | Train Loss: 0.3253, Val Loss: 0.4933
Epoch 13 | Train Acc: 0.8733, Val Acc: 0.8083 | Train Loss: 0.3137, Val Loss: 0.5027
Epoch 14 | Train Acc: 0.8696, Val Acc: 0.8833 | Train Loss: 0.3137, Val Loss: 0.3064
Epoch 15 | Train Acc: 0.8742, Val Acc: 0.7875 | Train Loss: 0.3007, Val Loss: 0.6624
Epoch 16 | Train Acc: 0.8836, Val Acc: 0.8917 | Train Loss: 0.2834, Val Loss: 0.2342
Epoch 17 | Train Acc: 0.8786, Val Acc: 0.8750 | Train Loss: 0.2895, Val Loss: 0.4030
Epoch 18 | Train Acc: 0.8887, Val Acc: 0.8917 | Train Loss: 0.2732, Val Loss: 0.2886
Epoch 19 | Train Acc: 0.8977, Val Acc: 0.8542 | Train Loss: 0.2528, Val Loss: 0.5200
Epoch 20 | Train Acc: 0.8896, Val Acc: 0.9000 | Train Loss: 0.2634, Val Loss: 0.2431
Saved model as 'best_eeg_transformer.pt'
```

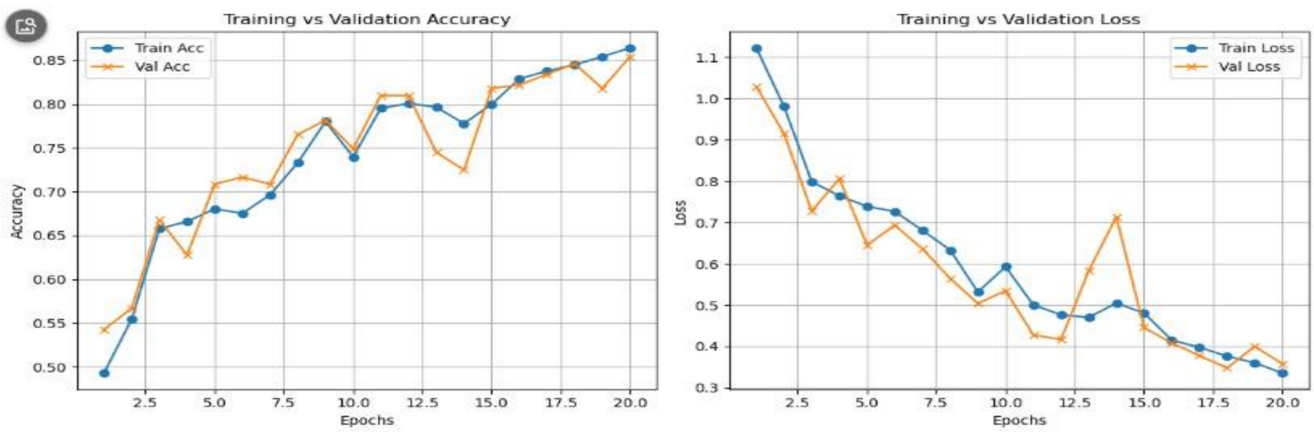
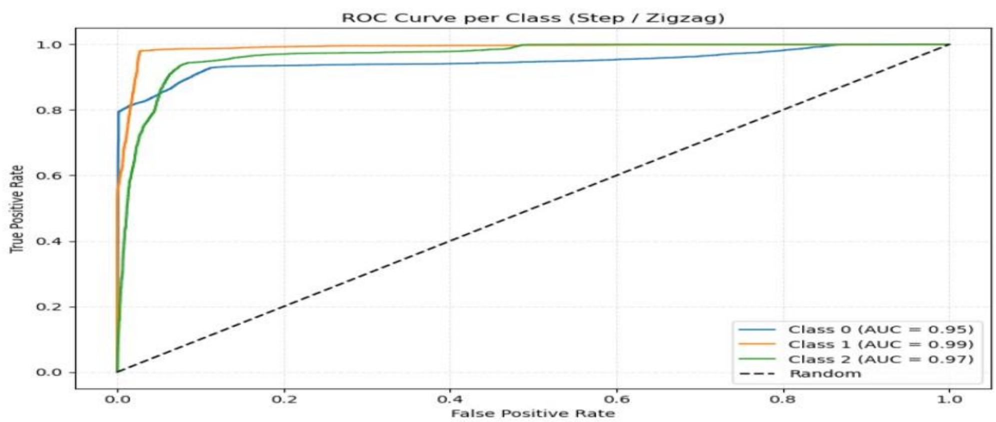


Figure 3: Training and Validation Accuracy/Loss Curves

The training and validation curves show a steady increase in accuracy and a decrease in loss across epochs, indicating stable learning and effective convergence of the model. The small gap between training and validation performance suggests minimal overfitting and good generalization capability.




```
Sensitivity Class 0: 0.78
Sensitivity Class 1: 0.98
Sensitivity Class 2: 0.95
```

Figure 4: ROC Curve for Multi-Class EEG Classification

The ROC curve analysis shows high AUC values (above 0.95) for all cognitive classes, demonstrating strong discriminative ability of the model. The confusion matrix further confirms that most predictions are correctly classified, with only minor confusion between closely related cognitive states.

```
Using device: cpu
Accuracy: 0.9019271383315733
Precision: 0.9141114719991333
Recall: 0.9019271383315733
F1 Score: 0.9013708160581929
Confusion Matrix:
 [[3979   63 1074]
 [  15 4921   82]
 [   2  250 4766]]
```

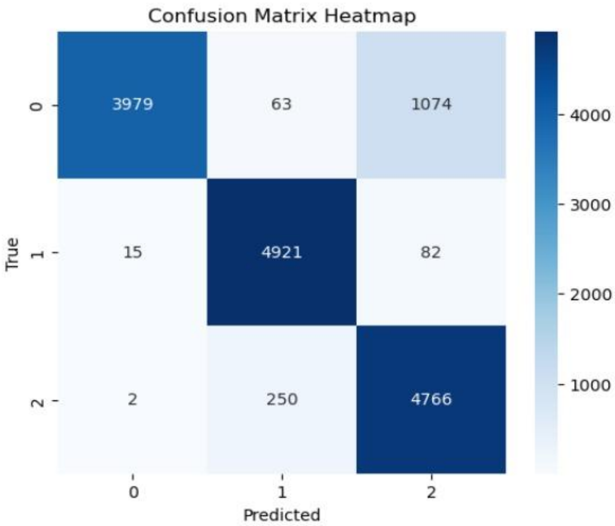


Figure 5: Confusion Matrix for Cognition State Classification

In addition to model performance, a user-friendly EEG cognition dashboard was developed to demonstrate practical implementation. The system allows users to input EEG data, analyse cognitive states, and visualize results through an interactive interface.

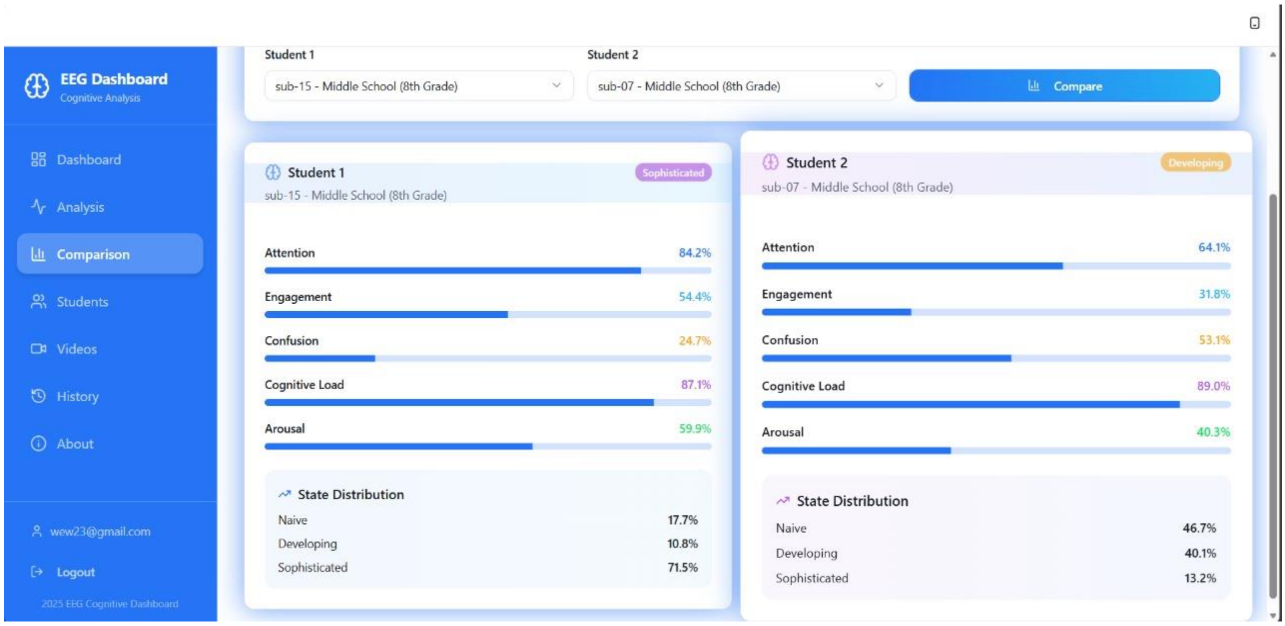


Figure 6: Cognitive State Comparison Interface

Overall, the results demonstrate that Transformer-based models are highly effective for analysing EEG signals and classifying cognitive states. The proposed system can be used in adaptive learning environments to monitor student engagement and provide personalized feedback. This approach has strong potential to improve learning outcomes by enabling understanding of students' cognitive processes.

References:

1. R. Y.-Y. Chan, "EEG Transformer for Classifying Students' Epistemic Cognition States in Educational Contexts," *IEEE Access*, vol. 13, pp. 23935– 23948, Feb. 2025.
2. E. Shi, S. Yu, Y. Kang, J. Wu, L. Zhao, D. Zhu, J. Lv, T. Liu, X. Hu, and S. Zhang, “MEET: A multi-band EEG transformer for brain states decoding,” *IEEE Trans. Biomed. Eng.*, vol. 71, no. 5, pp. 1442–1453, May 2024.
3. G. Siddhad, A. Gupta, D. P. Dogra, and P. P. Roy, “Efficacy of transformer networks for classification of EEG data,” *Biomed. Signal Process. Control*, vol. 87, Art. No. 105488, Jan. 2024.
4. M. Omer, M. Salmanpour, and M. Lataifeh, “EEG Data Obtained From Students As They Are Engaged in an Online Learning System,” 2024.
5. W. Sandoval, “Epistemic Cognition,” *International Society of the Learning Sciences (ISLS)*, 2024.
6. Gencoglu, S. Hussein, A. Salimi-Badr, and N. T. Heffernan, “EEG Transformer for classifying students' epistemic cognition states in educational contexts,” in *Proc. 16th Int. Conf. Educational Data Mining (EDM)*, 2023.
7. Z. Wan, M. Li, S. Liu, J. Huang, H. Tan, and W. Duan, “EEGformer: A transformer–based brain activity classification method using EEG signal,” *Front. Neurosci.*, vol. 17, pp. 1–13, Mar. 2023.
8. Kumar, B. Singh, and D. Tiwari, “EEG Transformer for Classifying Students' Epistemic Cognition States in Educational Contexts,” in *Proc. 2023 Int. Conf. on Cognitive and Learning Sciences (ICCLS)*, 2023.
9. M. Moinnereau, A. Rakotomamonjy, and A. Habrard, “A Transformer-Based Architecture for Epistemic Cognition Classification Using EEG Data,” *IEEE Access*, vol. 10, pp. 41638–41649, 2022.
10. Z. Zhang, G. Zhao, J. Li, and A. Cichocki, “EEG-based intention recognition using deep learning: A review,” *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1470–1483, 2021.
11. R. Singh and P. Singh, “Electroencephalography (EEG) based emotion recognition using machine learning techniques: A review,” *Materials Today: Proceedings*, vol. 33, pp. 1756–1761, 2020.
12. Craik, Y. He, and J. L. Contreras-Vidal, “Deep learning for electroencephalogram (EEG) classification tasks: A review,” *Journal of Neural Engineering*, vol. 16, no. 3, p. 031001, 2019.
13. Vaswani et al., “Attention is all you need,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.
14. A.Greene, R. Azevedo, and M. Wiedmann, “Cognitive and metacognitive processes during self-regulated learning with hypermedia: Issues in data collection, interpretation, and analysis,” *Educational Psychologist*, vol. 47, no. 2, pp. 104–123, 2012.

Future Scope:

Although the current system successfully classifies students' epistemic cognition states using EEG signals and Transformer-based deep learning, there are several directions in which this work can be extended and improved in the future.

- **Implementation with Low-Cost EEG Hardware:**
  Although the proposed system is designed as a low-cost EEG-based solution, the current work focuses on validating the model using publicly available EEG datasets.

In future, the system can be integrated with affordable EEG devices such as Muse or Emotiv headsets to enable real-time data acquisition and practical deployment in classrooms.

- **Real-Time Cognitive State Monitoring:**
  The system can be extended to support real-time EEG signal processing, allowing continuous monitoring of students during live learning sessions. This will help in identifying cognitive changes instantly and improve learning outcomes.
- **Early Detection of Learning Difficulties:**
  Future enhancements can focus on identifying specific learning difficulties such as lack of attention, cognitive overload, and slow comprehension. This will enable educators to provide early intervention and personalized support to students.
- **Integration with Educational Platforms:**
  The system can be integrated with Learning Management Systems (LMS) to create adaptive learning environments. Based on detected cognitive states, the system can recommend suitable learning materials and adjust content difficulty.
- **Extension to Multiple Cognitive and Emotional States:**
  The model can be expanded to include additional cognitive and emotional states such as stress, engagement, frustration, and motivation, providing a more comprehensive understanding of student learning behaviour.