Modeling the development of intuitive mechanics

Tianyi Li[1], Mengguo Jing[2], Zakir Makhani[3], Iris Oved[4], Nikhil Krishnaswamy[5], James Pustejovsky[6] & Joshua Hartshorne[7*]
[1]Department of Psychology and Neuroscience, Boston College; [2]Ohio State University; [3]Northeastern University; [4]Unaffiliated scholar; [5]Colorado State University; [6]Brandeis University; [7] Massachusetts General Hospital Institute of Health Professions

Introduction

Background
Children need time to learn and develop accurate predictions about object stability [1-2]. Previous research [3] suggests that children may learn to apply different rules to combine the information from weights and distances on both sides of the fulcrum. They start to predict from only noticing the weights, comparing the distance when weights are the same, and eventually shift to combine them using the production of the weight and distance.

Aim
In this study, we adapt the classical balance task and test whether computational models follow the same developmental stages as children learn.

The results would explain the gap between children and model simulations, and allow to refine our understanding of how children develop intuitive physics knowledge about balance.

Test materials

Following previous study [3], we designed **4 types of stimuli:**

- **Group 1 (G1):** Weight items: items have unequal weights but equal distances from the fulcrum
- **Group 2 (G2):** Distance items: items have equal weights but unequal distances from the fulcrum
- **Group 3 (G3):** Conflict-weight items: items have unequal weights and unequal distances, and will tip to the side with a larger weight
- **Group 4 (G4):** Conflict-distance items: items have unequal weights and unequal distances, and will tip to the side with a larger distance

Each group includes 10 variants of different width and height, with 100 randomized position on the platform, resulting in 1000 stimuli per group. The number for stay and fall examples are balanced.

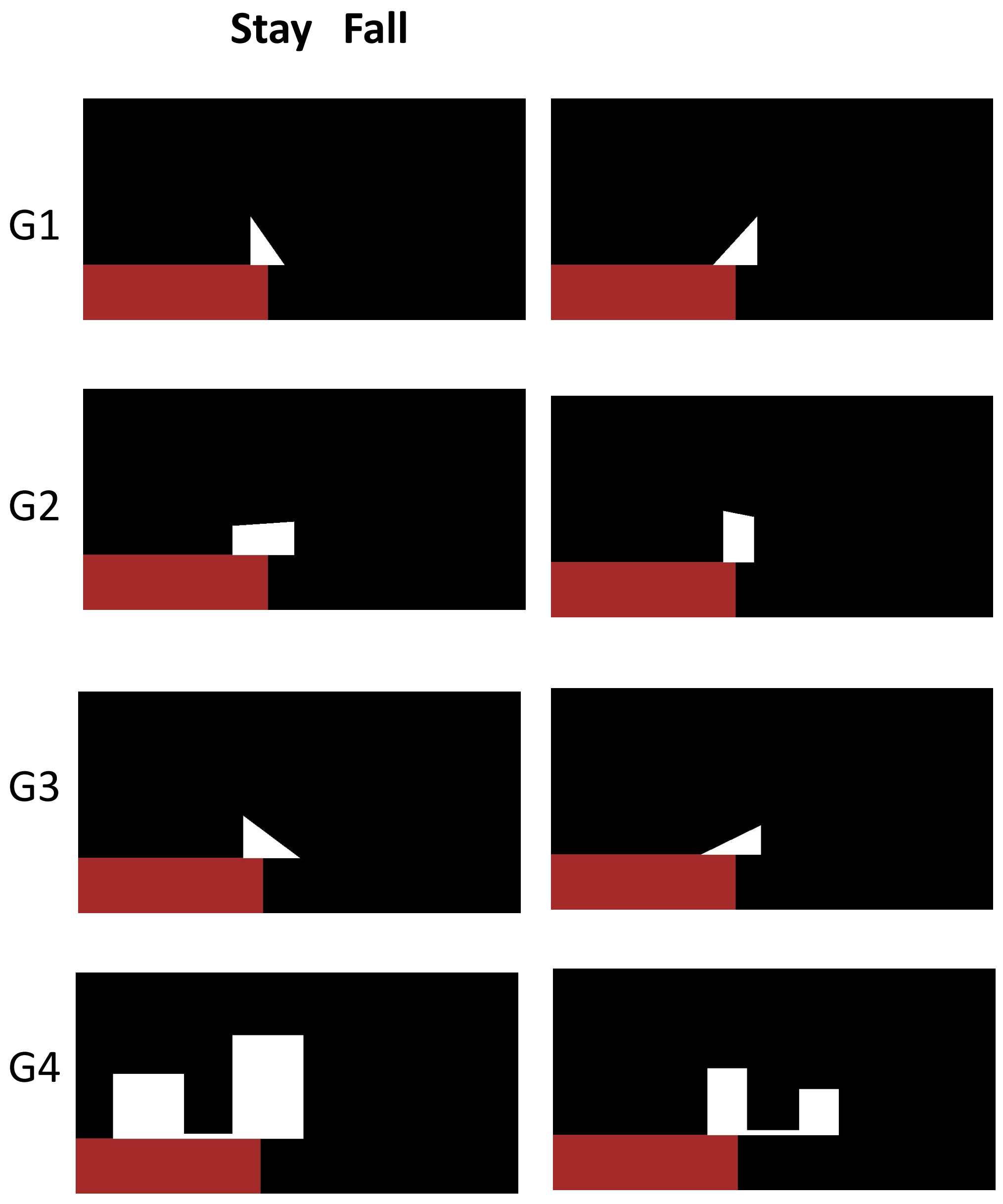

Figure 1. Examples of the 4 groups of the test stimuli

Methods

We implemented and trained Convolutional Neural Networks (CNNs). The architecture as Figure 2 shows.

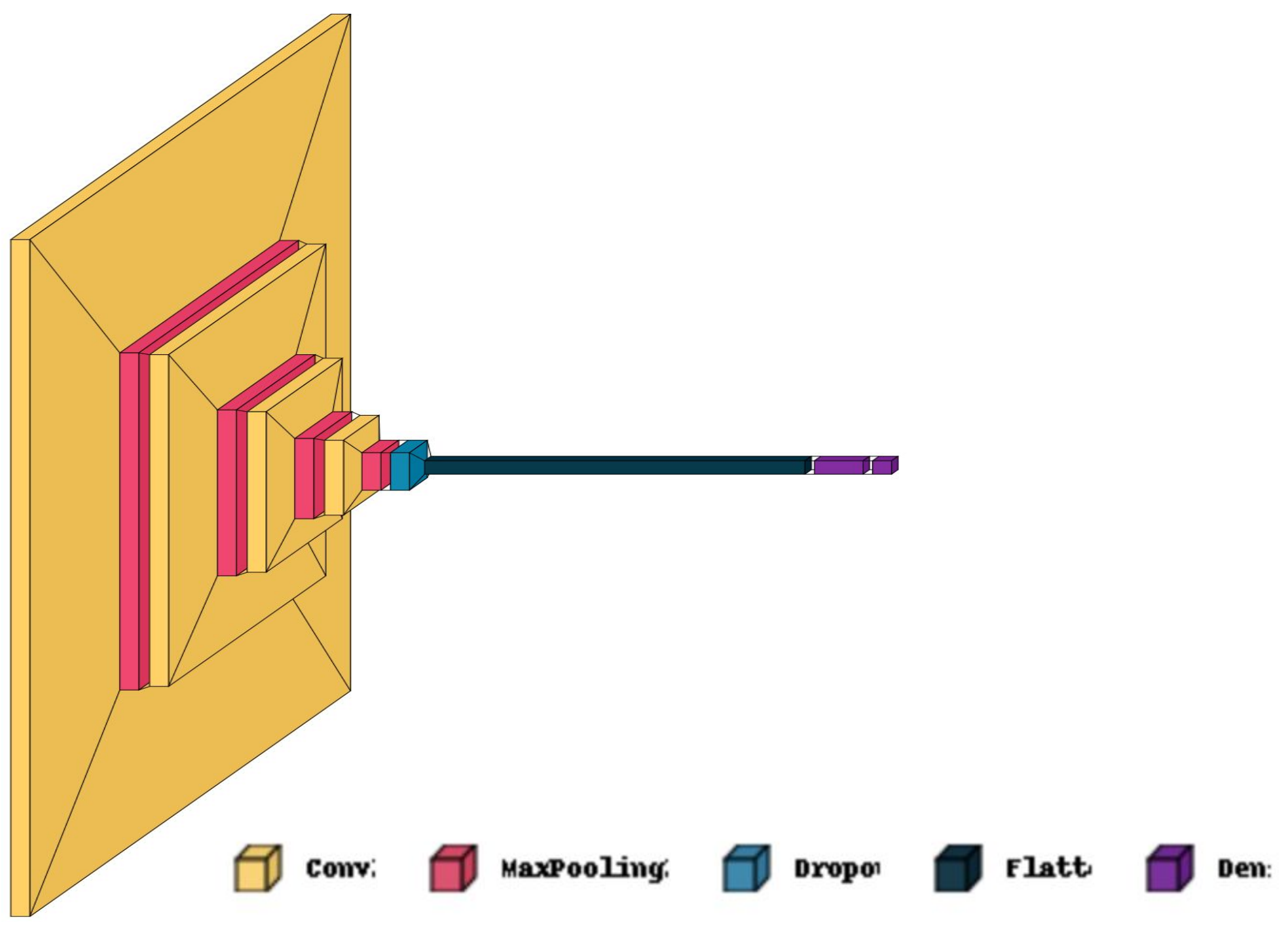

Figure 2. Architecture of the CNN model

Children may not see situations exactly designed as our 4 groups (Figure 1). We create more variants of the stimuli for training the model to learn to predict the balance status of objects in a more general setup. We use 4 novel types of stimuli for training (see Figure 3). Then test the trained model with the 4 types of groups we are interested in.

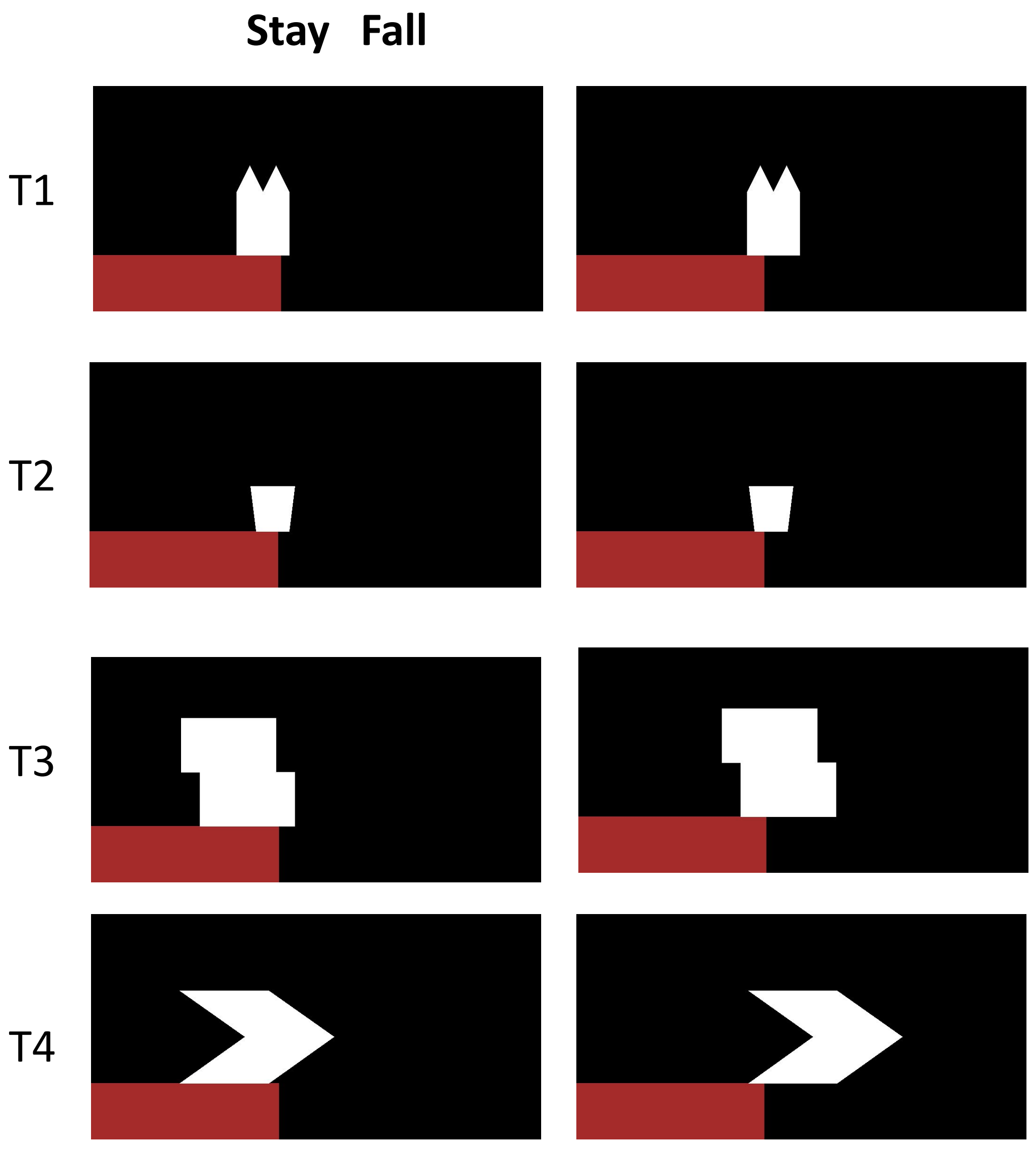

Figure 3. Examples of the 4 groups of the training stimuli

Results

- Overall, the models perform better on train/validation set than test set
- The trend for test on the 4 groups are stable for the models and training process we used:
  - Group 1-3 accuracy increase as training goes
  - Group 4 is relatively high at the beginning but then decrease during training
  - Group 3 accuracy remains lower than Group 1 & 2 although the training
- Comparing to what we expect from children based on previous evidence, the model learning curve is different

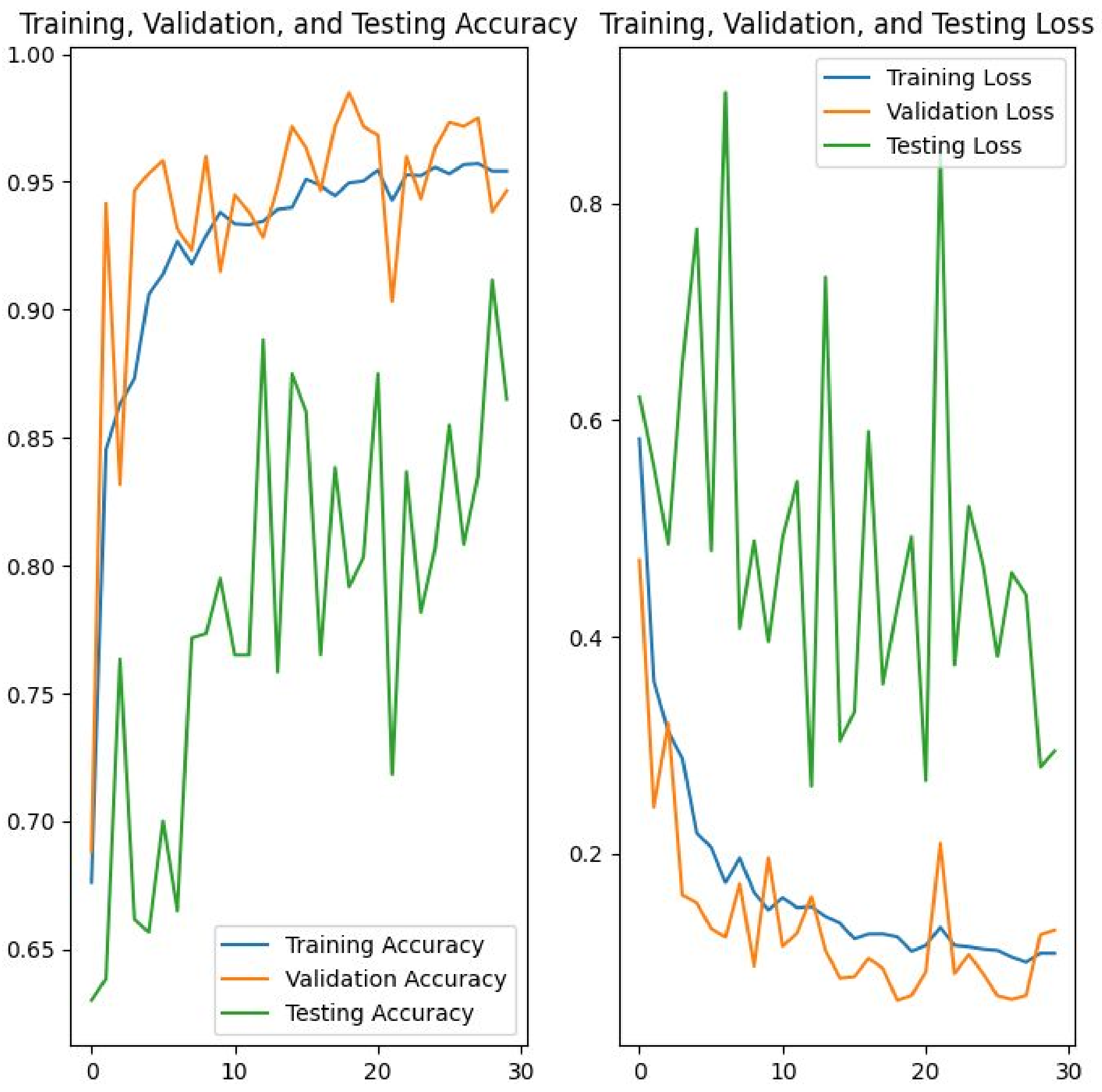

Figure 4. Accuracy and loss curves for train/validation/test set (from a single CNN)

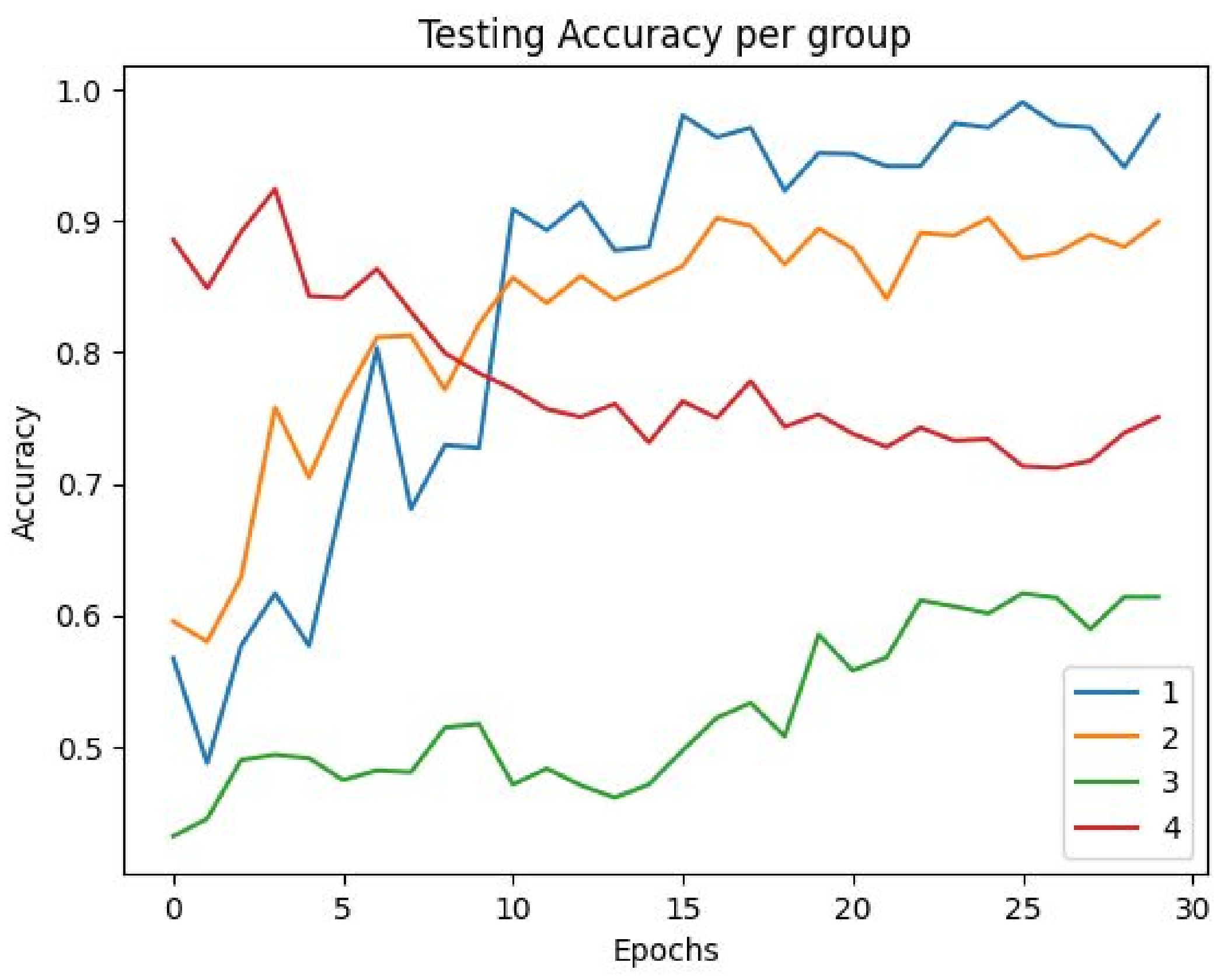

Figure 5. Test set accuracy per group (averaged for 10 CNNs)

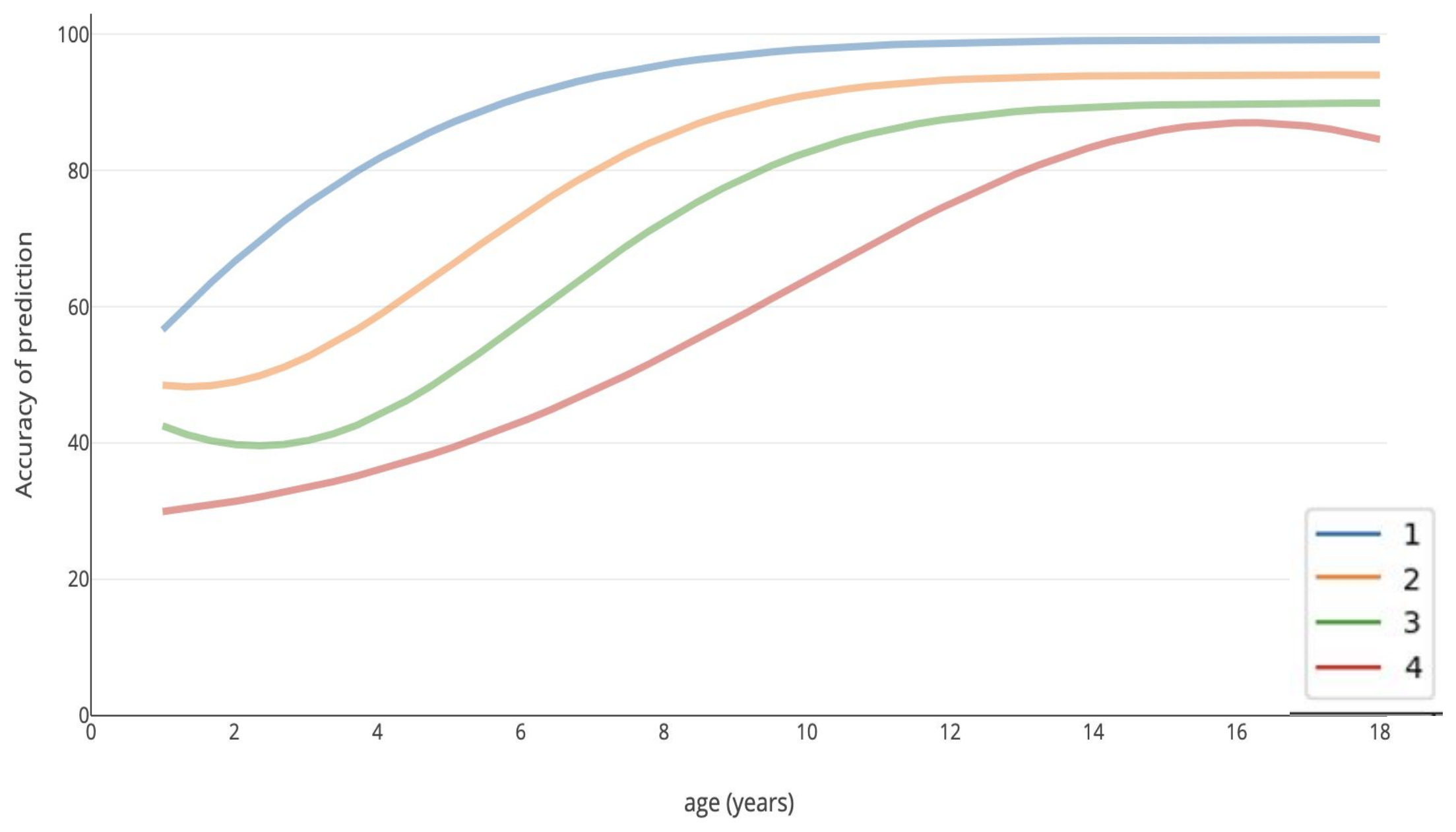

Figure 6. Expected trend in children (not based on actual data points)

Discussion & Conclusions

- CNNs may follow similar stages as children learn balance support task
- Explanations for why Group 4 decrease is not clear
  - Overfitting?
  - Different stimuli size/different latent features/position variants range for Group 4 stimuli
  - etc.etc.
- How about other kind of models? Would they go through similar stages in the balance task as children?

References

1. Karmiloff-Smith, A., & Inhelder, B. (1974). *If you want to get ahead, get a theory. Cognition,* 3(3), 195-212.
2. Lin, Y., Stavans, M., & Baillargeon, R. (2022). *Infants' physical reasoning and the cognitive architecture that supports it. Cambridge handbook of cognitive development,* 168-194.
3. Jansen, B. R. J., & Van Der Maas, H. L. J. (2002). The Development of Children's Rule Use on the Balance Scale Task. *Journal of Experimental Child Psychology*, *81*(4), 383–416.