A Appendix

A.1 Hyperparameters

A.1.1 PLay Pretraining

We train a text-conditioned PLay model. We remove the guideline condition and replace it with a text condition, which uses text embedding features from a BERT model with 12 layers, 12 attention heads, and hidden size 768. We inject text conditions through element-wise condition on pooled text embeddings and cross attention with the full text embedding.

The rest of the hyperameters that we used are equivalent to those in Cheng et al. (2023). We train the model on 8 Google Cloud TPU v4 cores for 40,000 steps with batch size 1024.

A.1.2 Reward Model Training

We include the hyperparameters for training in Table 2.

	Reward Model Pretraining		**Finetuning**
Method	**CLAY Pretrain Steps**	$\mathcal{D}_{\textbf{human}}$ **Train Steps**	**Optim. Steps**
Supervised Finetuning	x	x	70,000
RARE	1,400	600	200
Preference Reward	2,000	200	800
Chamfer Reward	49,000	7,000	400

Table 2: Reward Model Training Hyperparameters.

RARE and the Preference Reward Model have the same architecture as the denoising diffusion model used in PLay, with the exception that there is no time embedding, and there is an additional MLP layer that reshapes the output features and projects it to a scalar prediction. For the Chamfer Reward Model, we reduce the number of layers to 2, number of heads to 4, and key, query and value dimensions to 256 to prevent overfitting.

A.2 RLHF Hyperparameters

We train with sample batches of 256. In accordance with DDPO, we compute losses for a single timestep across denoising timesteps together. We set the PPO clip range to 1e-2. We use a batch size of 64 on 8 Google Cloud TPU v4 cores.

A.3 PLay Color Legend

We use the same color legend as in Cheng et al. (2023) to visualize the layouts. Colors for popular class elements are rendered in Figure 8.

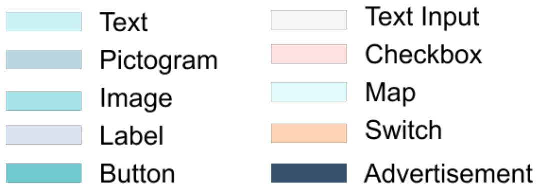
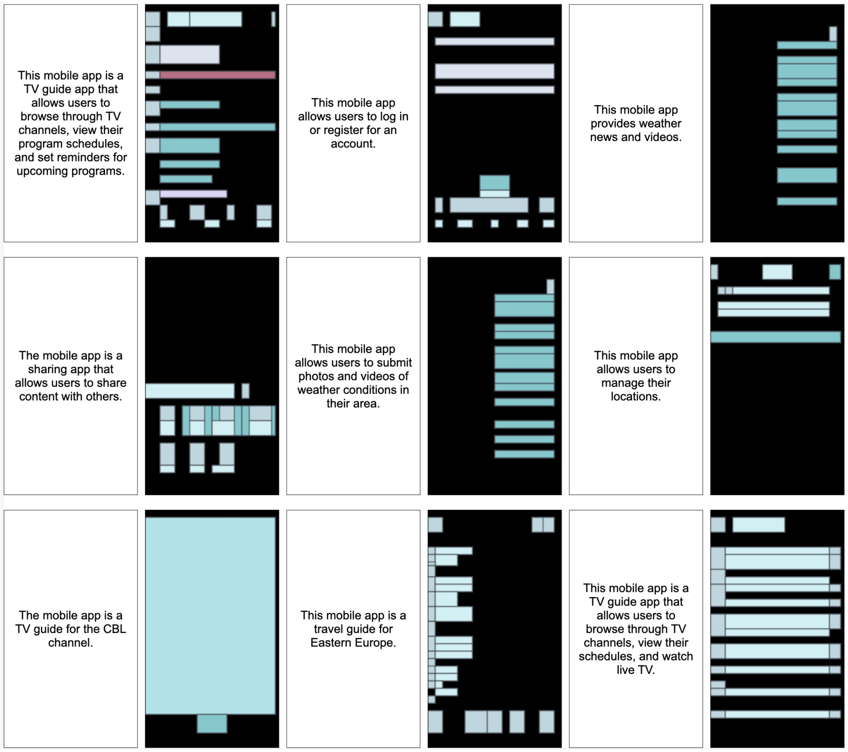


Figure 8: **Visualization Colors**

A.4 Additional Samples



This mobile app is a TV guide app that allows users to browse through TV channels, view their program schedules, and set reminders for upcoming programs.

This mobile app allows users to log in or register for an account.

This mobile app provides weather news and videos.

The mobile app is a sharing app that allows users to share content with others.

This mobile app allows users to submit photos and videos of weather conditions in their area.

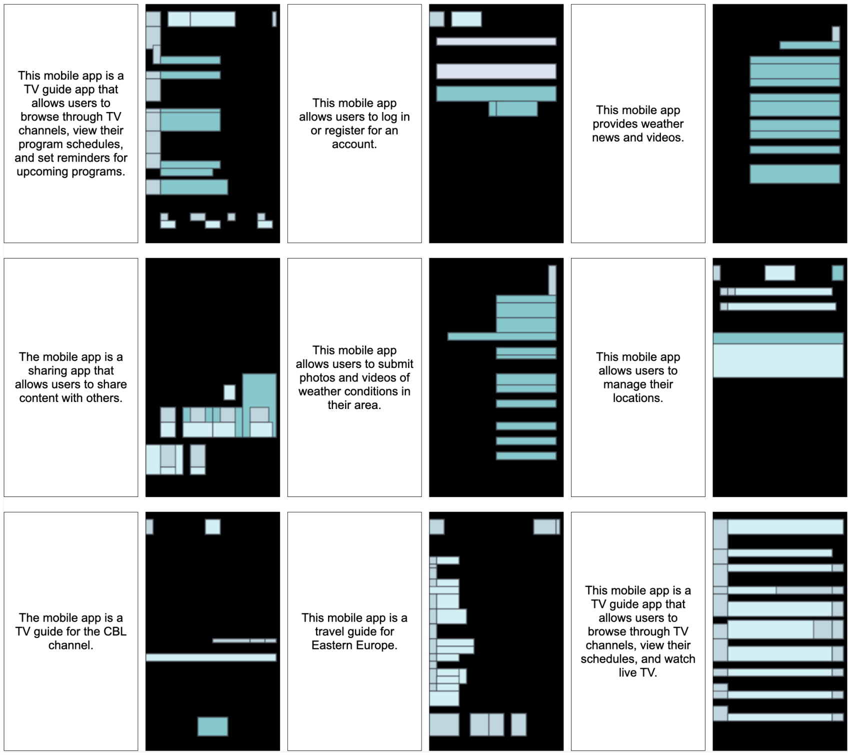
This mobile app allows users to manage their locations.

The mobile app is a TV guide for the CBL channel.

This mobile app is a travel guide for Eastern Europe.

This mobile app is a TV guide app that allows users to browse through TV channels, view their schedules, and watch live TV.

Figure 9: Non-cherrypicked samples from RLHF w/ RARE.



This mobile app is a TV guide app that allows users to browse through TV channels, view their program schedules, and set reminders for upcoming programs.

This mobile app allows users to log in or register for an account.

This mobile app provides weather news and videos.

The mobile app is a sharing app that allows users to share content with others.

This mobile app allows users to submit photos and videos of weather conditions in their area.

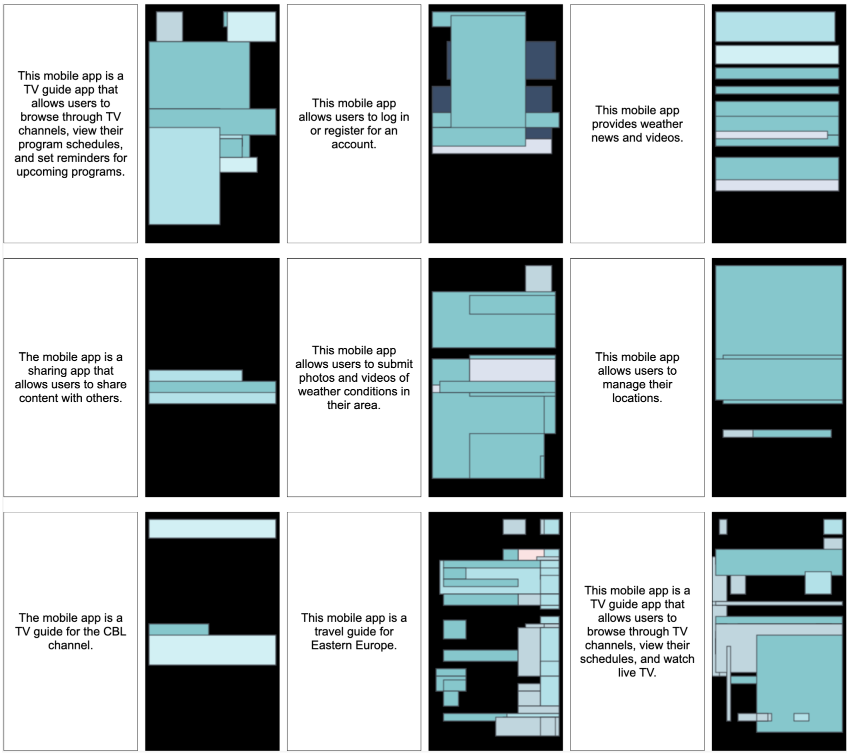
This mobile app allows users to manage their locations.

The mobile app is a TV guide for the CBL channel.

This mobile app is a travel guide for Eastern Europe.

This mobile app is a TV guide app that allows users to browse through TV channels, view their schedules, and watch live TV.

Figure 10: Non-cherrypicked samples from RLHF w/ a preference-based reward model.



This mobile app is a TV guide app that allows users to browse through TV channels, view their program schedules, and set reminders for upcoming programs.

This mobile app allows users to log in or register for an account.

This mobile app provides weather news and videos.

The mobile app is a sharing app that allows users to share content with others.

This mobile app allows users to submit photos and videos of weather conditions in their area.

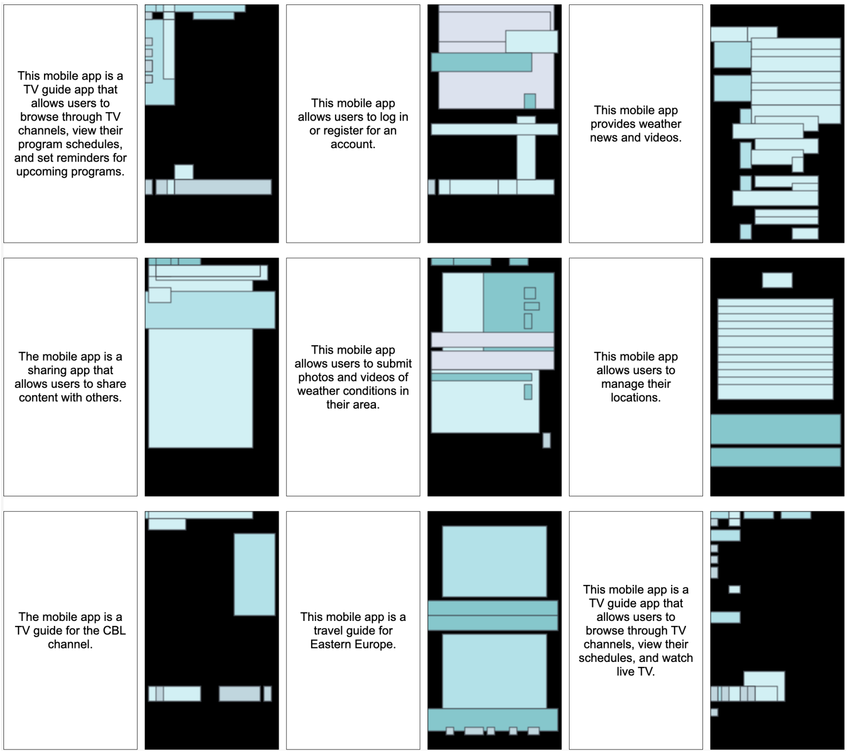
This mobile app allows users to manage their locations.

The mobile app is a TV guide for the CBL channel.

This mobile app is a travel guide for Eastern Europe.

This mobile app is a TV guide app that allows users to browse through TV channels, view their schedules, and watch live TV.

Figure 11: Non-cherrypicked samples from RLHF w/ a Chamfer distance reward model.



This mobile app is a TV guide app that allows users to browse through TV channels, view their program schedules, and set reminders for upcoming programs.

This mobile app allows users to log in or register for an account.

This mobile app provides weather news and videos.

The mobile app is a sharing app that allows users to share content with others.

This mobile app allows users to submit photos and videos of weather conditions in their area.

This mobile app allows users to manage their locations.

The mobile app is a TV guide for the CBL channel.

This mobile app is a travel guide for Eastern Europe.

This mobile app is a TV guide app that allows users to browse through TV channels, view their schedules, and watch live TV.

Figure 12: Non-cherrypicked samples from the Supervised Finetuning model.