Notes about CHRL human task

* The environmental response to each action sequence will be represented by a different symbol (sequence-symbol associations random or indicative of level?) -> each must be unique so that each is novel the first time around
* Symbols will be presented at random locations on the screen or according to level in the hierarchy?
* Symbols will stay on the screen until participants selects the next action -> needs to be able to move fast
* How many basic actions, how many levels, how long the action sequences?
* The same environment for everyone of a different one for every one?

Test conditions: Participants will interact with this general paradigm in different conditions, each of which will allow us to test specific hypotheses about hierarchical learning.

* Just explore, i.e., no rewards
  + show that learning occurs in the absence of rewards
  + determine which mechanisms underlie this learning? HRL suggests that sub-goal achievement acts as a pseudo-reward; test this.
* Transfer paradigms: after spending some time in one game, be transferred to a different game, with different background color, different symbols, at different locations for the responses; assess learning curves
  + Action sequences, at different levels, are either totally different from the ones in the original world; identical; or mostly identical but 1 out of many things has changed
  + Question: does interference impact different levels differently? Prediction: more basic levels will be affected more because they are more petrified, less flexible.
* Reach specific goals (collect as many objects of a certain kind as possible)
  + See how intrinsic (pseudo-rewards) and extrinsic (rewards) motivation are combined; how do extrinsic rewards affect sub-goal selection and curiosity?

Planned analyses and expected results:

Behavioral analyses.

* Evidence for the formation of discrete action chunks
  + Slower RTs between chunks than within chunks
  + Larger transition probabilities between actions within the same action sequence than across action sequences
  + Transition between item1 and item2 in a chunk increases the transition probability between imte2 and item3 in the same chunk (etc.)
* Evidence for hierarchical learning: lower-level chunks are combined into higher-level chunks
  + Look at higher-level sequences that are performed more often:
    - Transition probabilities are higher between two actions when these actions are the last and first actions of two chunks that occur in the sequence, than when they are not
    - RTs are shorter in the same cases
    - Both change gradually over time; plotting this change should reveal something about how the sequence is learned (faster change -> more learning)
    - These changes should be correlated between transitions within the same chunk, but should not be correlated to transitions within other chunks (that have been picked equally often, and are at the same level of hierarchy)
  + Look at action sequences at different levels:
    - Are lower-level ones learned sooner than higher-level ones? (RTs and transition probs are indices of learning)
    - How many repetitions are necessary to learn lower-level ones compared to higher-level ones? Does this depend on how proficient participants are at the building blocks? Or does it depend on WM?
    - We can assess whether people prefer to develop a small number of very abstract skills, or to acquire a broad of set of less abstract skills.

Computational modeling.

In order to create a good computational model of human behavior in this task, we will take a standard model fitting approach. The most relevant steps will be the following (for a full description, see Daw book chapter). We will first create alternative models to the one presented above, in which crucial elements, such as curiosity and sub-goal learning, will be subsequently removed. We will then fit each model to participants’ behavioral data and calculate model fits. The winning model is the one that is able to reproduce human behavior most closely. We will also conduct behavioral analyses on the simulated data in order to verify that the signatures of human behavior mentioned above are reproduced.

Once a model with a good fit to human data is found, which also reproduces relevant signatures of human behavior, it can be interpreted. We expect the hierarchical model to be the winning model, which would suggest that both curiosity and sub-goal selection are involved in human hierarchical learning. The model also allows introspection into the hierarchical learning process. We can assess whether model-based curiosity matches up with how often participants perform each sub-goal, and whether the accuracy of model-based policies matches up with behavioral markers of policy knowledge (short RTs within; item1->item2 predicts item2->item3; etc.)

Potential problems & alternative strategies:

* The current model might not fit -> this is normal, it always requires tweaking -> analyze the behavioral data carefully to inform the model
* Curiosity and/or sub-goal selection might not be relevant, i.e., the winning model might be a different one than the one expected -> this is fine; we expect this model to perform best based on previous research, but if another model is better, we will learn that other processes are indeed more relevant.
* We can’t find any model -> we still have the behavioral markers that can tell us many details about human hierarchical learning