We thank the reviewers for their thoughtful feedback. We appreciate the positive comments describing the paper as “new and interesting”, “technically sound”, "effectively supports claims", "well written" and “easy to follow”.

A common question among reviewers was how the approach can be extended to handle more complex visual environments. Indeed this is an important discussion point and a fruitful area for future research. We will integrate the following discussion into the paper.

As R1 points out, this work is the first to extend HER to visual domains without explicit goal conditioning. We show that for certain tasks, retroactive goal reassignment can be done by directly inserting a hallucinated goal into state trajectories while largely leaving the background unaltered. This allows HALGAN to focus on goal generation and not unnecessary background details, making its training sample efficient. Yet, in other visual environments, extra information in the state such as occlusion or background may influence goal generation. For this, the generative model in HALGAN can be extended to condition on extra variables such as the agent state(s) or a map of the environment if available. More broadly, one can think of future methods that use generative models to retroactively alter goals in visual states as lying on a range of how much of the original trajectory they alter and what state variables they condition on. It is also important to note that the advantage of hindsight methods is greatest when a large part of the original failed trajectory can be reused with a new goal.

Now we address individual reviewer questions.

R1 *“How does it handle more complicated visual input...”* HALGAN synthesizes goal images conditioned solely on desired relative location, hence is independent of clutter or distracting information that may occur in the state. Multiple possible solutions of goal image can be generated by the GAN if there is enough support in the training set. If state dependent goals are desired, the generative model will have to be conditioned on agent state (see above discussion).

“How does it enforce temporal consistency?” Temporal consistency is not explicitly enforced by HALGAN. Instead the model generates the hallucinated states in a transition independently, relying on the relative configuration to the final agent state. A change in relative configuration between two states will manifest in a different hallucination by the trained model, but there’s no guarantee that goal features independent of relative configuration will be temporally consistent. In practice, for our environments, we found that this was not a burden to RL training as only two consecutive states are used for making off-policy updates. An extension of our approach could maintain some recurrent, hidden state throughout the failed trajectory that is provided as input to a generative model during hallucinations.

R2 *“Are the 1000-samples used to train the HALGAN shown in Figure 3(f)...”* We include some details in section 5.3 Data Collection, but will expand on them further here and in the paper. The states in figure 3(f) are near goal snapshots, in which the relative configuration of the goal is known. In our experiments, we rely on a pre-collected dataset of the last 16 or 32 states of successful demonstrations, so that relative configurations can be automatically calculated using only the agent state as long as the agent ended at the goal. Alternatively, these data may be collected online as the agent explores by manually annotating a small set of failed initial episodes with relative goal location. Once HALGAN has sufficient training data, it can generalize to future episodes without annotation. The burden of collecting goal information for HER is not entirely eliminated, but can be significantly reduced to only a few thousand states. We did not enforce that the goal be visible in all collected states, but despite this there were enough data for the GAN to infer the object of interest.

“...why the generator of HALGAN does not input the s_t that it is trying to modify.” Please refer to common discussion above (L6-L15). For tasks where occlusion may play a large role, the generative model can be extended to condition on agent state or a short term memory over the trajectory.

"examples where this is the case in the real world could help..." The principle of visually hallucinating goals can be applicable to many other tasks such as avoiding collisions with objects (eg. negative penalty for hallucinations involving hitting pedestrians), following a human/object (eg. positive reward for hallucinations with person at constant distance), placing objects in visually identified zone (eg. hallucinating a visual marker where objects can be safely placed), etc.

R3 *“...how the authors intend to scale up this approach to more complex visual domains like Atari, DeepMind Lab etc.”* With enough training data, HALGAN should work in DeepMind Lab maze tasks of seeking out “apples”, which can be hallucinated on the background in the same way as we have shown in our environments. Atari games are generally not amenable to the hindsight family of algorithms as they do not have multiple (visual) goals that can be substituted in retrospect for each other. As such, we have not seen any examples in the literature that attempt HER in Atari or similar environments that do not possess this crucial property. On how to scale to other visual environments where goals may be dependent on state or occlusions may occur, please refer to common discussion above (L6-L15).

We once again thank all reviewers for their useful comments. We will include the responses in the final submission.