We are grateful for the reviewers' constructive feedback.

Novelty & limitations: Reviewer 4 (R4) correctly points out that there exists a rich literature on meta-learning learning rules for neural networks, and we did not adequately cite this literature. Regarding the high-level claims we made that credit assignment may be viewed as an optimization problem and that there may be effective alternatives to gradient descent for the problems we consider, we agree that our work is not the first to provide support for these ideas and rather complements the references the Reviewer mentions by proposing and analyzing a concrete instantiation of them. We agree with the Reviewer that our "specific implementation has sufficient novelty and should be of interest to the community" and that, going forward, more work is needed to compare it to/integrate it with alternatives. We will both cite the suggested references and be more nuanced in our claims about our specific advances in a final version.

R4 also points out an important limitation of this work, and most modern meta-learning studies in general, namely that the meta-learned learning strategies are tailored to rather narrow task distributions. We agree that scaling our method, e.g. by meta-training on a large, diverse set of tasks, is a high priority for future work. As R2 suspects, computational constraints make doing so a challenge – we think tackling this challenge is worthy of another paper. We believe that our paper, which uses contemporary (if imperfect) benchmarks from the meta-learning literature, is worth disseminating to the community and will facilitate progress toward these even more ambitious goals.

Training & performance: R3 and R4 bring up the fact that our reported results involve plasticity in only the last several layers of the network. To answer R3's question, this is not because the meta-training becomes unstable when more plasticity is allowed – for instance, in the regression task, we have found that enabling plasticity in all layers gives similar performance to enabling it in 3 layers. We suspect that the performance saturates as a function of the number of plastic layers because of the nature of the tasks, rather than a fundamental limitation of our algorithm, as we find the same trend holds for gradient-based meta-learners. We suspect that different or more complex tasks might require even deeper plasticity. We will include these results and discussion in the camera-ready version of the paper.

R3 asks whether the feature reuse baseline would be stronger if we used a shallower network. We have now checked on the regression task, and the performance is much worse using a 3-layer network (1.35 MSE) and about the same using a 6-layer network (0.05 MSE) compared to the 9-layer network that we used (0.05 MSE). Thus, we think ours is indeed a fair baseline – it is not suffering from the network depth.

R2 asked how the learning speed of our method compares to that of the gradient-based baseline. Both the outer-loop and inner loop learning speed are comparable. For instance, on the i.i.d. Omniglot task, both methods take 20,000 epochs to reach 5% error and 30,000 epochs to reach 3% error. In the inner loop, FLP learns a bit more quickly in early iterations. We will add learning and meta-learning trajectories to the final version of this paper.

R4 asked how we used batches. In classification experiments, no batches are used (this is what "one example at a time" referred to). In the regression experiments, the inner loop consists of 400 size-32 batches (to allow more examples to be used in the inner loop while maintaining computational tractability). Note that each example is only presented once. We will be more clear about these methods in a final version of the paper.

Implementation choices: Several Reviewers touched on assumptions concerning the nature of the feedback, including fixed (within-lifetime) rather than plastic feedback weights and direct feedback from the readout layer. We made the implementation decision of fixed feedback weights for simplicity, and to demonstrate how (remarkably) far one can get with this approach. Going forward, plasticity of feedback weights is a natural extension, and may indeed be important for scaling the method to harder and more diverse problems. Regarding the use of direct feedback from the output, rather than from each hidden layer to the previous one, this was motivated by biological plausibility considerations, as it remains an open problem whether and how error signals can be multiplexed with feedforward signals and also transmitted backward to earlier network layers (though there are exciting proposals, e.g. Payeur et al. *bioRxiv* 2020).

R1 brought up the use of targets vs. errors for feedback. The idea of using both simultaneously is intriguing – we are interested in exploring this, but it will take some time. We should note that our implementation decisions – targets for classification, errors for regression – yield better performance both for FLP and the gradient-based baseline, so this does not appear to be a quirk of our method. We will elaborate on this in the analysis section, as suggested.

R1 mentioned the difficulty in comparing our method to backpropagation, given that the outer loop is able to specify weight initializations. We want to emphasize that our gradient-based baseline also used a meta-learned initialization for fair comparison. We agree that full specification of the initialization is likely unbiological (at least for sufficiently complex animals), and as mentioned in the Discussion, we hope to address this issue in future work.

R2 pointed out that backpropagation is used to learn the feedback weights in the outer loop. This is true, and we do not attempt to model this outer loop optimization in a biological fashion. We consider the outer loop to be roughly analogous to the biological processes of evolution and development and thus not subject to inner-loop locality constraints. That said, we hope that follow-up work will model this outer-loop learning in a more biologically plausible fashion.