We thank the reviewers for their constructive and valuable comments. The resulting revisions and additional experiments significantly strengthened the paper. In the following three sections we summarise and address all major concerns in detail. All other comments will be addressed as well, but not discussed here due to space limitations.

Biological Plausibility. Several reviewers raised concerns whether DDTP and DRL are still biologically plausible (rev. comments 2.iii, 3.1, 3.2, 4.5). (i) The purpose of our work was not to validate TP and its variants to be bio-plausible. Instead, we aimed to mathematically analyze the principles of TP/DTP optimization, to uncover strengths and weaknesses and to establish a theoretical framework that allows us and other researchers to address the latter. For example, we linked TP to GN and GD and showed that layer-wise DTP training of feedback paths leads to inefficient parameter updates. Next, we provided a theory-derived solution to address this (DRL). Even if such an improved TP variant turns out to be less bio-plausible we think this is still highly valuable information as it sets new grounds for future discussions in the field. (ii) Regarding the direct or skip feedback connections used in DDTP, we clarify that numerous anatomical studies of the mammalian neocortex consistently reported such direct feedback connections in the brain. In primate visual cortex, both V4 and area MT back-project to V1 (Ungerleider et al., Cereb. Cortex **18**, 2007; Rockland & Van Hoesen, Cereb. Cortex. **4**, 1994). We therefore argue in the revised paper that the flexibility to allow for direct feedback is a major advance in bio-plausibility, compared to methods that only allow strict layer-wise feedback. (iii) DRL requires coordinated noise level alteration to separate reconstruction loops in time which might be biologically questionable. Coordination in time has been used for several major bio-plausible learning methods (Akrout et al., NeurIPS 2019; Kunin et al., ICML 2020) and it is still an open question whether the brain could implement this. Currently, there are several promising paths towards overcoming this need for coordination in time. A first option would be to design a noisy estimator for the DRL derivative where all layers can be noisy simultaneously, similar to Lansdell et al. (ICLR 2020), making use of the correlation between the noise perturbation of layer i and the relevant noise perturbation on the target originating from reconstruction loop i. This would directly address reviewer comment 3.1 on how layer i can filter out the target perturbation that originates from its reconstruction loop. A second option would be to not learn the feedback weights explicitly through a reconstruction loss, but to use a dynamical control system for the inversion (Podlaski & Machens, arXiv, 2020) and adapt it such that it approximates GNT. While both options are interesting they require further examination and testing which would go beyond the scope of this work.

New Experimental Results. Based on suggestions by reviewer 3 we performed new experiments to benchmark the ability of the new TP variants to minimize the training loss. Table 1 shows the new performance results on Fashion-MNIST (other datasets will also be included in the paper) which reveal that the optimization performance of DDTP-linear is strikingly similar to BP while the DTP/DFA methods are inferior by at least one order of magnitude. Complementing the frozen-MNIST experiments in the paper, these new results show that DRL methods substantially improve optimization by feeding back more useful training signals deep into the network, as predicted by our theory. Furthermore, the new results indicate that DDTP-linear (for simple tasks) converges to fixed points of similar depth as BP, even though it does not converge to true local minima of the loss function (rev. comment 1.1).

Table 1: Training loss of last epoch for Fashion-MNIST (mean ± SD for n = 5 seeds).

BP	$\cdot(6.46 \pm 0.25) \cdot 10^{-5}$
DDTP-linear	$\mathbf{(1.03 \pm 0.15) \cdot 10^{-5}}$
DTPDRL	$(1.36 \pm 0.48) \cdot 10^{-3}$
DDTP-RHL	$(3.51 \pm 0.80) \cdot 10^{-3}$
DDTP-control	$(3.88 \pm 2.63) \cdot 10^{-3}$
DTP	$(4.07 \pm 0.42) \cdot 10^{-2}$
DTP (pre-trained)	$(2.73 \pm 0.67) \cdot 10^{-2}$
DFA	$(1.98 \pm 0.24) \cdot 10^{-2}$

From Theory to Practice. All reviewers suggested a more elaborate discussion on how our theoretical insights translate into a practical/experimental setting (rev. comments 1.1, 2.ii, 3.4, 4.2, 4.4). (i) We now discuss in greater detail how the propagated targets for DRL methods are not exactly equal to GNT because of the limited capacity of the feedback parameterization, limited training iterations for the feedback path and the approximation of λ by weight decay and other approximations (see also lines 179-184; 672-721). We discuss as well Figure 2, showing that experimental methods still remove the inefficiencies of DTP, and Figures 4 and S3-S5, demonstrating that our methods well approximate GNT. However, for upstream layers, future studies are required for further improvement, e.g. by investigating better feedback parameterizations or by using dynamical inversion (Podlaski & Machens, arXiv, 2020). A better alignment between targets and GNT in upstream layers will likely improve the performance on more complex tasks such as CIFAR. (ii) We elaborate that λ is approximated by weight decay and is negligible in practice due to the observed implicit damping (lines 1204-1232, rev. comment 1.4 and 4.3). (iii) We now discuss in the paper that although mini-batches of 1 are rarely used on GPUs, they are highly relevant for neuromorphic engineering and bio-plausible networks that use online learning. (iv) Finally, we detail that Theorem 4 applies to general forward mappings and that nothing prevents the GNT framework from being applied to CNNs and other feed-forward architectures. As a proof-of-concept, we now include a small CNN (Conv5x5x32; Maxpool3x3; Conv5x5x64; Maxpool3x3; FC512; FC10) on CIFAR10 with DDTP-linear and DFA with FC feedback. We achieved promising results: test error of $24.38 \pm 0.29\%$ (BP), $23.99 \pm 0.31\%$ (DDTP-linear) and $30.00 \pm 0.74\%$ (DFA), indicating that our theory also applies to CNNs. For comparing with DTP and DTPDRL, careful design of the feedback pathways is needed, which is outside of the scope of this theoretical work.