My title is a little crazy. "Convergence of a Human-in-the-Loop Policy-Gradient Algorithm With Eligibility Trace Under Reward, Policy, and Advantage Feedback".

And my abstract is the following:
"Fluid human--agent communication is essential for the future of human-in-the-loop reinforcement learning in robotics. An agent must respond appropriately to feedback from its human trainer even before they have significant experience working together. Therefore, it is important that learning agents respond well to various feedback schemes human trainers are likely to employ. This work analyzes the COnvergent Actor--Critic by Humans (COACH) algorithm under three different types of feedback---policy feedback, reward feedback, and advantage feedback. We find that COACH can behave sub-optimally, but prove a close variant of COACH converges to optimal policies in all three. We compare our COACH variant with two other reinforcement-learning algorithms: {Q}-learning and TAMER."