FP32 vs MX Loss (with/without attention)