	LR=0.0005	Epochs=20	OPT=SGD	UNSEEN=VAL,NIH	BS = 32		
	BASELINE				**SWAD**		
	Run	Accuracy Seen	Accuracy Unseen		Run	Accuracy Seen	Accuracy Unseen
Baseline vs Swad	1	0.8112	0.5667		1	0.8214	0.5818
	2	0.7321	0.5575		2	0.7232	0.5817
	3	0.764	0.5293		3	0.7231	0.5338
Low epochs I think reason for	4	0.7539	0.5817		4	0.7539	0.5318
not much improvment, Paper says	5	0.7538	0.5718		5	0.7939	0.5818
5,000 iterations gives	Avg	0.763	0.5614		Avg	0.7631	0.56218
good result. With 20 epochs we have around 660					Dif From Base	0.0001	0.00078

	LR=0.0005	Epochs=60	OPT=Adam	UNSEEN=VAL,NIH	BS = 32		
	BASELINE				**SWAD**		
	Run	Accuracy Seen	Accuracy Unseen		Run	Accuracy Seen	Accuracy Unseen
Baseline vs Swad	1	0.7539	0.5818		1		
	2				2		
	3				3		
Trying more epochs this time 60 epochs	4				4		
gives about 2000 iterations which might	5				5		
be enough to see some overfitting happen	Avg	0.7539	0.5818		Avg	#DIV/0!	#DIV/0!
					Dif From Base	#DIV/0!	#DIV/0!

	LR=0.0005	Epochs=60	OPT=Adam	UNSEEN=VAL,NIH	BS = 32		
	BASELINE				**SWAD**		
	Run	Accuracy Seen	Accuracy Unseen		Run	Accuracy Seen	Accuracy Unseen
Swad vs Swad	1				1		
With/Without Shuffling	2				2		
	3				3		
	4				4		
Trying swad with and without shuffling	5				5		
to see results	Avg	#DIV/0!	#DIV/0!		Avg	#DIV/0!	#DIV/0!
					Dif From Base	#DIV/0!	#DIV/0!