Average loss during training across different batch sizes