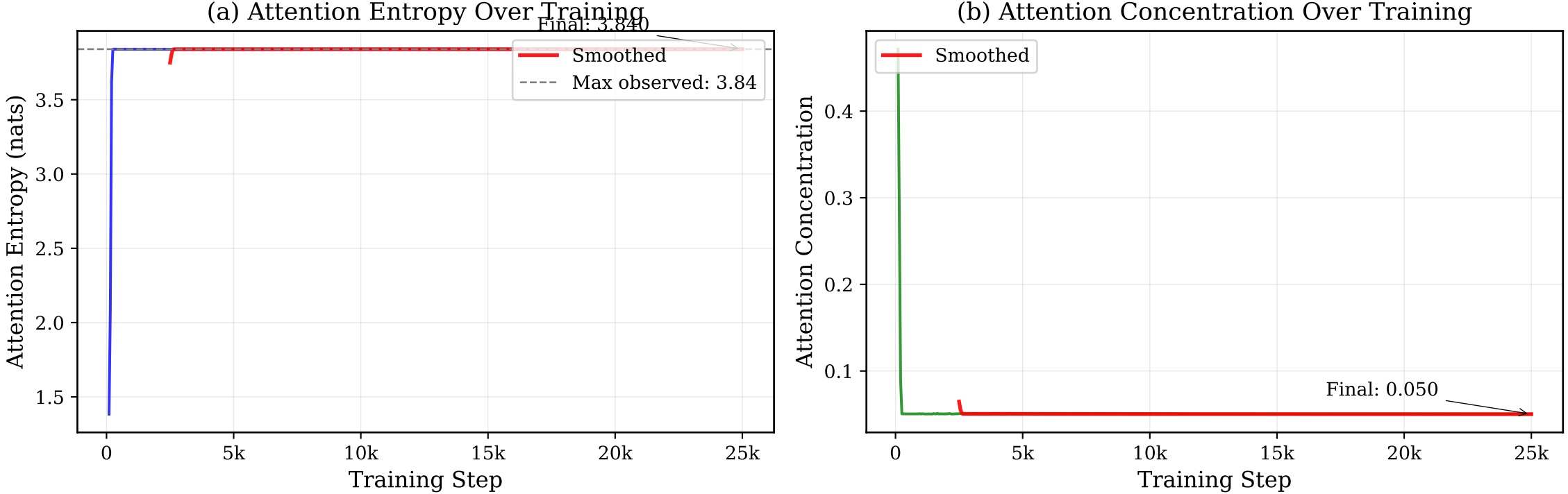

Lower entropy = sharper attention (focused). Higher entropy = uniform attention (diffuse).