Modern neural networks can easily fit their training set perfectly. Surprisingly, they generalize well despite being "overfit" in this way, defying the bias–variance trade-off. A prevalent explanation is that stochastic gradient descent has an implicit bias which leads it to learn functions that are simple, and these simple functions generalize well. However, the specifics of this implicit bias are not well understood. In this work, we explore the hypothesis that SGD is implicitly biased towards learning functions that are smooth. We propose several measures to formalize the intuitive notion of smoothness, and conduct experiments to determine whether these measures are implicitly being optimized for. We exclude the possibility that smoothness measures based on first derivatives (the gradient) are being implicitly optimized for. Measures based on second derivatives (the Hessian), on the other hand, show promising results.