This paper proposes a framework for evaluating the reliability of obfuscating transformations in program code. Many obfuscating transformations have been proposed and implemented in obfuscation tools to prevent malicious analysis and tampering by reverse engineering, such as substituting for an alternative instruction, complicating arithmetic expressions, and flattening control flow. However, there is no framework for evaluating the reliability of obfuscating transformations, in other words, whether this transformation makes the code harder to analyze and comprehend while preserving the input–output relationship for some program code. The proposed framework applies obfuscating transformation the collection of prepare program code, evaluates the reliability of obfuscating transformation in program code from whether all obfuscated program code are successfully tested, whether the original program code is hard to estimate from the obfuscated program code, and whether the program code has been transformed on the whole. The case study evaluates the 37 obfuscating transformations implemented in the well-known obfuscation tools, Tigress and Obfuscator–LLVM. The characteristics of the 14 obfuscating transformations and the reasons why the input–output relationship could not be preserved are investigated from the code that failed test cases. The proposed framework enables obfuscation tool developers and users to evaluate the reliability of obfuscating transformations independently of the obfuscation tools. The result shows that all 37 obfuscating transformations changed programs overall to a certain degree, but 14 obfuscating transformations could not preserve the input–output relationship.