Error control codes are widely used to increase the reliability of transmission of information over various forms of communications channels. The Hamming weight of a codeword is the number of nonzero entries in the word; the weights of the words in a linear code determine the error-correcting capacity of the code. The $r^{th}$ generalized Hamming weight for a linear code $C$, denoted by $d_r(C)$, is the minimum of the support sizes for $r$-dimensional subcodes of $C$. For instance, $d_1(C)$ equals the traditional minimum Hamming weight of $C$. In 1991, Feng, Tzeng, and Wei proved that the second generalized Hamming weight $d_2(C) = 8$ for all double-error correcting $BCH(2^m, 5)$ codes. We study $d_3(C)$ and higher Hamming weights for $BCH(2^m, 5)$ codes by a close examination of the words of weight 5.