**S2. Text.** Establishing reliable response patterns with the rtt

According to Tryon (1957) one first conceptualizes some property, X, of a group of individuals (p 229). In this study, the property X will be the correct identification or sorting of fear and happy mood odors and a control odor on one trial. Next, one define[s] the property Xm in terms of objective specifications that directly lead to the taking of test-sample observations X1, X2, ... Xm believed to elicit the defined behavior, X” (p 229). In this study, such test samples (X1, X2…) are a series of trials in which participants correctly identify three (happy, fearful, and control) odors from the same donors. Lastly, one computes a “composite total score, Xt“ (i.e., the number of correct choices over all trials). From such information we can calculate “within individual variance” and ultimately a reliability score rtt (see Tyron 1957 for computational details) for the detection of fear and happy odors. Using this method, reliability is calculated for every trial and using this we can determine the exact number of trials needed to reach a specified reliability. A high ***rtt***score means that observed phenotypic differences were not due to chance but demonstrate reliable patterns of response. The formula for rtt is

For each trial (1 to n) we can compute a sample variance Vi (i = 1, 2, 3, 4, ....., n). In addition, for each composite score X**t** we can calculate a corresponding variance (Vt) for the composite score. Where **n** is the number of trials and V**t** is the variance of the composite score in trial **n**. **Vi** is the sum of the single trial variances (**V1 + V2 + ... Vn**).