In the context of first impressions, consistency usually stands for consensus between the target’s personality measures, self-report, parent-report, peer-report and matching behaviour. For instance, if a person is extroverted, it is likely that this will be shown in reports produced by close family and acquaintances of target, as well as evident from target’s behaviour. In the other words, the extroversion rating obtained will be consistent across raters. Most studies choose to investigate consensus between some of these aspects. The higher correspondence between the reports and behaviour, the more likely it is that the perception of behaviour observed will be accurate. Below is an example of study that illustrates consistency across perceivers in different exposure time conditions.

Bar, Neta, & Linz (2006) argued that there are two types of threat: “active” threat and “passive” threat. Active threat refers to an active representation of emotional state, specifically anger, which infers immediate danger, whereas passive threat refers to a passive representation of danger, i.e. a person with neutral expression may be perceived as dangerous in general but not dangerous immediately. Following that, we may expect that people will infer threat in actively threatening faces quicker than in passively threatening faces. Thus Bar and colleagues (2006) investigated how quickly people can form impressions that are consistent across observers, independent of their validity. Participants were shown a series of faces and asked to rate to what extent they belonged to a threatening person. The authors inferred the “correctedness” of the answers by observing their consistency with judgements made after longer exposure (26 and 39 ms vs. 1700 ms). They found that threat judgements were highly consistent across 39 ms and 1700 ms exposure conditions (Pearson’s *r=.774, p<.001*), suggesting that 39 ms is sufficient to form consistent threat impressions. In contrast, there was no significant correlation between 26 ms and 1700 ms condition (*r=.256, p>.1*), suggesting that at least some degree of conscious awareness is required to infer threat.

However, consistency does not always mean accuracy. Ratings may be highly correlated but inaccurate if observers share biases or physical stereotypes (Zebrowitz & Collins, 1997). For instance, some observers may tend to believe that people possessing babyish facial traits are in fact more childlike, that they require care, are weak, warm, naive and helpless (Zebrowitz & Montepare, 2008). This may not at all be true; however, such particular set of facial features can mislead observers and their ratings may be high in consistency, although not accurate.