Formative assessment marks were displayed on the notice board. We saw a batch of students in different moods, discussing about the marks they had scored. As a student of MBBS, we used to wonder why half of our batch mates would score more and the other half, of which, we were a part, would get low scores despite satisfactory performance? We got an answer for this question, after becoming a teacher in a medical school. The complexities of assessment, with multiple factors, examiner leniency and stringency contribute to the 'hawk-dove effect'. 'Hawk-dove effect' applies not only to marking theory papers but also to assessment in practical and clinical exams. Hawks tend to fail most candidates, expecting high standards from students both in the theoretical and practical performance, whereas doves tend to pass most candidates.¹

We remember few of the teachers who became famous as terrorist hawks. We were praying god that such hawks do not become examiners for our batch. This tendency continues even after three decades of our undergraduate study.

Though the hawk-dove nomenclature, is well known in literature. There are suggestions to use alternate terms for the same such as, 'Stringent and lenient', or 'candidate centered and patient centered', malignant examiners etc. Hawks or patient centered examiners always belong to the group of expecting high standards from candidates. Candidate centered examiners have sympathy for the student and pass them, whereas patient centered examiners aim to maintain high standards and expect the student to have very good theory and practical knowledge so that he becomes a competent doctor.¹²³

To overcome these hawk-dove effect in assessment of written examinations we introduced a marking system, in department of pathology where each examiner is assigned to mark particular questions. He will evaluate and mark only those particular questions of the entire batch. For example, a formative examination of 60 marks theory paper, will be evaluated by three examiners, each examiner is assigned to evaluate 20 marks. Either the marks or the number of question are divided amongst the examiners. Each examiner marks all the students for 20 marks i.e, each paper is marked by all three examiners.

This ensures uniform scores. If one examiner is a hawk, he harms all the students and if he is a dove, he helps all the students thus overcoming the gross effects of Hawk and Dove.

Other methods of assessment, like multiple choice questions, objective structured clinical examination, objective structured practical examinations etc. avoid the examiner bias, and protect the students from malignant examiners.

Based on these characteristics, we may classify examiners as 'benign' and 'malignant' examiners as well. We recommend university authorities also to adopt an evaluation system which can overcome the 'hawk-dove' effect.

In an editorial, Jankaria B highlights the fact that examiners should not forget to be decent human beings, in response to an email received by a DNB student. The student said 'it is really very pathetic state when results are out anytime with complete wash out of candidates in some cities and overall pass percentage of around 20-25% maximum'. students constantly keep highlighting the maliciousness of some of the examiners who go out of their way to insult students and to demotivate them.²

This type of response we see in medical schools from undergraduate as well as postgraduate students.

In many examinations if one internal examiner is bad, the others balance him/her out and if the external examiner is a terror, the internal examiner usually manages to smoothen things out. The editorial of jankaria b concludes with the following message, which every examiner has to read and follow.

Examiners need to remember that what goes around, comes around! The students will not forget these insults and barbs and, as has happened to "terrors" in the past, it
is very possible that after retirement, these examiners who usually have no sympathy among their former students will be summarily consigned to the dustbin of history. He repeats ’Passing or not passing a student is entirely the onus of the examiner based on the student’s performance. But the least he/she can do is try and be a decent human being whatever he/she may think of the student and his/her ability’.  

It is an accepted fact that examiners do differ in the leniency and stringency and the effect can be estimated using “Rasch modeling”. MacManus IC et al analyzed MRCP(UK) Membership of the Royal Colleges of Physicians of the United Kingdom clinical examination(PACES Practical Assessment of Clinical Examination Skills)using Rasch modeling, to assess the examiner leniency and stringency(hawk-dove) effect. They concluded that reason for the differences are not clear, but there are some demographic correlates, and the effects appear to be reliable across time. Account can be taken of difference, either by adjusting marks or perhaps more effectively and more justifiably, by pairing high and low stringency examiners, so their raw marks can be used in the determination of pass of fail.

References:
2. Jankharia B. Why examiners should not forget to be decent human beings. India J Radiol Imaging 2011;21:79-80