

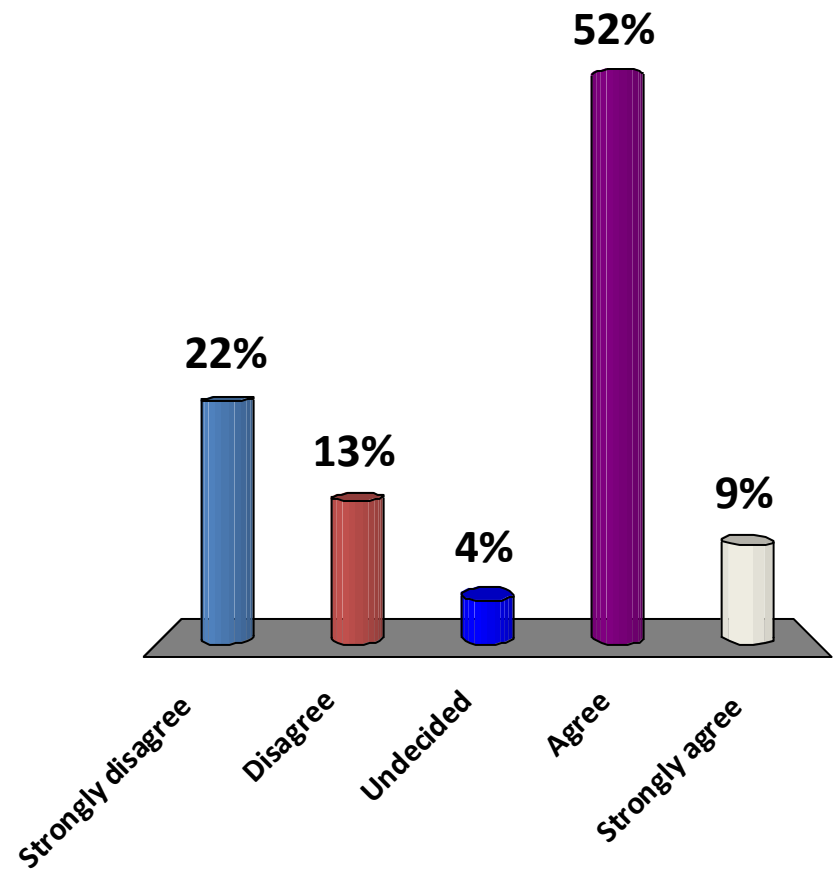
Hawks and Doves

Zeev Goldik (MD) (MPH)

CESMA

The exam is a teaching experience, and therefore, every exam is a lesson and the examiner is a teacher

- A. Strongly disagree
- B. Disagree
- C. Undecided
- D. Agree
- E. Strongly agree



During the exam my sympathy is with the candidate

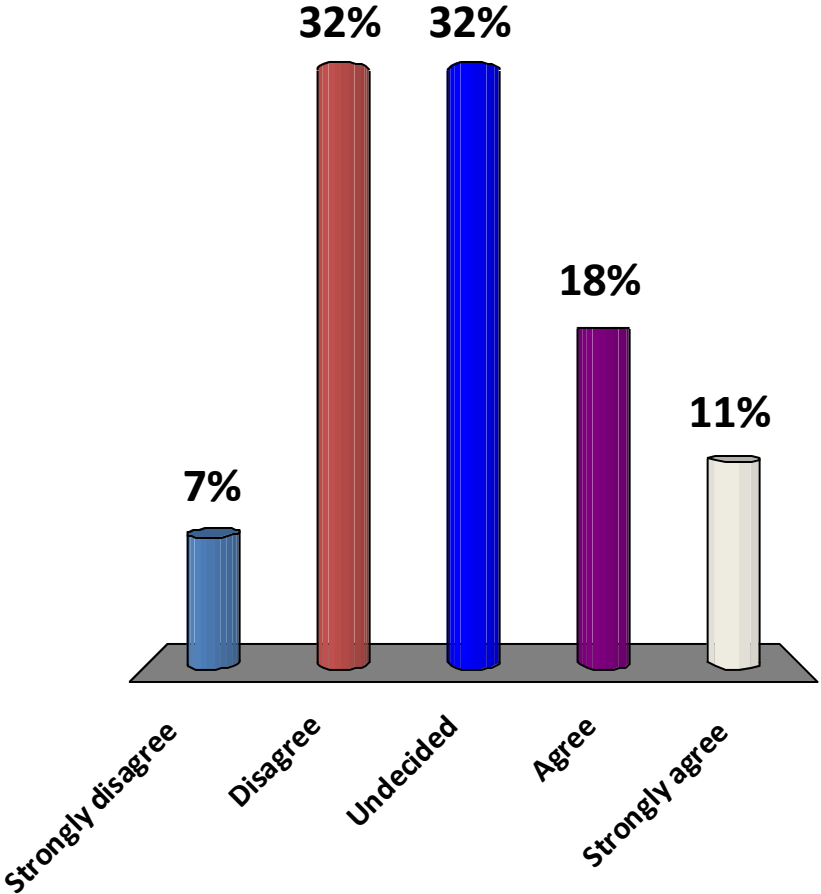
Strongly disagree .A

Disagree .B

Undecided .C

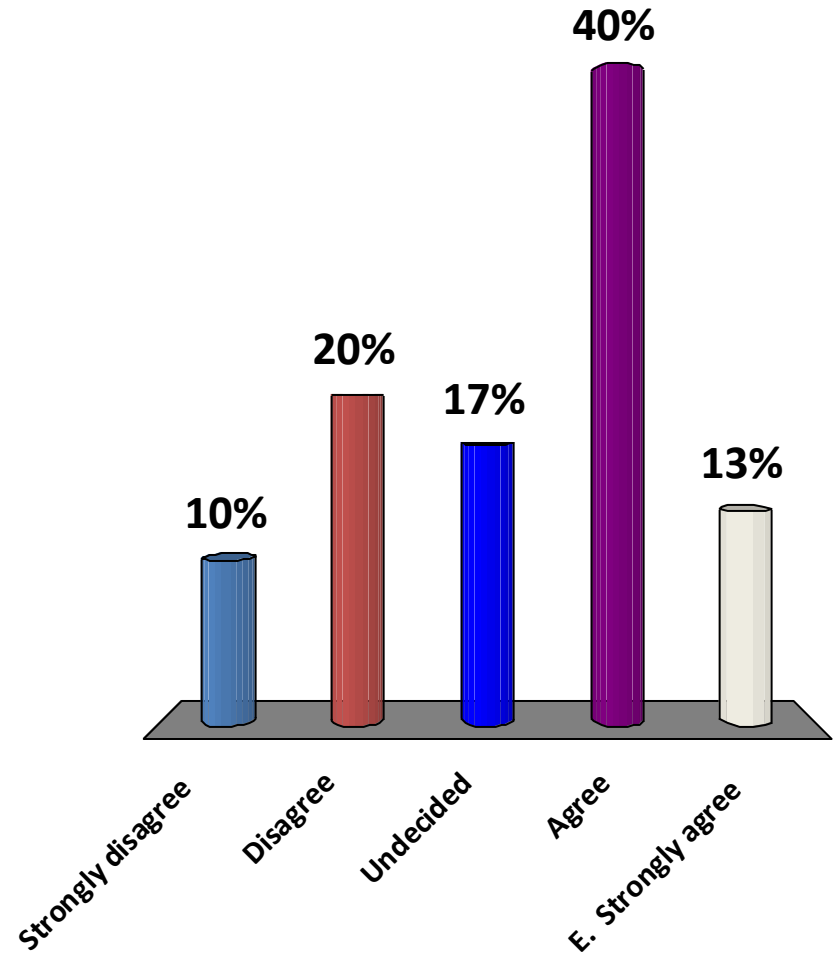
Agree .D

Strongly agree .E



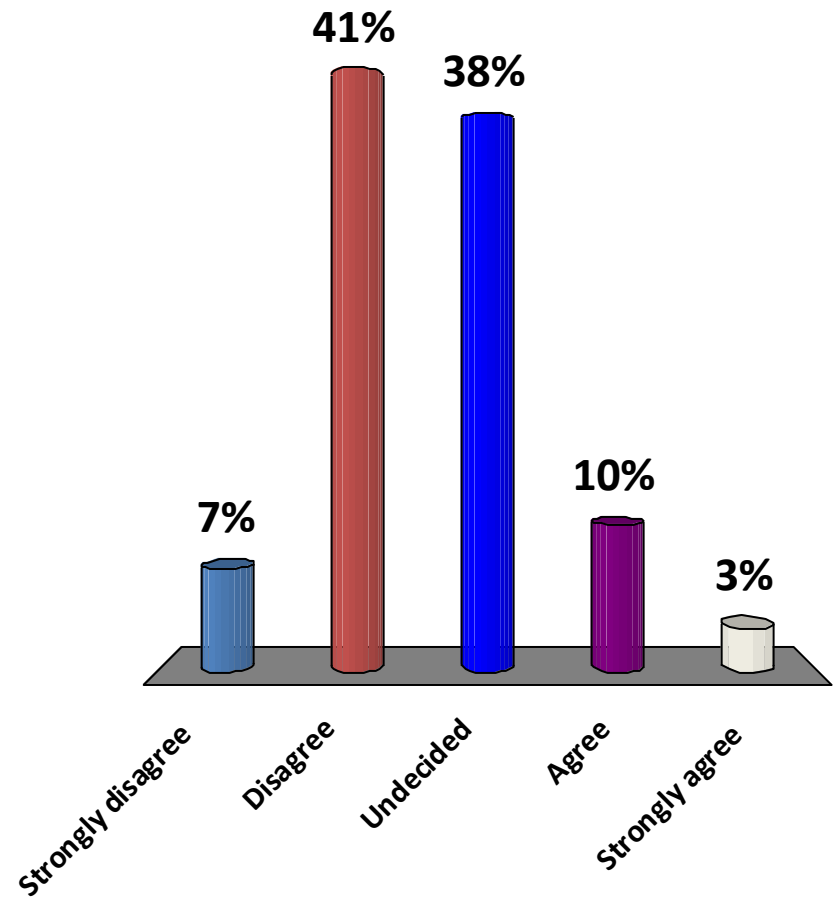
During the exam my sympathy is with the virtual patient

- A. Strongly disagree
- B. Disagree
- C. Undecided
- D. Agree
- E. E. Strongly agree



In general, the borderline candidate should pass?

- A. Strongly disagree
- B. Disagree
- C. Undecided
- D. Agree
- E. Strongly agree



DC COMICS

THE NEW 52!

#1
\$2.99

HAWK & DOVE



LAURENCE

GATES
SEIFELD

DIRECT SALES



000000 000000

000000 000000

Examiners

- Hawks and Doves
- Stringent and lenient
- Candidate centered and patient centered

Candidate centered

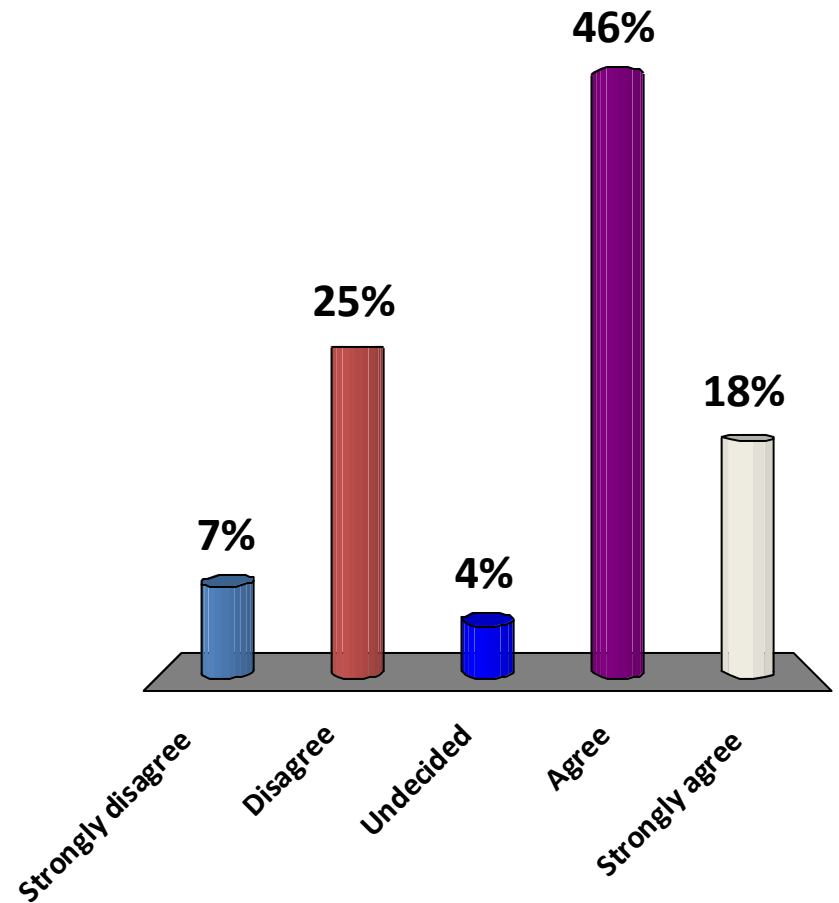
- Their sympathy are primarily with the candidates, of whom they wish to pass as many as possible

Patient centered

- Their primary aim is to maintain clinical standards at a high level so the patients are protected and provided with competent doctors

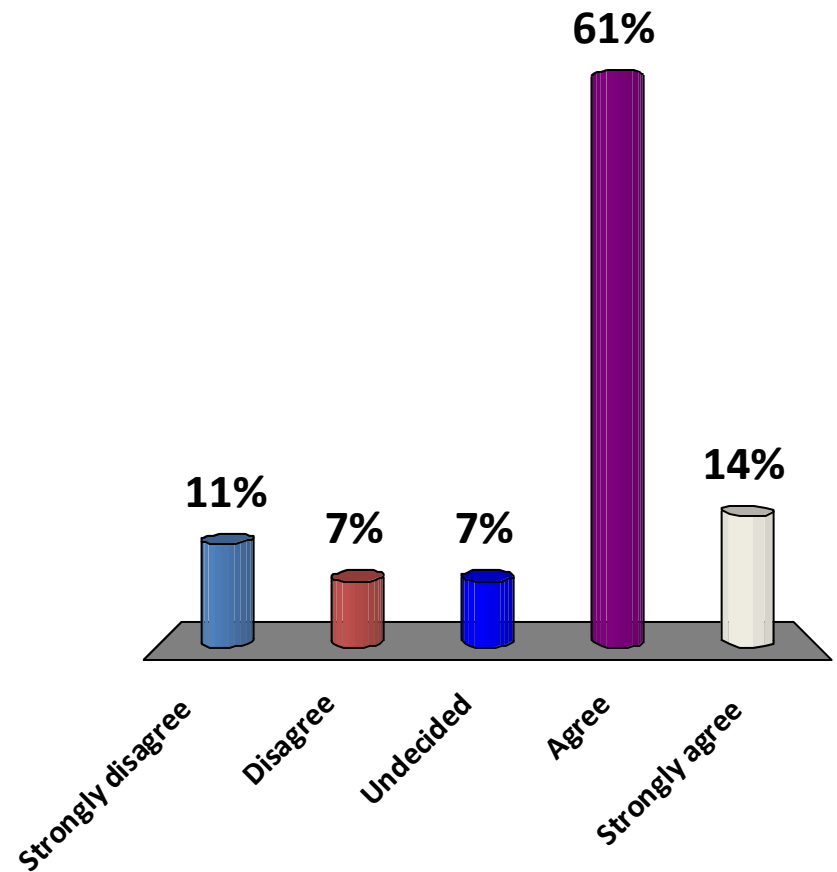
Good or back luck in exams exists

- A. Strongly disagree
- B. Disagree
- C. Undecided
- D. Agree
- E. Strongly agree



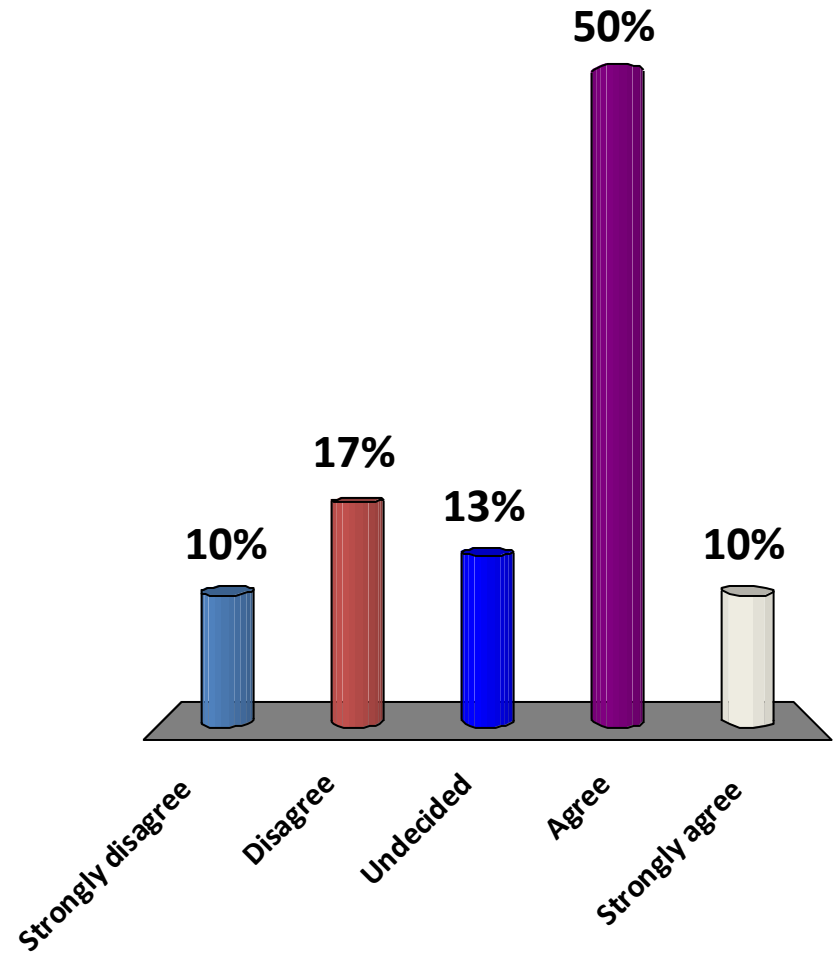
Results depend on the stringency of examiners

- A. Strongly disagree
- B. Disagree
- C. Undecided
- D. Agree
- E. Strongly agree



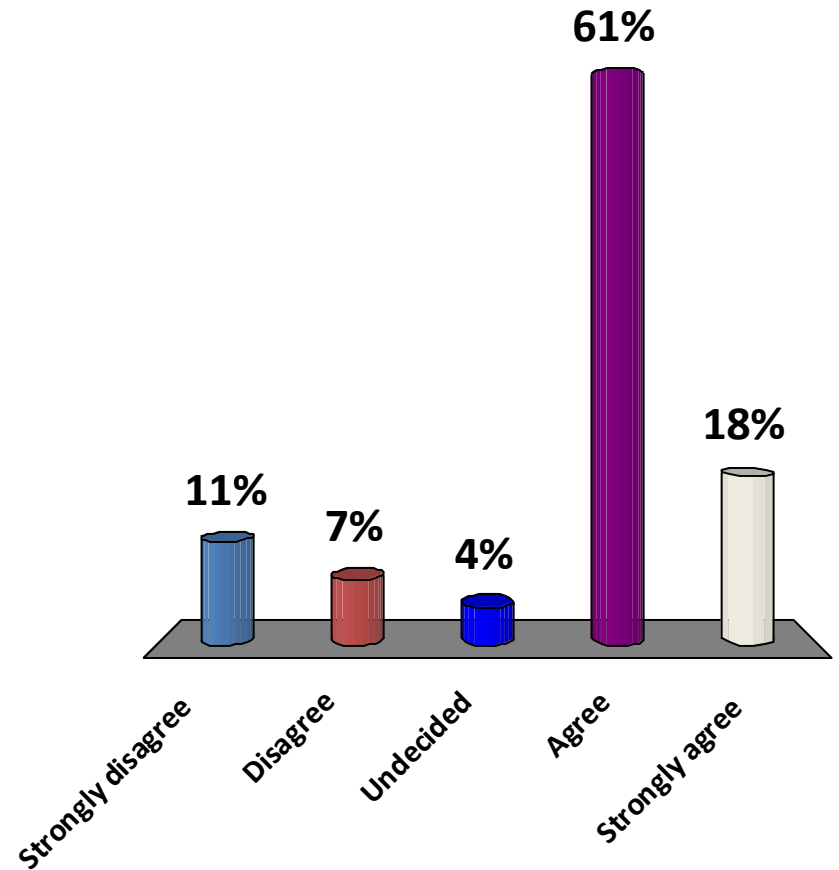
Results depend on the marking system

- A. Strongly disagree
- B. Disagree
- C. Undecided
- D. Agree
- E. Strongly agree



Results depend on test's difficulty

- A. Strongly disagree
- B. Disagree
- C. Undecided
- D. Agree
- E. Strongly agree



Candidate's performance

- Candidate ability
- Test difficulty
- Examiner stringency
- Marking scale

Case specificity

- Because candidates are not equally proficient at all clinical tasks they have areas of weakness and strength, and hence can get lucky or unlucky in the particular cases they happen to see, sometime seeing cases with which they are familiar and other times seeing cases with which, for a host of reasons, they are unfamiliar.

Marks

- 87% of systematic variance- difference in candidates
- 12% differences in examiners
- 1% Station type

Examiner's marks

- 8% Clear fail
- 27% Fail
- 40% Pass
- 25% Clear Pass

Examiners specificity

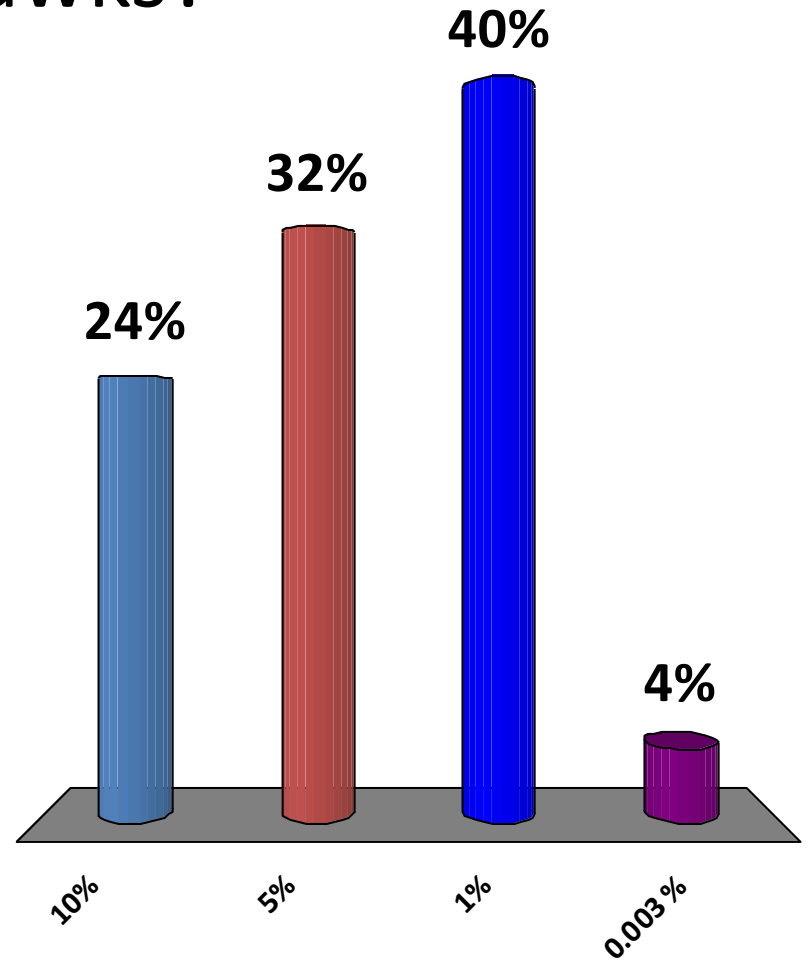
- A candidate's marks depending on the particular examiner (s) they happen to see

I see Professor Sanderson
is as hawkish as ever....



What is in your opinion the usual percentage of extreme doves examiners or extreme hawks?

- A. 10%
- B. 5%
- C. 1%
- D. 0.003 %



“Doves” and “Hawks”

Step One

- Identify potential extreme examiners:
 - All examiners whose average score for a station was more than three standard deviations above (**potential dove**) or below (**potential hawk**) the average score for all remaining examiners
 - (33) out of 2182
 - Catching the Hawks and Doves: A Method for Identifying Extreme Examiners on Objective Structured Clinical Examinations” – July 20th 2011

Doves” and “Hawks”

Step Two

- Identifying distribution of ratings from extreme examiners compared with distribution for all examiners
- Determine whether the examiner demonstrated adequate variability in their candidate ratings for a given station
- (17) (16 eliminated)

Doves” and “Hawks”

Step Three: Cohort criterion

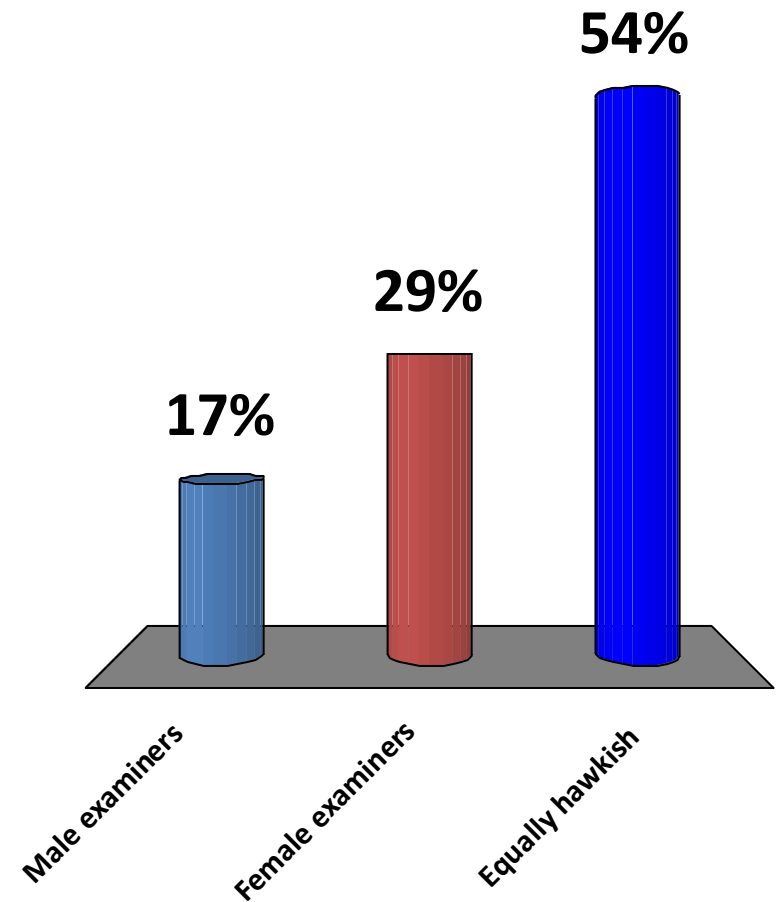
- Determine whether the candidate cohort seen by the examiners in question demonstrated adequate variability
- From 17 only 7 remained (7)
- 7 out of 2182

Effect of training

- “All assessments that depend on human raters are vulnerable to mischief due to raters
- Some studies have found null or even negative effects of training
- - Bernardin, HJ, & Buckley, MR. Strategies in rater training. *Academy of Management* 1981; 6: 205-222.

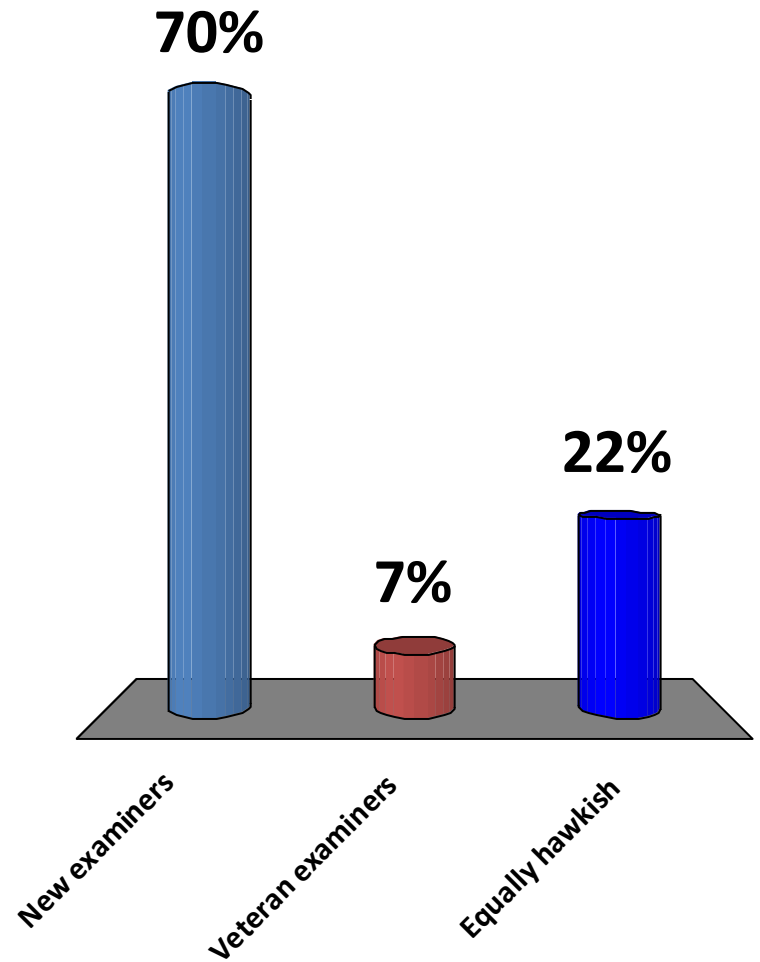
Males and females – who is more hawkish?

- A. Male examiners
- B. Female examiners
- C. Equally hawkish



Young or experienced – who is more hawkish?

- A. New examiners
- B. Veteran examiners
- C. Equally hawkish



Examiner's Gender Difference

- None

Young or experienced

- Experienced are slightly hawkish...

The language problem

- The slow candidate
- Crucial?
- Translate?
- Be telegraphic!

LIFE WAS
MUCH EASIER
WHEN Apple
AND *BLACKBERRY*
WERE JUST
FRUITS

Setting the standard

- Good mix of examiners!
- Good mix of questions
- Good mix of exam modalities
- Appropriate marking system

Thanks