



**Sharing reliable and valid OSCE stations so as to improve
cross-institutional assessment strategies:
Are we equipped for it?**

Thomas JB Kropmans

Senior lecturer MIME, CMN&HS



QPERCOM
RECRUIT



QPERCOM
OBSERVE



QPERCOM
CHOICE



QPERCOM
ENTRUST



Methods



QPERCOM
OBSERVE

- Signed Informed consent of 20 Universities
- 11 responded positively despite NDA
- 50% negotiated anonymity in data analysis
- QA (quantitative) penultimate year 16-17
 - Classical psychometrics
 - G-theory analysis
 - 2 way symmetric design Students X Observations (R)
 - Absolute G-coefficient plus SEM



Number of OSCE, Stations & Examiners



| Location | Penultimate year | Year of Analysis | Number of OSCE | Number of Stations (min-max) | Nr of Candidates (min - max) | Multiple scenarios/station | Nr of examiners/station (min-max) |
|-------------------|------------------|------------------|----------------|------------------------------|------------------------------|----------------------------|-----------------------------------|
| Central Europe 1 | 3 | 2016-2017 | 4 | 8 - 16 | 36 - 373 | 5 | 12 - 18 |
| Central Europe 2 | 5 | 2016 - 2017 | 2 | 14 | 205 | 4 - 7 | 4 - 8 |
| Northern Europe 1 | 4 | 2016-2017 | 5 | 4 - 16 | 8-28 | 1 | 1 - 2 |
| Northern Europe 2 | 4 | 2016-2017 | 7 | 8 | 126 | 1 | 4 |
| Western Europe 1 | 4 | 2016-2017 | 4 | 7 - 9 | 91 - 96 | 1 | 4 - 6 |
| Western Europe 2 | 4 | 2016-2017 | 2 | 6 - 7 | 340 | 1 | 6 - 8 |
| Western Europe 3 | 4 | 2016-2017 | 3 | 7-5 | 148 - | 1 | 3 - 3 |
| Western Europe 4 | 4 | 2016-2017 | 1 | 5 | 119 - | 1 | 5 |
| Western Europe 5 | 4 | 2016-2017 | 1 | 4 | 16 | 1 | 1 |
| Western Europe 6 | 3 | 2016-2017 | 1 | 3 (9 - 12) | 263 | 1 | 23 - 47 |
| Western Europe 7 | 4 | 2016 - 2017 | 1 | 6 | 60 | 1 | 1 - 3 |



Divergence

| Divergence (min-max) | Nr students scoring below cut score | Nr students failing on the GRS |
|----------------------|-------------------------------------|--------------------------------|
| 20.6 - 50.4% | 645 | 435 |
| 3.0 - 18% | 295 | 205 |
| 4.7 - 6.9% | 4 | 31 |
| 0.9%-8.8% | 200 | 8 |
| 8-14.9% | 114 | 16 |
| 12.5 - 26.8% | 0 | 20 |
| 3 - 19.8% | 68 | 2 |
| 3.8%-11.8% | 65 | 0 |
| 0 | 5 | 10 |
| 33.5 - 47.6% | 36 | 20 |
| 0 - 12.9% | 46 | N/A |



Advanced measures

| Cut-score | BRA (min-max) | OSCE MIS | Absolute G-coeff. | Cronbach's Alpha (min-max) | Standard Error of Measurement (Abs) | Error associated with stations |
|-----------|---------------|----------|-------------------|----------------------------|-------------------------------------|--------------------------------|
| 65% | 41.6 - 68.3% | 9 | 0.59 | 0.61 - 0.89 | 4.58% | 6.70% |
| 53% - 88% | 59.2 - 86.0% | 8 | 0.57 | 0.44 - 0.66 | 5.01% | 13.40% |
| 70% | 30.8 - 74.3% | 8 | 0.42 | 0.24 - 0.74 | 6.37% | 46.60% |
| 65%-70% | 44.1 - 77.5% | 1 | 0.44 | 0.25 - 0.75 | 5.90% | 28% |
| 50% | 45.8 - 53.9% | 9 | 0.77 | 0.69 - 0.84 | 3.20% | 21% |
| 0% | 55.2 - 63.1% | 7 | 0.58 | 0.75 - 0.88 | 5.37% | 3.80% |
| 53%-87% | 52.7 - 88.4% | 7 | 0.75 | 0.25 - 0.84 | 3.24% | 49.80% |
| 51%-61% | 51.4 - 62.1% | 0.5 | 0.45 | 0.55 - 0.73 | 6.12% | 31.80% |
| 50% | 0/0 | 0.5 | 0.36 | 0.51 - 0.88 | 7.78% | 19.69% |
| 44% - 65% | 53.7 - 70.6% | 4 | 0.3 | manually | 10.33% | 6.70% |
| 50% | N/A | 4 | 0.34 | 0.43 - 0.91 | 5.49% | 54% |



Standard Error of Measurement



| Location | Cut-score | Standard Error of Measurement (Abs) | Cut-score minus 1 SEM | Cut-score plus 1 SEM | Highest score | Lowest score | Range | Mean |
|-------------------|-----------|-------------------------------------|-----------------------|----------------------|---------------|--------------|-------|--------|
| Central Europe 1 | 65% | 4.58% | 60.42% | 69.58% | 0.86 | 0.47 | 39% | 70.70% |
| Central Europe 2 | 53% - 88% | 5.01% | | | 0.96 | 0.48 | 48% | 77.60% |
| Northern Europe 1 | 70% | 6.37% | 63.63% | 76.37% | 0.91 | 0.25 | 66% | 73.30% |
| Northern Europe 2 | 65%-70% | 5.90% | | | 0.92 | 0.59 | 33% | 80.90% |
| Western Europe 1 | 50% | 3.20% | 46.80% | 53.20% | 0.71 | 0.46 | 25% | 59.90% |
| Western Europe 2 | 0% | 5.37% | | | 0.95 | 0.45 | 49% | 74.50% |
| Western Europe 3 | 53%-87% | 3.24% | | | 0.97 | 0.75 | 22% | 88.00% |
| Western Europe 4 | 51%-61% | 6.12% | | | 0.89 | 0.54 | 35% | 71.60% |
| Western Europe 5 | 50% | 7.78% | 42.22% | 57.78% | 0.90 | 0.59 | 32% | 78.50% |
| Western Europe 6 | 44% - 65% | 10.33% | | | 1.00 | 0.44 | 56% | 79% |
| Western Europe 7 | 50% | 5.49% | 44.51% | 55.49% | 0.72 | 0.46 | 26% | 61.10% |



Conclusions

- Examiners are confused!
- Discrimination Global Rating Scores and professional clinical practice
- Wide divergence in marks and GRS

- Examiners training e.g. examiners feedback
- Multiple examiners (SP + Examiner)
- Divergence rules (like in Qpercom RECRUIT)

Sharing reliable and valid OSCE stations so
as to improve cross-institutional
assessment strategies:
Are we equipped for it?



- *Educational Informatics* provides insight in QA through Advanced Assessment Solutions (IT)
 - Fail & Borderline performance is unclear to examiners
 - Divergence between examiners is troublesome
- Sharing means caring, are we ready for it?



thomas.kropmans@qperc
om.ie



**OWN YOUR LEARNING
SELFISHLY**

**SHARE YOUR LEARNING
GENEROUSLY**

@ROBERT_SCHUETZ