scoring/feedback to enhance catalytic educational effects of our M4 OSCE.
Summary of Work: We provide OSCE scoring/feedback at the individual-station level, and the overall-exam level. For the 2012-13 cohort(n=171), we calculated new skill-domains to provide additional scoring/feedback: Data-Gathering, tabulated from History/Physical Examination checklists across seven stations; Communication, from six station checklists; and Clinical-Reasoning, from five post-encounter notes and assessments, an oral case presentation, and written evidence-based medicine and critical values exams.
Summary of Results: Across domains, Mean(SD) were similar. Data-Gathering:83%(5%); Communication:82%(8%); Clinical-Reasoning:83%(5%)). All inter-domain correlations were significant($p<0.01$). Communication yielded weak inter-domain correlations with Clinical-Reasoning($r=0.25$) and Data-Gathering ($r=0.22$); Data-Gathering and Clinical-Reasoning were moderately correlated($r=0.45$); OSCE-Overall had the strongest correlation with Clinical-Reasoning($r=0.87$). Students ≥2SD below average on Communication scored variously on the OSCE-Overall: some poorly, others well. In contrast, students with low Clinical-Reasoning scores fell in the bottom quartile on OSCE-Overall.
Conclusions: Adding domain-specific scoring to our OSCE identified needs for communication and clinical-reasoning curricular enhancements. It also improved remediation for students with domain-related performance patterns. Some who did well on the OSCE overall learned Communication improvements were still warranted; patterns for students with poor Clinical-Reasoning scores revealed some needing data-gathering improvements, while others could data-gather but needed to improve synthesis.
Take-home Messages: Domain-specific scoring/feedback expanded catalytic educational effects of our OSCE, by identifying areas for curricular enhancement, and improving feedback and remediation processes for students.

OTT-OB-6-3
Benchmarking the attainment of clinical competence

Bunmi Malau-Aduli (*James Cook University, Medical Education Unit, School of Medicine and Dentistry, Townsville, Australia)*, Peta-Ann Teague (*James Cook University, School of Medicine and Dentistry, Townsville, Queensland*), Richard Turner (*University of Tasmania, School of Medicine, Hobart, Tasmania*), Karen D'Souza (Deakin University, *School of Medicine, Melbourne*), David Garne (*University of Wollongong, Graduate School of Medicine, Wollongong*), Ian Wilson (*University of Wollongong, Graduate School of Medicine, Wollongong*)

Background: The sharing of co-developed Objective Structured Clinical Examinations (OSCEs) by medical schools for the assessment of clinical competence could serve as a responsive and flexible means of entrenching assurance of high quality assessment processes. It has the added benefit of benchmarking and evaluating comparable competence standards.
Summary of Work: This study was undertaken to explore the feasibility of utilising shared OSCEs to benchmark clinical performance of students in four geographically dispersed Australian medical schools. Four shared OSCE stations were co-developed by the participating medical schools and embedded in the end of year examinations for the assessment of clinical performance in the early clinical phase of the course. Returned checklist, global and total scores from 1670 student results were then analysed using SAS analytical package to compare mean scores and clinical competence levels.
Summary of Results: Data analysis revealed similar clinical competence patterns in the performance of the medical students, indicating comparable standards. The degree of difficulty for the shared OSCE stations was largely similar for participating schools, although mean total student scores varied between schools.
Conclusions: Benefits of benchmarking are available to collaborating medical schools through identification of common curriculum areas requiring specific focus and the sharing of assessment approaches. Similarly, relative underperformance by a school in a particular assessment item may indicate deficiency that can be remediated in order to achieve comparability with its peers.
Take-home Messages: Sharing of assessment materials can provide common defensible, reliable, valid, robust and standardised assessments which in turn, enhance transparency and accountability.

OTT-OB-6-4
The impact of OSCE case information length on examiner markup behavior

Marguerite Roy (*Medical Council of Canada, Ottawa, Canada)*, Caroline Marshall (*Medical Council of Canada, Ottawa*)

Background: Being an OSCE examiner is demanding. Examiners must observe candidate performance in real time, track targeted case information for scoring decisions, complete rating tools correctly, monitor SP performance, and be attuned to the assessment standard. The risk of cognitive overload is high. Cognitive load theory (Sweller, 1983) builds on the premise that working memory has limited capacity and posits that when tasks, materials, or situations impose excessive load on an already limited system, task performance may suffer (Taveres & Eva, 2013). Excessive task-irrelevant load can be effectively offset by redesigning materials or developing strategies to better support task-relevant demands (Chandler & Sweller, 1992; Moreno & Mayer, 1999).
The current study explored the relationship between the complexity of case information sheets provided to aid examiners in using rating scales to score and spontaneous strategies they adopted as they rated candidates.
Summary of Work: Case information sheets for a high stakes OSCE involving 232 examiners rating 548 candidate performances across multiple sites were collected. The length of case information per station was