

Adapting the McMaster-Ottawa Scale and Developing Behavioral Anchors for Assessing Individual and Team Performance in a Team Observed Structured Clinical Encounter (TOSCE)

Keck School of Medicine of USC

Désirée Lie, MD, MSED, Win May, MD, PhD, Regina Richter Lagha, PhD, Christopher P. Forest, PA-C, Yvonne Banzali, MBA, MFT, Kevin Lohenry, PhD, PA-C
Keck School of Medicine of USC, Division of Physician Assistant Studies



INTRODUCTION

BACKGROUND:

Rating scales for interprofessional team performance do not provide adequate behavioral anchors for allowing formative feedback to teams and individuals. The Team Observed Structured Clinical Encounter (TOSCE) provides an opportunity to adapt and develop an existing scale for this purpose.

STUDY AIMS:

- Retool the McMaster-Ottawa Scale for evaluating individual student and team performance, and test feasibility in a TOSCE.
- Evaluate the ability of faculty raters to use the retooled scale to accurately distinguish different levels of student and team performance.

METHODS

- McMaster-Ottawa Scale (six constructs) was retooled to add behavioral anchors and converted from 9-point to 3-point scale (Table 1).
- 16 Students (four professions per team: medicine, physician assistant, pharmacy and occupational therapy) were trained to perform at three levels (below, at, and above expected), as individuals and teams using behavioral anchors.
- 16 Blinded faculty raters from five professions were trained using prepared videos of different performance levels to evaluate student and team performances using the retooled scale.

COMPETENCIES	INDIVIDUAL RATING		
	Below	At	Above
1. Communication	1	2	3
2. Collaboration	1	2	3
3. Roles & Responsibilities	1	2	3
4. Collaborative Patient-Family Centered Approach	1	2	3
5. Conflict Manage Resolution	1	2	3
6. Team Functioning	1	2	3
GLOBAL RATING SCORE	1	2	3

Table 1. Modified McMaster-Ottawa Scale. Detailed instructions were provided for 3-point scoring in each category.

RESULTS

DATA ANALYSIS:

- Descriptive statistics as well as generalizability theory (G theory), using SPSS and GENOVA, were used to examine the ability of faculty to accurately rate individuals and teams.

FEASIBILITY (Aim 1)

- One TOSCE station (stroke) was successfully administered to four teams (16 students total) over four hours, with four faculty rating each student and team.
- Trained students performed at pre-assigned competency levels.
- Faculty were comfortable rating up to four students per station during the course of a 35-minute encounter.



Figure 1. Faculty trainers and faculty raters

Source of Variance	df*	1 Faculty Rater 6 Competencies†	2 Faculty Raters 6 Competencies†	4 Faculty Raters 6 Competencies†
Faculty (r)	10	0.04234 (9.93)	0.02117 (6.31)	0.01058 (3.73)
Student (p)	15	0.23441 (55.00)	0.23441 (69.86)	0.23441 (82.75)
Competency (c)	5	0.00042 (0.10)	0.00042 (0.13)	0.00042 (0.15)
pr	150	0.12243 (28.72)	0.06122 (18.25)	0.03061 (10.81)

Table 2. Estimated variance components for student performance scores on TOSCE
* df indicates degrees of freedom
† Variance component (% of total variance)

RATER ACCURACY (Aim 2)

- Accuracy of raters varied for students (38% to 81%), with the majority of rating errors in the direction of over-rating student performance.
- Accuracy of raters for team performance (50% to 100%) was greater than for individual students.
- G study: students x faculty raters x competency indicated variation attributable to interaction of student and rater (pr), i.e. relative standing of students varied by rater (0.03061, or 11% of total variance) (Table 2).
- G study: faculty raters x students nested within teams (s:t) indicated some variation (0.00650, or 25% of total variance) in rater ability to accurately score student performance (Table 3).
- High proportion of error variance due to interaction of rater and team (0.00487, or 18% of total variance), i.e. relative standing of teams varied from rater to rater.



Figure 2. Students clarifying their roles before seeing the patient

Source of Variance	df*	1 Team, 1 Student/ Team†	4 Teams, 4 Students/ Team†
Faculty (p)	14	0.00650	0.00650 (24.81)
Team (t)	3	0.00620	0.00155 (5.92)
Student (s): Team (t)	12	0.07123	0.00445 (16.98)
pt	42	0.01949	0.00487 (18.59)
ps,t,e	168	0.14127	0.00883 (33.70)

Table 3. Estimated variance components for faculty ability to correctly identify student performance level on TOSCE

TOSCE



Figure 3. Student team conducting H&P on SP



Figure 4. Student team huddling in a separate room

CONCLUSIONS

- The TOSCE can be administered as an evaluation of individual student and team competencies in an interprofessional setting.
- Faculty raters demonstrate a 'leniency error' rating students, even with prior training.
- We recommend at least two faculty per station for accuracy of rating.
- We also recommend either Rater Error Training or Frame-of-Reference Training with an emphasis on an increase in the number of observations for lower-performing students.
- Future studies will examine psychometric properties of the retooled scale.

REFERENCES

- Olson R, Bialocerkowski A. *Med Educ.* 2014;48:236-246.
- McMaster-Ottawa TOSCE Toolkit. Available at <http://fhs.mcmaster.ca/tosce/en>
- Solomon P, Marshall D, Boyle A, et al. *J Interprof Care.* 2011;25(4):302-304.
- Richter RA, et al. *Acad Med.* 2012;87:1077-1082.
- Feldman M, et al. *J Contin Educ Health Prof.* 2012;32:279-286.

ACKNOWLEDGEMENT:

This study was supported by the Health Resources and Services Administration Bureau of Health Professions grant, Physician Assistant Training in Primary Care.

Contact: Desiree Lie - dlie@usc.edu