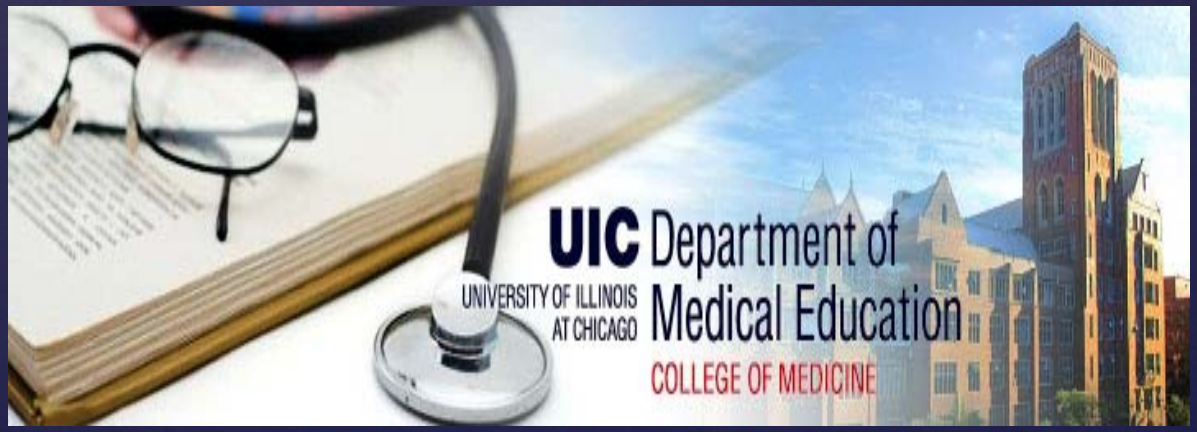
Assessing the Professionalism of Pediatric Residency Applicants: Validity and Feasibility Evidence for the Professionalism Mini- Evaluation Exercise

N. Bajwa, MD, N.V. Vu PhD, R. Yudkowsky, MD, MHPE, Y.S. Park, PhD

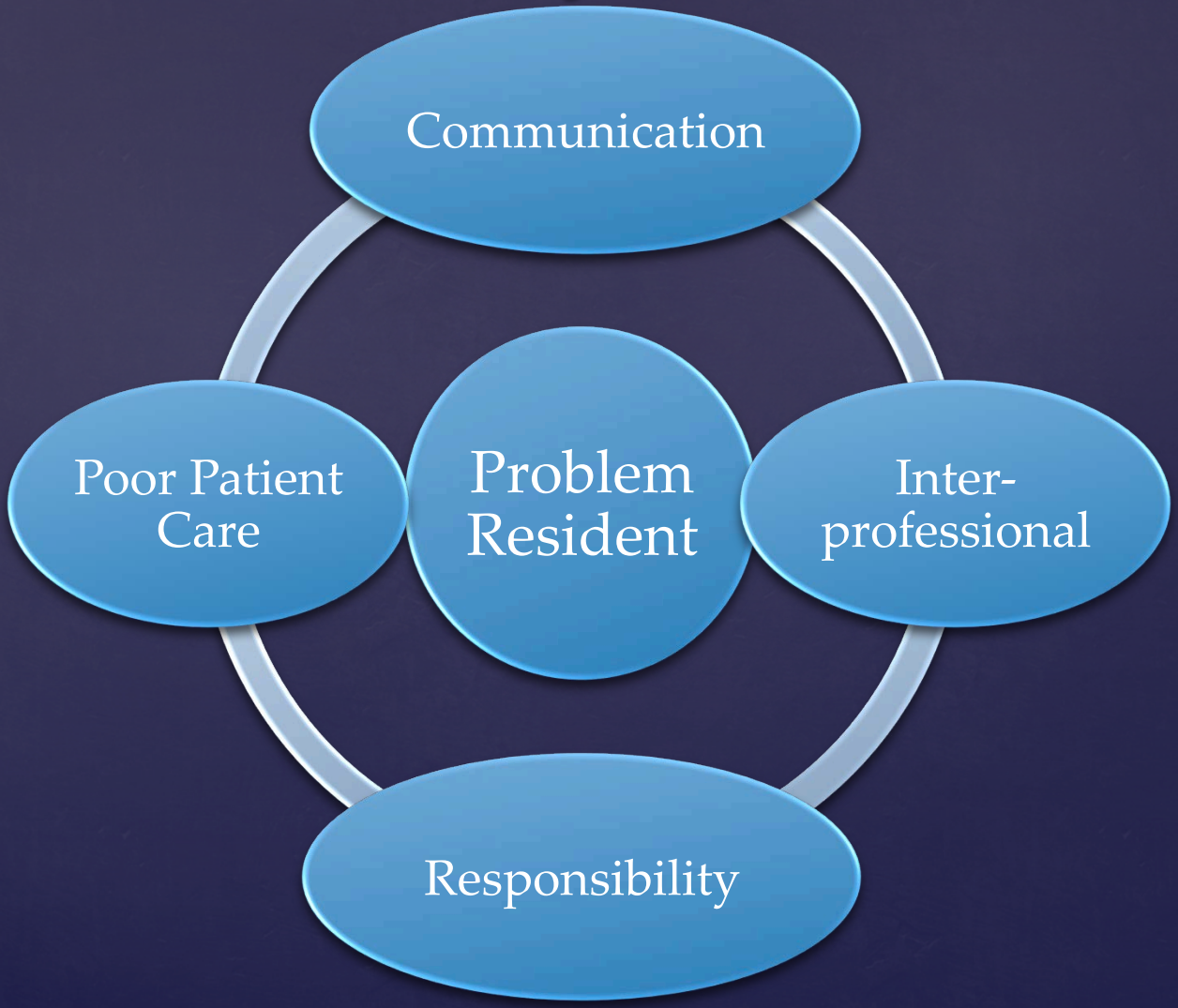
UNIVERSITÉ DE GENÈVE
FACULTÉ DE MÉDECINE
Département de pédiatrie



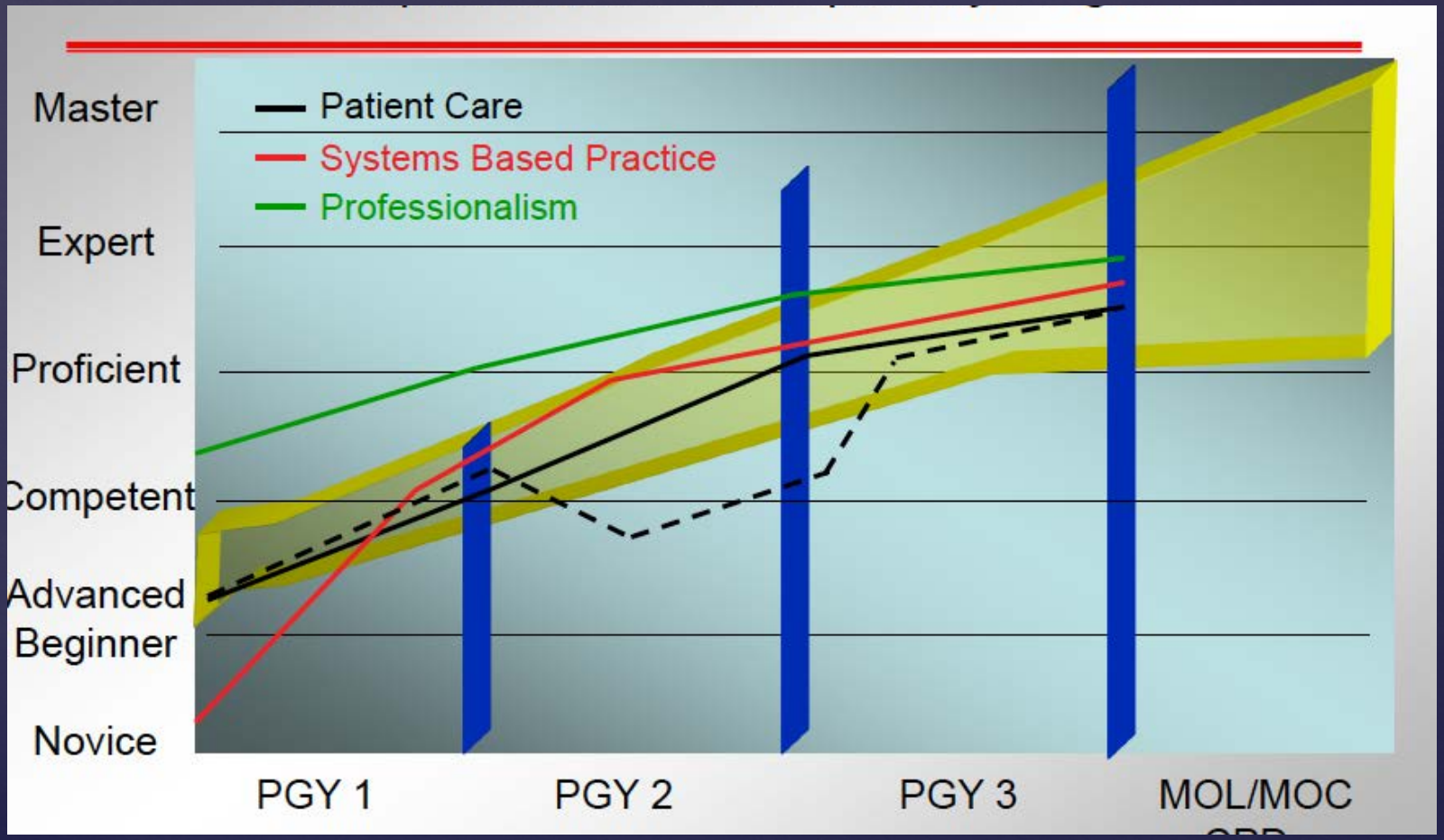
Outline

- Introduction and background of the study problem
- Conceptual frameworks
- Purpose of the study
- Methods
- Results
- Analysis and limitations
- Conclusions
- Future Directions

The Study Problem

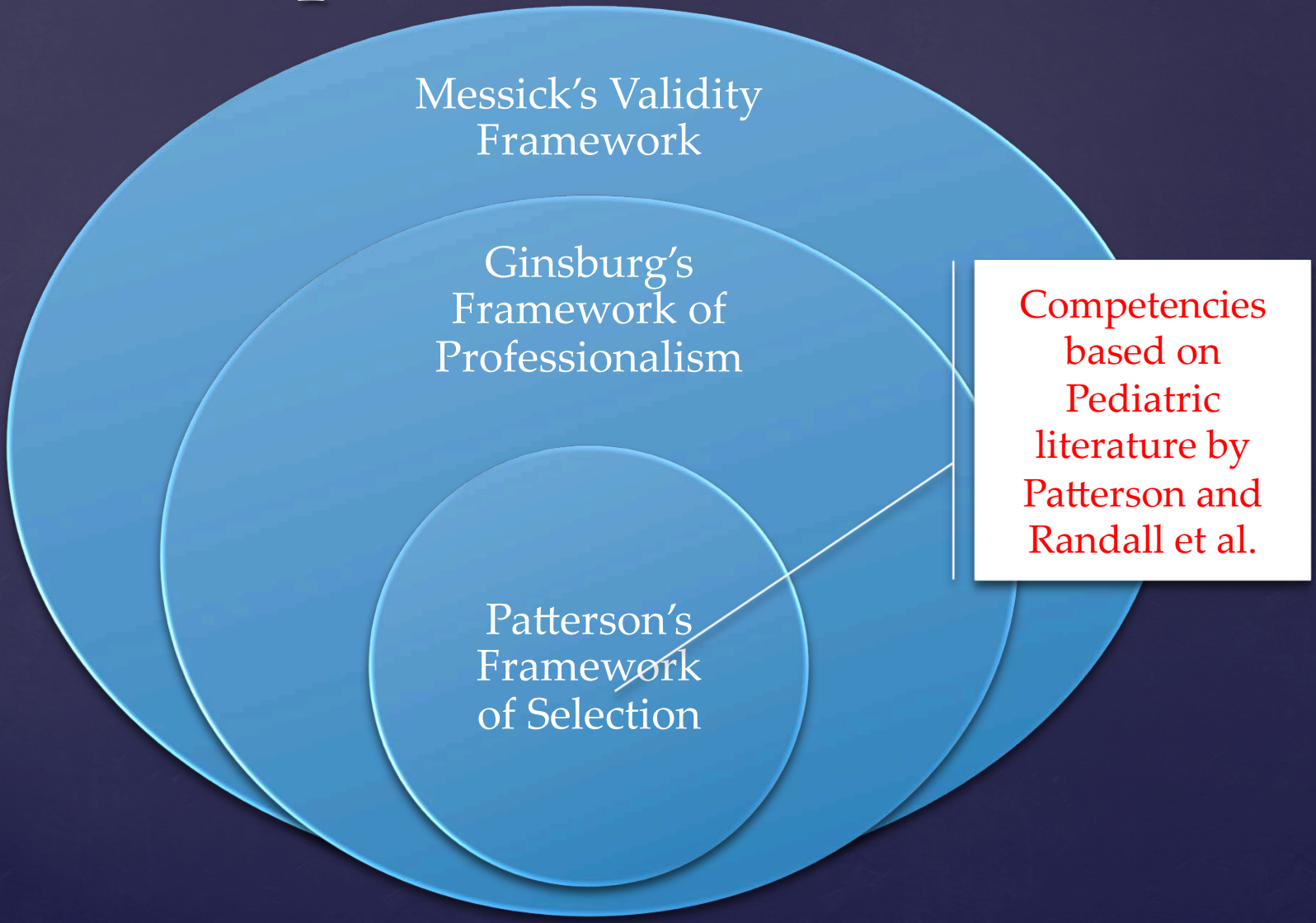


The Study Problem

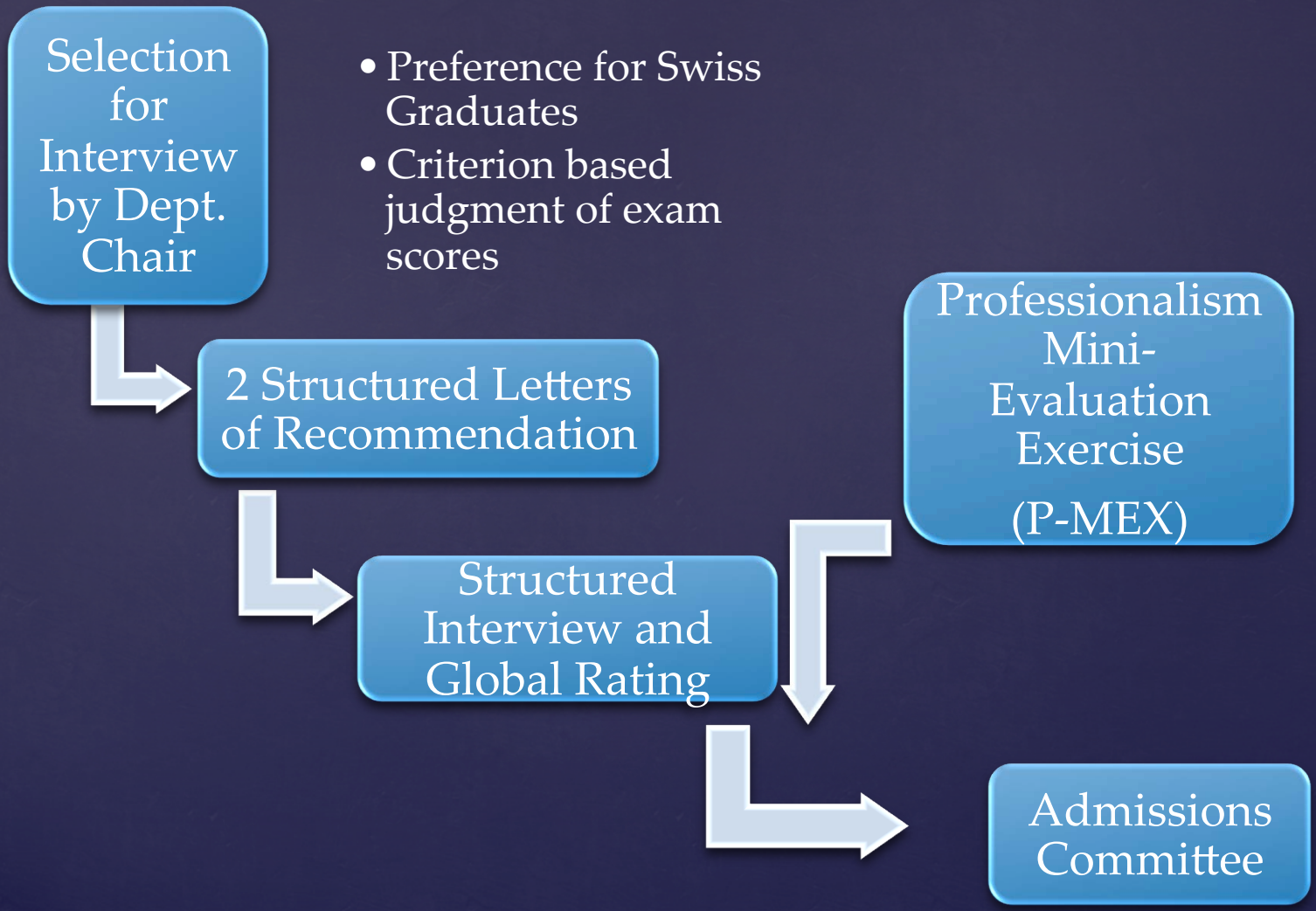


Nasca T. Graduate Medical Education in the United States. Vision and General Directions for the Next Ten Years.; November 7, 2010; Washington D.C.

Conceptual Frameworks



Pediatric Residency Admissions Process



Purpose of the Study

- *To gather validity and feasibility evidence for the use of the P-MEX in assessing professionalism in a pediatric residency admissions process.*

Sample

- All applicants who were selected for an interview were approached to participate in the study.
- 73 applicants out of 225 who applied for an interview, 32%.

Professionalism Mini-Evaluation Exercise (P-MEX)

- Based off of the Mini-Clinical Evaluation Exercise (Mini-CEX)
- Cruess et al., from McGill
- Direct observation during a patient consultation
- 21 Item evaluation form based on:
  - Doctor-Patient Relationship
  - Reflective skills
  - Time Management skills
  - Interprofessional Relationships
- Scale: Unacceptable, below expectations, meets expectations, and above expectations

P-MEX

- Construct validity in assessing the professionalism behaviors of medical students and residents (Canada and Japan)
- Limitations:
  - Need for multiple assessments (8-12)
  - Low reliability if junior residents are used as raters

Until now the P-MEX has never been used in residency admissions.

P-MEX Standardized Patient Cases

- University of Geneva Faculty of Medicine Standardized Patient Program
- 2 cases in 2012, 3 cases in 2013
- Rater training

alternativechoicespa.blogspot.com, www.womenshealth.gov, Image from NVIC

Structured Letter of Recommendation (SLR)

- 2 letters per applicant
- 6 competencies:
  - Professional integrity
    - 2013, 4 yes/no questions regarding lack of collegiality, poor documentation, excessive absences, irregular presence at teaching conferences
  - Scientific curiosity
  - Organizational skills and autonomy
  - Teamwork and collaboration
  - Patient management skills
  - Communication skills
- Overall assessment: lower tier, middle tier, higher tier, and top 10%

Structured Interview (SI)

- 1 interview per applicant with 2 faculty interviewers
- 3 structured questions based on desired competencies
- Dreyfus scale with behavioral based definitions used to rate responses
- Interviewers were trained in a training workshop and given an interview guide

Global Rating (Global)

- Based on SI, CV, and overall impression
  - Blind to SLR and exam scores
- Same 2 faculty interviewers independently rated the applicant
- 5 point scale (A+, A, B+, B, C)

2013 Composite Score

- Based on Kane's formula for composite reliability
- 2012 data (OR to predict admission) used to assign weights to the SLR, SI, and P-MEX
  - SLR: 18%
  - SI: 25%
  - P-MEX: 57%
- In 2013, the composite score of each applicant was made available to the admissions committee

Validity Methods

- Content
  - Item Analysis
  - Qualitative Feedback from applicants and faculty examiners: thematic analysis, using grounded theory, constant comparative method
- Response process
  - Inter-rater reliability with intraclass correlation
- Internal structure
  - Generalizability theory analysis
  - $p \times c \times r \times i$ design
  - Variance components from 7 different blocks of data were aggregated using Brennan's method.*

Validity Methods

- Relationship to other variables
  - Correlations between selection instruments
- Consequences
  - Logistic regression to predict admission
  - Pseudo R-squared was calculated to examine the contribution of predictors in the admissions process with and without P-MEX scores.
  - Wilcoxon signed-rank test comparing rank list with composite score vs. global rating

Feasibility

- Estimated based on resources used:
  - Faculty effort measured in hours for rater training and assessment of applicants
  - Cost of standardized patients

- Comparison between the admissions process in:
  - 2012: SLR, SI, and Global
  - 2013: SLR, SI, Global and P-MEX

Results
Sample and Demographics

- 70/73 applicants participated
  - 31 applicants in 2012
  - 39 applicants in 2013
- 87% women
- 29% foreign graduates
- 56% in their final year of medical school
- 74% had prior experience with SPs.

Results: Content

- Mean P-MEX score: 3.13/4 (SD=0.23)
- Average of 2.8/21 items (SD=1.70) not completed or marked not applicable

Items not completed	**Items noted "below expectations"**
Admitted error/omissions	Recognized and met patient needs
Was available to colleagues	Extended his/herself to meet patient or parent needs
Maintained patient confidentiality	Ensured continuity of patient care
	Completed tasks in a reliable fashion

Results: Content

- Both applicants and faculty felt that the P-MEX was an appropriate way to assess professionalism.
- Themes from faculty examiners included the desire to rate non-verbal communication skills, body language.

It's true that the cases were really well chosen and were representative of what we frequently see in pediatrics and the manner in which people react …I think it is a fairer evaluation…it's good to stress us a little and see how we work in that situation.

Results: Response Process

Selection Instrument	Intraclass correlation (standard error)
SLR	0.51 (SE=0.04)
SI	0.66 (SE=0.07)
Global	0.87 (SE=0.12)
P-MEX	0.36 (SE=0.02)

68% Exact agreement

Results: Internal Structure

- G-coefficient: 0.45, Phi coefficient 0.43

Table 1: Generalizability study results from the 2012 and 2013 combined P-MEX data (n=70)

Effect	Variance Component	%Variance Component
p	0.017	6.8%
c	0.001	0.4%
r	0.001	0.3%
i	0.005	2.1%
p x *c*	0.025	10.0%
p x *r*	0.005	2.0%
p x *i*	0.012	4.7%
c x *r*	0.000	0.1%
c x *i*	0.002	0.7%
r x *i*	0.003	1.0%
p x *c* x *r*	0.025	10.1%
p x *c* x *i*	0.020	8.0%
p x *r* x *i*	0.001	0.6%
c x *r* x *i*	0.002	0.9%
p x *c* x *r* x *i*	0.128	52.2%

Results: Internal Structure

- Decision study results, 2 raters per case
  - 7 cases, G-coefficient of 0.6
  - 10 cases, G-coefficient of 0.65

Number of cases	2013 data Phi-Coefficient	G-Coefficient
4	0.51	0.52
5	0.54	0.55
6	0.57	0.58
7	0.59	0.60
8	0.60	0.62
9	0.62	0.63
10	0.63	0.65

Results: Relationship to Other Variables

Correlations with P-MEX scores	r (p-value)
SLR scores	0.25 (p= .036)
SI scores	0.34 (p= .004)
Global	0.48 (p< .001)
Acceptance decision	0.56 (p< .001)

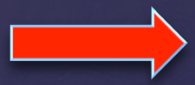
Pediatric Residency Admissions Process

Results: Consequences

Table 2. Odds ratios of scores from the selection instruments in relation to the acceptance decision (n=70)

	Accept/Not accept		
	Odds Ratio	Standard Error	*P*-value
Structured Interview	1.92	.76	.10
Structured Letter of Recommendation	1.22	.42	.56
P-MEX 2012 OR 3.4, p= .028	4.37	1.80	< .001
Pseudo R-squared		.31	
Pseudo R-squared (without P-MEX)		.11	

Results: Consequences

Composite Score Elements	Composite Reliability
SLR + SI	0.51
SLR+ SI+ 2 P-MEX Cases	0.66
SLR + SI+ 3 P-MEX Cases	 0.74

Results: Consequences

- Mean composite score was 0.77 (SD=0.07).
- Wilcoxon signed-rank test to compare rank lists created with the global rating (old method) vs. composite score (new method)
  - $z = 5.51$, $p < .001$

Results: Feasibility

2012 SLR, SI, Global

2 faculty hours per applicant.
No direct costs

+

P-MEX, 3 SP cases 30 CHF/hour,
2 faculty raters per case
1.5 hours per applicant
66 USD/ Applicant

2013 SLR, SI, Global, P-MEX, Composite

=

3.5 faculty hours per applicant
66 USD/ applicant

Analysis and Limitations

- Single specialty, single European institution.
- Evidence of construct irrelevant variance:
  - Low ICC (0.36)
  - Interaction of *p x c x r* (10.1%).
  - Need for improved rater training
- Evidence of case-specificity: *p x c* (10%)
- Evidence of construct under-representation (G-coefficient 0.45)
- Difficult in terms of feasibility to increase the number of cases (faculty hours).

Conclusions

- Novel approach to assessing professionalism in the residency admissions process
- Adds to the validity evidence for the P-MEX in the literature.
- The combination of the SLR, SI, and P-MEX improves the reliability of the admissions process (0.74).

Conclusions

- Both applicants and faculty found that the use of the P-MEX enhanced the fairness of the process.
- Admissions committee also found the process acceptable and feasible
- Important to assess professionalism!

Future Directions

- Determine the predictive validity of the P-MEX
- Detailed qualitative and content analysis
- Standard setting exercise for the P-MEX
- Improved rater training for the P-MEX

Merci!

- Special thanks to:
- Yoon Soo Park, PhD (committee chair)
- Rachel Yudkowsky, MD, MHPE (committee member)
- Prof. Nu Viet Vu (committee member)
- DME faculty
- Janet Settle, MA, MEd