Not just for learning:

Simulation for assessment of competency

Leah Mallory, MD

Micheline Chipman RN

Faculty Development Conference 2023

Disclosure

- No disclosures or conflicts of interest

Objectives

- Explore key differences between formative and summative assessment
- Describe the process needed to establish sufficient validity to use simulation for summative assessment
- Discuss examples to consider simulation as a means to implement just-in-time assessment

Competency based training vs apprenticeship

- Old model- Apprenticeship
  - Six week clerkships, clinical rotations for nursing students, years of residency
- New model- Competency-based
  - ACGME- "Next Accreditation System"/ "Milestones"
  - AACN- "The Essentials"/ Domains and Concepts
  - APPs- 6 core competencies
- **New model assumes robust assessment methodologies**

➢ **Simulation is structured, reproducible, realistic- an ideal venue to assess**

Formative Assessment: assessment *for* learning

- Deliberate Practice- Identify areas for improvement, create opportunities to practice, reflect, try again
- Key Principles
  - Longitudinal during training program
  - Discover knowledge gaps
  - Improve performance and identify further learning and development
  - Develop the professional role
  - Occurs often
  - Lower stakes
    - » Feedback personalized to individual learner

Simulation debrief: formative assessment

- Curiosity about why the learner performed the way they performed
- Discover their *frame* of thinking
- Are there knowledge gaps that created those frames?
- Shift frames for future use

Summative Assessment: assessment *of* learning

- Key Principles
  - Associated with a grade or official rating
  - Assessment of competence
  - Occurs less frequently
  - High stakes
    - » Move to the next level or certify
  - **Requires valid and reliable evaluation tools**
  - Evaluators must be trained on the tools

Designing a summative assessment

- Define the objectives, knowledge, and skills to be assessed
- Design appropriate simulations (task trainer, scenario) or educational experience
- Select or develop assessment tools
  - Ensure **validity**
  - Determine **reliability**
- Evaluators
  - Training on the tool is a must
  - Avoid knowing the student if possible
  - Independent rating occurrences

Oermann et al. Journal of Nursing Education, 2016, 55(6):323-328

Validity

- Validity is relative- there is no such thing as "validated"- a tool has validity evidence associated with it's use in a specific context
- At a bare minimum, for summative assessment you must prove your tool:
  - Measures skills, knowledge and attitudes that they were intended to measure
  - Can accurate interpretations about competence be made?

Reliability

- Reliability –refers to the consistency, stability and dependability
- At a bare minimum, for summative assessment you must prove your tool:
  - Reproduces the same results by different evaluators
  - Reproduces the same results at a different time by the same evaluator
  - Training is a must for evaluators

A Case Study- What to assess using sim?

Targeting Simulation-Based Assessment for the Pediatric Milestones: A Survey of Simulation Experts and Program Directors

Leah A. Mallory, MD; Sharon Calaman, MD; Marjorie Lee White, MD, MPPM, MA; Cara Doughty, MD, MEd; Karen Mangold, MD, MEd; Joseph Lopreiato, MD, MPH; Marc Auerbach, MD, MSci; Todd P. Chang, MD, MAcM

From the Division of Pediatric Hospital Medicine, Department of Pediatrics, Tufts Unive
Children's Hospital at Maine Medical Center, Portland, Maine (Dr Mallory); Division of P
Drexel University College of Medicine/St Christopher's Hospital for Children, Philadelphia
Department of Pediatrics, University of Alabama at Birmingham, Birmingham, Ala (Dr
Department of Pediatrics, Baylor College of Medicine/Texas Children's Hospital, Houst
Medicine, Department of Pediatrics, Ann & Robert H. Lurie Children's Hospital of Chica
University, Chicago, Ill (Dr Mangold); Department of Pediatrics, Uniformed Services Un
Lopreiato); Section of Pediatric Emergency Medicine, Department of Pediatrics, Yale U
Auerbach); and Division of Emergency Medicine and Transport, Department of Pediatri
Medicine at University of Southern California, Los Angeles, Calif (Dr Chang)
The authors declare that they have no conflict of interest.
Address correspondence to Leah A. Mallory, MD, Division of Pediatric Hospital Medicin
Children's Hospital at Maine Medical Center, Portland, ME 04102 (e-mail: mallol@mmc
Received for publication June 9, 2015; and in revised form September 22, 2015. accep

ABSTRACT

OBJECTIVE: To determine which of the 21 general pediatrics milestone subcompetencies are most difficult to assess using traditional methodologies and which are best suited to simulation-based assessment.

METHODS: We surveyed 2 samples: pediatric simulation experts and pediatric program directors. Respondents were

"demonstrate
strengths and d
to assessment
work," "clinica
tion." Program
agreement for

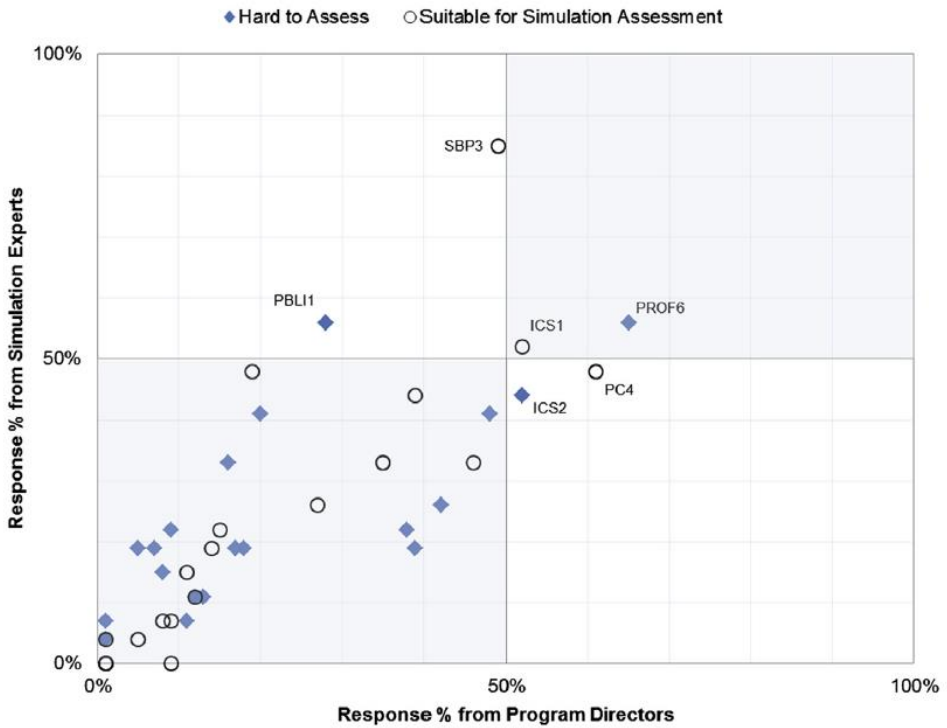

Figure. *Survey responses by expert group.*

3 Milestones- Examples of CBA

- IP Teamwork- *Works in an Interprofessional Team to enhance patient safety and improve patient care quality.*
- Empathy- *Demonstrate the insight and understanding into emotion and human response to emotion.*
- Ambiguity- *Recognize that ambiguity is part of clinical medicine and utilize appropriate resources in dealing with uncertainty.*

Interprofessional Teammember

ORIGINAL RESEARCH

Development of a Simulation-Based Interprofessional Teamwork Assessment Tool

Zia Bismilla, MD, MEd
Tehnaz Boyle, MD, PhD
Karen Mangold, MD, MEd
Wendy Van Ittersum, MD
Marjorie Lee White, MD, MPPM, MA
Pavan Zaveri, MD, MEd
Leah Mall

ABSTRACT

Background The Accreditation Council for Graduate Medical Education (AC
identify essential skills and develop means of assessment with supporting va
trainees on a milestone subcompetency related to working in interprofessio
any role on an interprofessional team in a variety of scenarios would be v
assessment.

Objective We developed a tool for simulation settings that assesses inter

Methods In 2015, existing tools that assess teamwork or interprofessional
reviewed for appropriateness, generalizability, adaptability, ease of use, an
included in a Delphi method with multidisciplinary pediatrics experts using
2017 to develop an assessment tool.

Results Thirty-one unique tools were identified. A 2-stage review narrowe
Twenty-two pediatrics experts participated in 4 rounds of Delphi surveys,
Sixteen items reached consensus for inclusion in the final tool. A global 4-
developed.

Conclusions A novel tool to assess interprofessional teamwork for individ
using a systematic review and Delphi methodology. This is the first step to e
tool for competency-based assessment.

Introduction

In the Accreditation Council for Graduate Medical Education (ACGME) Milestones projects, residents assess eith
and focus
The go

Simulation-Based Interprofessional Teamwork Assessment Tool (continued)

Question	Novice	Beginner	Competent	Proficient	Not Obs Not Ab Evalu
10. Participates in a collaborative relationship with other team members.	Does not establish collaborative relationships with others.	Occasionally establishes collaborative relationships with others.	Frequently establishes collaborative relationships with others.	Consistently establishes collaborative relationships with others.	Not observ not able evaluate
11. Practices active listening through closed-loop communication.	Does not use closed-loop communication.	Occasionally uses closed-loop communication.	Frequently uses closed-loop communication.	Consistently uses closed-loop communication.	Not observ not able evaluate
12. Works with other team members to shift roles to address urgent/emergent events when appropriate.	Does not establish collaborative relationships with others.	Occasionally establishes collaborative relationships with others.	Frequently establishes collaborative relationships with others.	Consistently establishes collaborative relationships with others.	Not observ not able evaluate
13. Listens respectfully to the expressed needs of all team members, including patient and family, in delivering care.	Does not listen to needs of team members, including patient and family.	Occasionally listens to expressed needs of team members, including patient and family.	Frequently listens to team members in an active and respectful manner.	Consistently listens to team members in an active and respectful manner and ensures a common understanding of care decisions.	Not observ not able evaluate
14. Provides care in a way that is mindful of the patient and their family.	Does not interact with patient/family.	Several examples where communication or care do not respond to patient/family needs.	Communicates treatment plan to patient/family, but fails to respond appropriately to their needs.	Most communication is appropriate and mindful of patient/family.	Not observ not able evaluate
15. Is open to opinions from other team members.	Does not acknowledge opinions of other team members.	Acknowledges some team member opinions, but ignores others.	Responds respectfully, but does not solicit others' opinions.	Solicits and responds respectfully to others' opinions.	Not observ not able evaluate
16. Contributes to team debriefing	Does not share information with the team. When asked, does not give input. Does not volunteer ideas despite prompting.	Shares information inconsistently with the team. When asked, gives input occasionally. Volunteers ideas occasionally with prompting only.	Shares information frequently with the team. When asked, gives input. Volunteers ideas with prompting.	Shares information consistently with the team. When asked, gives input. Volunteers ideas without prompting. Will give feedback if warranted.	Not observ not able evaluate

Empathy

Validation of a Modified Jefferson Scale of Empathy for Observers to Assess Trainees

Leah Mallory, MD; Rebecca Floyed, MD; Cara Doughty, MD; Tonya Thompson, MD, MA; Joseph Lopreiato, MD, MPH; Todd P. Chang, MD, MAcM

From the Tufts University School of Medicine (L Mallory), The Hannaford Simulation Center at Maine Medical Center, Maine; Dell Medical School (R Floyed), University of Texas at Austin, Tex; Baylor College of Medicine (C Doughty), Texas Children's Hospital Simulation Center, Baylor College of Medicine/Texas Children's Hospital, Houston, Tex; University of Arkansas for Medical Sciences (T Thompson), Little Rock, Ark; Uniformed Services University of the Health Sciences (J Lopreiato), The Val G. Hemming Simulation Center, Silver Spring, Md; and Keck School of Medicine of USC (TP Chang), Children's Hospital Los Angeles, Los Angeles, Calif

The authors have no conflicts of interest to disclose.

Address correspondence to Leah Mallory, MD, Medical Director, The Hannaford Si
Center, 22 Bramhall St, Portland, ME 04102 (e-mail: mallol@mmc.org).

Received for publication July 10, 2019; accepted June 7, 2020.

Abstract

Objective: "Demonstrate insight and understanding into emotion" is a competency amenable to simulation-based assessment. The Jefferson Scale of Patient Perceptions of Physician Empathy (JSPPPE) has validity evidence for patients to assess provider empathy. A version adapted for a third-party observers does not exist. Our aim was to modify the JSPPPE and use recorded standardized encounters to obtain validity evidence.

Methods: This cross-sectional study used video and data collected from 2 pediatric residencies. In 2018, 4 raters reviewed 24 videos of 12 interns communicating with standardized patients (SP) in 2 encounters and completed a modified JSPPE for observers (JSEO). Reliability between raters was established using Intraclass Correlations (ICC). JSEO mean scores were correlated to Essential Elements of Communication (EEC), JSPPPE, and faculty composite interpersonal communication (IC) scores using Spearman Rank.

Results 0.755). W an outlier. raters was correlated (rho = 0.6

Conclus modified recorded educators of this t appealing

Keywor ric Milest

Academi

Jefferson Scale for Physician Empathy for Observers

Instructions: We would like to know the extent of your agreement or disagreement with *each* of the following statements *about the physician named below*. Please use the following 7-point scale and write your rating number from 1 to 7 on the underlined space before each statement (1 means you Strongly Disagree, and 7 means you Strongly Agree with the statement, a higher number indicates more agreement).

1-----2-----3-----4-----5-----6-----7
Strongly Disagree *Strongly Agree*

Dr. (Name of the physician in here) _____

1. ____Can view things from the patient's perspective (see things as the patient would).
2. ____Asks about what is happening in the patient's daily life.
3. ____Seems concerned about the patient and their family.
4. ____Understands the patient's emotions, feelings and concerns.
5. ____Is an understanding doctor.

Figure 1. *Jefferson Scale for Physician Empathy for Observers.*

Maine Medical Center MaineHealth | Tufts University Sch Medicine

Ambiguity

"STATUS"

(Scalable Tolerating Ambiguity/Uncertainty Tool Utilizing Simulation)

- Instructor Assessement
  - *11 skills*
  - *3 pt Likert*
- Self Assessment
  - *11 attitudes*
  - *5 pt Likert Scale*

	Internship Class		
	Late Interns (n=10)	Early Interns (n=9)	Estimate (95% CI)
Instructor Assessment (IA)			
Q1 – Assessed their readiness to hear bad news?	1.40 ± 0.52	1.56 ± 0.68	-0.16 (-0.74, 0.43)
Q2 – Asked for their perspective and values when discussing treatments or making recommendations?	1.10 ± 0.74	1.22 ± 0.67	-0.12 (-0.81, 0.56)
Q3 – Acknowledged his/her limits in speaking knowledgably about this?	1.90 ± 0.32	1.28 ± 0.67	**0.62 (0.13, 1.12)‡**
Q4 – Only used simple non-medical language to suit the patient or family's understanding?	1.35 ± 0.47	1.22 ± 0.44	0.13 (-0.32, 0.57)
Q5 – Led a nuanced discussion of the strengths and weaknesses of the data?	1.00 ± 0.00	1.11 ± 0.33	-0.11, (-0.33, 0.11)
Q6 – Encouraged questions?	1.20 ± 0.42	1.33 ± 0.50	-0.13 (-0.58, 0.31)
Q7 – Asked them to repeat back explanations to ensure their comprehension?	1.60 ± 0.52	2.00 ± 0.00	**-0.40 (-0.76, -0.04)‡**
Q8 – Balanced the conversation between reality and hope/optimism?	1.00 ± 0.47	1.00 ± 0.00	0 (-0.33, 0.33)
Q9* - Offered to collaborate together on a course of action?	2.00 ± 0.00	2.00 ± 0.00	---
Q10 – Offered resources or sought the help of experts to answer questions?	1.70 ± 0.48	1.61 ± 0.49	0.09 (-0.38, 0.56)
Q11 – Sympathized while encouraging them to move forward?	1.10 ± 0.57	1.56 ± 0.53	-0.46 (-0.99, 0.08)
Overall IA Score	15.35 ± 1.56	15.89 ± 1.45	-0.54 (-2.01, 0.93)
Learner Self-Assessment (LSA)			
Q1 – The uncertainty of patient care troubles me.	2.60 ± 1.17	3.44 ± 1.13	-0.84 (-1.96, 0.27)
Q2 – I try to avoid situations with uncertain outcomes.	1.90 ± 0.57	2.33 ± 1.22	-0.43 (-1.34, 0.47)
Q3 – When I am uncertain of a diagnosis, I imagine all sorts of bad scenarios…	2.30 ± 0.82	3.22 ± 0.83	**-0.92 (-1.72, -0.12)‡**
Q4 – I usually feel anxious when I am not sure of a diagnosis.	2.60 ± 1.26	3.56 ± 0.88	-0.96 (-2.02, 0.11)
Q5 – Not being sure of what is best for a patient is one of the most stressful parts of being a physician.	3.10 ± 1.10	4.00 ± 0.50	**-0.90 (-1.74, -0.06)‡**

Simulation for "just in time" assessment

Effect of Just-in-time Simulation Training on Tracheal Intubation Procedure Safety in the Pediatric Intensive Care Unit

Akira Nishisaki, M.D.,* Aaron J. Donoghue, M.D., M.S.C.E.,† Shawn Colborn, R.R.T.,‡ Christine Watson, R.N.,§ Andrew Meyer, M.D.,‖ Calvin A. Brown III, M.D.,# Mark A. Helfaer, M.D.,** Ron M. Walls, M.D.,†† Vinay M. Nadkarni, M.D.‡‡

ABSTRACT

Background: Tracheal intubation-associated events (TIAEs) are common (20%) and life threatening (4%) in pediatric intensive care units. Physician trainees are required to learn tracheal intubation during intensive care unit rotations. The authors hypothesized that "just-in-time" simulation-based intubation refresher training would improve resident participation, success, and decrease TIAEs.

Methods: For 14 months, one of two on-call residents, nurses, and respiratory therapists received 20-min multidisciplinary simulation-based tracheal intubation training and 10-min resident skill refresher training at the beginning of their on-call period in addition to routine residency education. The rate of first attempt and overall success between refresher-trained and concurrent non–refresher-trained residents (controls) during the intervention phase was compared. The incidence of TIAEs between preintervention and intervention phase was also compared.

Results: Four hundred one consecutive primary orotracheal intubations were evaluated: 220 preintervention and 181 intervention. During intervention phase, neither first-attempt success nor overall success rate differed between refresher-trained residents *versus* concurrent non–refresher-trained residents: 20 of 40 (50%) *versus* 15 of 24 (62.5%), $P = 0.44$ and 23 of 40 (57.5%) *versus* 18 of 24 (75.0%), $P = 0.19$, respectively. The resident's first attempt and overall success rate did not differ between preintervention and intervention phases. The incidence of TIAE during preintervention and intervention phases was similar: 22.0% preintervention *versus* 19.9% intervention, $P = 0.62$, whereas resident participation increased from 20.9% preintervention to 35.4% intervention, $P = 0.002$. Resident participation continued to be associated with TIAE even after adjusting for the phase and difficult airway condition: odds ratio 2.22 (95% CI 1.28–3.87, $P = 0.005$).

Conclusions: Brief just-in-time multidisciplinary simulation-based intubation refresher training did not improve the resident's first attempt or overall tracheal intubation success.

* Attending Physician, Department of Anesthesiology and Critical Care Medicine, Center for Simulation, Advanced Education and Innovation, ‡ Clinical Specialist for Pediatric Intensive Care Unit, Respiratory Care, Department of Respiratory Care, § Pediatric Intensive Care Unit Level 4 Nurse, Department of Nursing, ‖ Clinical Fellow, Division of Neonatology, Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania. † Assistant Professor, Department of Anesthesiology and Critical Care Medicine, Center for Simulation, Advanced Education and Innovation, The Children's Hospital of Philadelphia, and Department of Pediatrics, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. # Assistant Professor, †† Professor, Department of Emergency Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. ** Professor, Department of Anesthesiology and Critical Care Medicine, The Children's Hospital of Philadelphia, and Department of Pediatrics, School of Medicine, University of Pennsylvania. ‡‡ Associate Professor, Department of Anesthesiology and Critical Care Medicine, Center for Simulation, Advanced Education and Innovation, The Children's Hos-

Central Venous Catheter Dress Rehearsals

Translating Simulation Training to Patient Care and Outcomes

Amy K. Scholtz, MSN;
Anne Marie Monachino, MSN;
Akira Nishisaki, MD, MSCE;
Vinay M. Nadkarni, MD;
Evie Lengetti, MSN

Introduction: Central line–associated blood stream infection (CLABSI) is a preventable burden to our current health care system. Inconsistencies in knowledge and practice of central venous catheters (CVC) dressing change procedures are associated with CLABSI. We hypothesized that participation in a "just-in-time" and "just-in-place" CVC dressing change program would improve nurses' knowledge, confidence, and psychomotor performance on mannequins (eg, T1 outcomes). Moreover, this simulation program would be associated with improved procedural competence on real patients (T2 outcomes) and hospital CLABSI rate (T3 outcomes).

Methods: We conducted a prospective before and after timed series study at a large urban children's hospital. This program provided an opportunity to practice a CVC dressing change using a simulated patient chest/arm. Cognitive and psychomotor skills were evaluated using a pre–self-assessment/post–self-assessment, written knowledge test and direct observation using a standardized checklist. Central line–associated blood stream infection rates were monitored monthly by the Office of Quality and Patient Safety.

Results: Five hundred twenty-four inpatient nurses participated in this program between November 2008 and May 2010. Knowledge and self-confidence improved significantly (knowledge, 4.1 [0.7] vs. 4.6 [0.5], $P < 0.001$; self-confidence, 4.1 [0.8] vs. 4.6 [0.6], $P < 0.001$). Of 2469 real-patient CVC dressing changes observed, dress rehearsal trainees required fewer corrective prompts (9% vs. 21%, $P < 0.001$), and CLABSI rates decreased from 5.3/1000 to 2.9/1000 line days ($P < 0.001$) during the study.

Discussion: This program improved nurse's knowledge, self-confidence, and psychomotor skill performance on mannequins (eg, T1 outcomes). These improvements were associated with improved procedural competence on real patients (T2 outcomes) and a temporal association with decreased hospital CLABSI rates (T3 outcomes).

(*Sim Healthcare* 8:341–349, 2013)

Advantages and Challenges: with "just in time" assessment

- Advantages
  - Intuitively makes sense
  - Review anatomy/protocol/process before clinical intervention
  - Build confidence and preparedness
- Challenges
  - Increase confidence does not translate to competence
  - Fidelity
  - Unrealistic context (IP team or family member present)
  - Busy clinical units
  - Faculty preparedness

Conclusions

- Simulation is an excellent modality for assessment:
  - Structured (can highlight a variety of knowledge, skills, or attitudes)
  - Reproducible (can run a group of learners through standardized scenarios)
  - Timely (can schedule according to need, evaluate real time or asynchronously via recorded encounter)
- Key question: formative vs summative?
- More wiggle room with formative assessment
- Summative assessments- must consider validity evidence
- "Just in Time" Simulation Assessment- makes sense but evidence is mixed

Exercise

- As educators, what do you find difficult to assess with current modalities?
  - Is your assessment need summative or formative?
  - What current method is used to assess this?
  - Would simulation present a solution for assessment?
    - » How would the simulation be structured?
    - » What would the assessment tool look like?
    - » How would you ensure validity?