The Effect of Faculty Training and Personality Characteristics on High Stakes Assessment of Simulation Performance

Ann Holland, PhD, RN, Bethel University
Deborah Bambini, PhD, PhD, WHNP-BC, CNE, CHSE, Grand Valley State University
Linda Blazovich, DNP, RN, CNE, St. Catherine University
Dorie Fritz, MSN, RN, Normandale Community College
Vicki Schug, PhD, RN, CNE, St. Catherine University
Jone Tiffany, DNP, RN, CNE, CHSE, ANEF, Bethel University

Conflicts of Interest and Disclosures:

This study was supported, in part, by a grant from the National League for Nursing, Washington DC.

Session Objectives:

- Describe the need for evaluator training and development of shared mental models for high stakes assessment of student performance in simulation
- Describe an experimental study design that includes a training intervention developed to prepare faculty for performance evaluation
- Discuss results of the study and implications for nursing education programs and future research

Background for Study

High-Stakes Assessment in Simulation

High Stakes Assessment in Simulation

- Simulation is evolving from a teaching/learning strategy used for formative assessment to…..
- A summative and high stakes assessment method
  - NCSBN study (Hayden, Smiley, Alexander, Kardong-Edgren, & Jeffries, 2014)
  - NLN study (Kardong-Edgren, Oermann, Rizzolo, & Odom-Maryon, 2017; Rizzolo, Kardong-Edgren, Oermann, Jeffries, 2015)
- How do we ensure validity and reliability in high stakes assessment in simulation?
- How do we ensure fair and equitable assessment?

High stakes assessment:

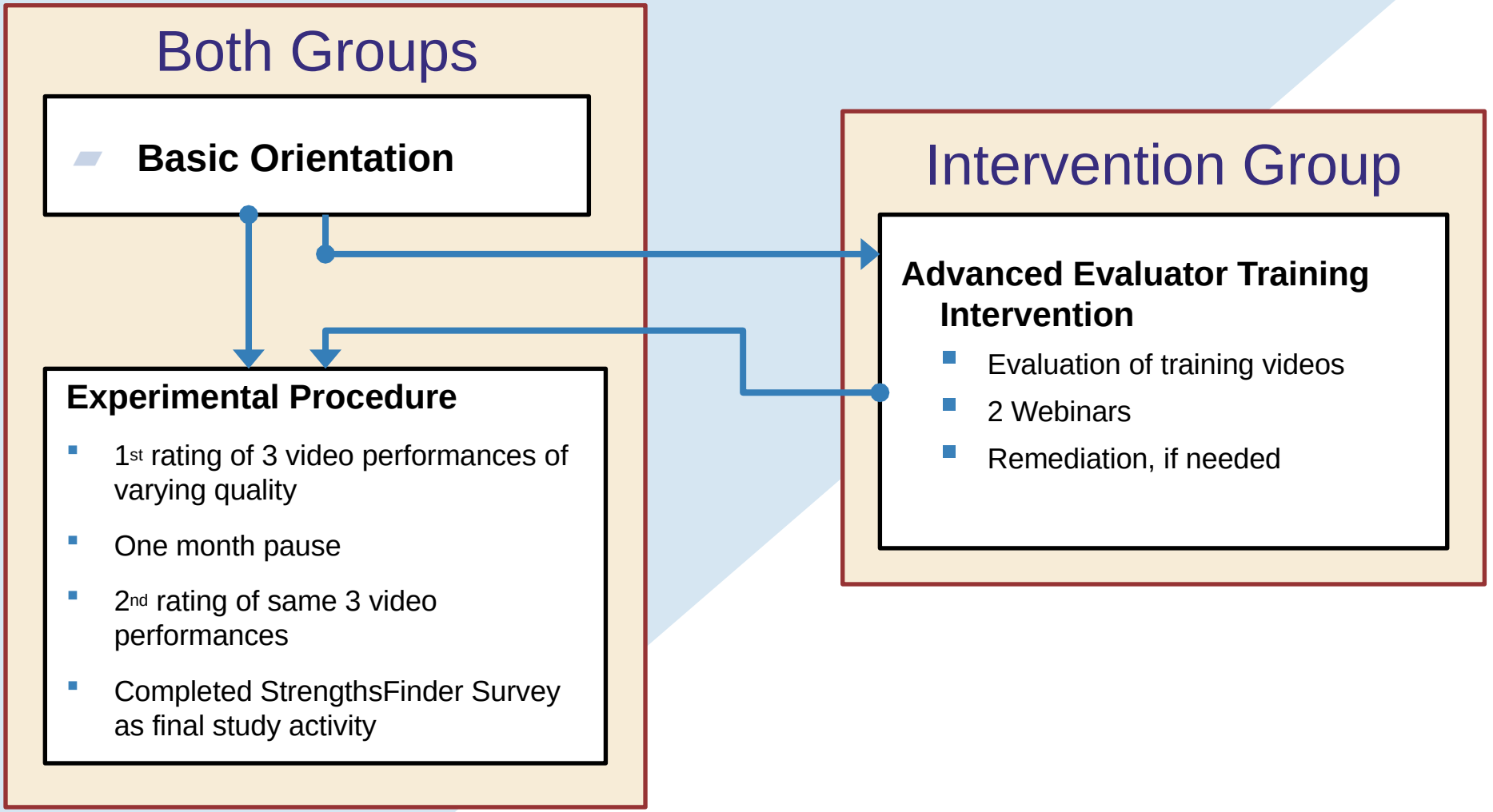
“an evaluation process associated with a simulation activity that has a major academic, educational, or employment consequence . . .” (Meakim et al., 2013, p. S7).

.

Research Study Design

Experimental, Randomized, Controlled Study

- What is the effect of (a) a training intervention and (b) faculty personality characteristics on faculty ability to achieve intra/inter-rater reliability when evaluating student performance during simulation?
- Multi-methods: IRR, qualitative survey
- Conceptual Framework
- Nationwide recruitment of 102 pre-licensure RN faculty
- Control and Intervention Groups


Both Groups
Basic Orientation
Experimental Procedure
1st rating of 3 video performances of varying quality
One month pause
2nd rating of same 3 video performances
Completed StrengthsFinder Survey as final study activity
Intervention Group
Advanced Evaluator Training Intervention
Evaluation of training videos
2 Webinars
Remediation, if needed

Data Collection Instruments

- Demographic Survey
- Creighton Competency Evaluation Instrument (CCEI)
- 10 video-recorded student performance videos
- Clifton StrengthsFinder Inventory
- StrengthsFinder Survey

Findings/Results

Participants

- 75 of the 102 participants completed the study
- Retained equal numbers in Control (37) and Intervention (38) groups
- No statistically significant differences between the groups in demographic characteristics such as age, years of teaching, years of teaching with simulation

Inter-rater Reliability: Total Score

Video	Intervention Group		Control Group	
	ICC	Kappa	ICC	Kappa
1	.869	.769	.794	.527
4	.969	.541	.738	.311
6	.934	.411	.768	.294
11	.922	.461	.771	.363
19	.962	.658	.861	.584
28	.952	.705	.567	.615

ICC
0.75-1.0: excellent
0.60-0.74: good
0.40-0.59: fair
<0.40: poor

Kappa
>0.80: high
0.60-0.79: substantial
0.40-0.59: moderate
<0.40: poor to fair

Intra-rater Reliability: Total Score

Video	Intervention Group		Control Group	
	ICC	Kappa	ICC	Kappa
1	.811	.951	.912	.527
4	.825	.885	.844	.541
6	.858	.723	.770	.503
11	.896	.821	.879	.522
19	.957	.862	.953	.829
28	.926	.891	.576	.616

ICC
0.75-1.0: excellent
0.60-0.74: good
0.40-0.59: fair
<0.40: poor

Kappa
>0.80: high
0.60-0.79: substantial
0.40-0.59: moderate
<0.40: poor to fair

Personality characteristics

- Most participants thought the StrengthsFinder results were accurate
- Participants explained how knowing their strengths could help with the evaluation process
  - “As I learn more about myself, I feel I can provide more consistent evaluation”
- Participants reflected on how their strengths may make it more difficult to evaluate students fairly
  - Too strict, too lenient, don’t like conflict, difficulty being consistent

Study Conclusions

Conclusions

- A structured training method that included a model evaluation and practice evaluations contributed to higher inter and intra rater reliability.
- Providing participants with a video recorded “model” evaluation was a useful method to initiate a SMM in a diverse group of evaluators.
- Live discussions allowed participants to express their values and beliefs about clinical competence in attempts to reach consensus.
- Participants who were initially resistant agreed, through continued discussion, to adopt a reasonable, if not perfect, SMM.
- The accuracy of video performance evaluation was highly dependent on the quality of audio and video capture and the

Study Implications

Education and Research

Implications for Education

- Communication and sharing of individual beliefs and values about competency must occur among team members.
- An intentional process must be planned that allows time for consensus to occur.
- Evaluators must be trained in the use of a valid evaluation tool.
- The process should include practice evaluations using live or high-quality video recorded performances of differing competency levels.

Implications for Research

- Additional studies are needed to test and confirm the effects of training methods on reliability.
- Funding sources are needed to support large scale experimental studies.
- The effect of faculty personality characteristics on student performance evaluation requires more study.
- Studies are needed to examine evaluator decision-making using checklists versus global rating scores.

Questions?

References

Hayden, J. K., Smiley, R. A., Alexander, M., Kardong-Edgren, S., & Jeffries, P. R. (2014). Supplement: The NCSBN national simulation study: A longitudinal, randomized, controlled study replacing clinical hours with simulation in prelicensure nursing education. *Journal of Nursing Regulation, 5*(2), C1-S64.

International Nursing Association for Clinical Simulation & Learning. (2016). INACSL standards of best practice: Simulation℠ participant evaluation. INACSL Standards Committee. *Clinical Simulation in Nursing*, *12*, S26-S29. http://dx.doi.org/10/1016.j.ecns.2016.09.009

Kardong-Edgren, S., Oermann, M. H., Rizzolo, M. A., & Odom-Maryon, T. (2017). Establishing inter- and intrarater reliability for high-stakes testing using simulation. *Nursing Education Perspectives, 38*(2), 63-68. doi: 10.1097/01.NEP.0000000000000114

References

McComb, S., & Simpson, V. (2014). The concept of shared mental models in healthcare collaboration. *Journal of Advanced Nursing*, *70*(7), 1479-1488. doi:10.1111/jan.12307

National League for Nursing. (2012). Fair testing guidelines for nursing education. Retrieved from http://www.nln.org/fairtestguidelines

Rizzolo, M. A., Kardong-Edgren, S., Oermann, M. H, & Jeffries, P. R. (2015). The National League for Nursing project to explore the use of simulation for high-stakes assessment: Process, outcomes, and recommendations. *Nursing Education Perspectives, 36*(5), 299-303. doi:10.5480/15-1639