

# High Fidelity Simulation Is Associated With Good Discriminability In Emergency Medicine Residents' In-Training Examinations

**Shou-Yen Chen**

Chang Gung Memorial Hospital Linkou Main Branch: Chang Gung Memorial Hospital

**Chung-Hsien Chaou**

Chang Gung Memorial Hospital Linkou Main Branch: Chang Gung Memorial Hospital

**Shiuan-Ruey Yu**

Chang Gung Medical Research Center

**Yu-Che Chang**

Chang Gung Memorial Hospital Linkou Main Branch: Chang Gung Memorial Hospital

**Chip-Jin Ng**

Chang Gung Memorial Hospital Linkou Main Branch: Chang Gung Memorial Hospital

**Pin Liu** (✉ [sisterweber@yahoo.com.tw](mailto:sisterweber@yahoo.com.tw))

Chang Gung Memorial Hospital <https://orcid.org/0000-0001-7872-5405>

---

## Research article

**Keywords:** emergency department, high-fidelity simulation, in-training examination, resident training

**Posted Date:** October 7th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-80182/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Medicine on June 18th, 2021. See the published version at <https://doi.org/10.1097/MD.00000000000026328>.

# Abstract

## Background

In-training examinations (ITEs), arranged during residency training, evaluate the residents' performances periodically. The ITE usually consists of written and oral forms. There is limited literature focusing on the effectiveness of emergency medicine (EM) resident ITEs in the format of simulation-based examinations, as compared to traditional oral or written tests. Our primary objective is to compare high fidelity simulation with written and oral tests in the evaluation of the residents with different seniority in EM training programs.

## Methods

During the 5-year study period, 8 ITEs were administered to 68 EM residents, and 253 ITE scores were collected. Different ITE scores were calculated and presented as mean and standard deviation. Discrimination of ITE scores between different training years of residency was examined using one-way ANOVA test.

## Results

The high fidelity simulation scores correlated to the EM training year, and residents in their fourth training year (R4) had the highest scores consistently, followed by R3, R2, and then R1. The oral test scores had similar results but not as consistent as the high fidelity simulation tests. The written test scores distribution failed to discriminate the residents' seniority. The high fidelity simulation test had the best discriminative ability and better correlation between different EM residency training years comparing to oral or written forms.

## Conclusions

High fidelity simulation tests had a good discriminative ability and were well correlated to the EM training year. We suggest high-fidelity simulation should be a part of ITE of EM training programs and maybe adopted in EM board certification examination in the future.

# Introduction

The goal of postgraduate medical education is to facilitate the resident's acquisition of medical knowledge and clinical skills, and to support the achievement of the competency required to practice in a medical specialty(1). It is imperative to evaluate the residents' performance periodically, in order to facilitate them to overcome their weakness and to ensure the quality of the residency program. The Accreditation Council for Graduate Medical Education (ACGME) and the Council of Emergency Medicine Residency Directors (CORD) emphasize the value of practical and reliable assessment tools to evaluate the ability of residents and the effectiveness of the residency programs.(2-4)

In-training Examination (ITE), also known as in-service examination, was firstly introduced in 1963 to residents in medical specialty by the American Academy of Orthopedic Surgeons.(5) In either oral or written forms, ITE has been adopted by various medical specialties as a powerful and multi-functional assessment tool to measure the performance of residents.(6–9) For emergency medicine (EM) residents, the ITE is not merely a signboard of their current academic performance but also a chance to prepare for the board exam. The ITE offers an opportunity for them to review their deficiencies in medical knowledge and to improve themselves.(10) Prior studies have found a positive correlation between EM ITE scores and the American Board of Emergency Medicine (ABEM) written board certification scores.(11–13) Similar correlation of the ITE and board exam score was also documented in the field of internal medicine and its subspecialty of cardiovascular disease.(14, 15)

EM in-training examination, as well as board-certification qualifying examination, usually consists of written and oral tests.(6, 13) Some of these types of examinations measure the degree of medical knowledge and clinical skills, but may not correlate to the overall clinical performance in EM residency, which is supposed to advance with EM residency training year. Furthermore, clinical productivity (number of patient encounters per hour), as a "real-life" in the emergency department (ED), correlated poorly with academic performance when measured by oral ITE and oral/written ABEM scores.(13) ED is a complicated lieu, which is characterized by time limits, various independent tasks with undetermined priorities, a rapidly changing environment full of interruptions, and the need for education and training(12). The progress of multitasking ability by EM residency training year may not be directly reflected by written and oral test scores.

High fidelity simulations, using computer-controlled mannequins, are being used throughout medical education to emulate real patient encounters since the 1960s.(16, 17) In addition, simulation is increasingly being used as an assessment tool particularly for procedures.(18, 19) Simulation-based teaching was also integrated into the curriculum, including system-based practice and communication skill training course for EM residents.(20–22) Simulation-based assessment can be used to evaluate the EM residents' competency in differential diagnosis, resuscitation, procedures and the year-to-year progression of performance, both formatively and summatively.(23) Some literature supports the use of simulation-based assessment tools in the evaluation of residents, as in anesthesiology residency.(24–26) Some even advocated for the use of simulation-based tests in board certification exams.(19, 27, 28)

Currently, there is limited literature focusing on the effectiveness of EM resident ITEs in the format of simulation-based examinations, as compared to traditional oral or written tests. Our primary objective is to compare high fidelity simulation with written and oral tests in evaluation of the residents with different seniority in EM training programs.

## Method

### Study setting

This is a retrospective cohort study of EM resident physicians. The study was conducted at a university-affiliated tertiary teaching hospital with a 3,600-bed capacity and an estimated annual ED volume of 180,000 patient visits. There are 63 board-certified EM faculty members within the department. The residency program accepts 7–10 resident physicians each year. The study was approved by our institutional review board (IRB no.202000099B0).

## Participants and data collection

Our EM training program administers in-training examination (ITE) to all residents biannually, usually in February and August. All residents needed to participate in the examinations unless specific conditions were met, such as severe illness. The first-year resident (R1) physicians would not take the examination in August because they usually registered in the same month. The ITEs were supervised by the residency program director and organized by the education committee within the department.

The ITE consists of 3 different forms: written examination, oral examination, and high-fidelity examination (oral and high-fidelity examination was canceled in 2018 due to equipment problems). Written examination involved multiple choice questions (MCQs) and short answer question (SAQ) stations. MCQs had mixed content of EM. SAQs stations included electrocardiogram (ECG) reading, image test (radiograph, computed tomography, ultrasound), and other different EM themes. Oral simulation tests included 2 to 3 stations with different EM topics, such as internal medicine, pediatrics, toxicology, emergency medical system (EMS)...etc. The examiner of each station was EM board certified faculty member. High-fidelity simulation was operated with computer-controlled mannequin and 2 designed scenarios with different topics were used in each ITE.

We collected five years of ITE data. Eight ITEs were included in the study. Only one ITE was administered in 2017 because of a change in the management level. Average scores from written, oral simulation and high-fidelity exams were calculated. The written test, oral test and high-fidelity examination scores were the mean scores across all stations.

## Statistical analysis

Data were analyzed using SPSS software (version 13.0 for Windows; SPSS, Chicago, IL). In the descriptive analysis, categorical variables were presented as numbers and percentages. Different ITE scores were calculated and presented as mean and standard deviation. The reliability of ITE was evaluated using Cronbach's alpha coefficient. The discrimination of ITE scores between different training years of residency was examined using one-way ANOVA test. A  $p$  value  $< 0.05$  was considered statistical significance.

## Results

During the 5-year study period, 8 ITEs were administered to 68 EM residents training in our EM training program, and 253 ITE scores were collected. Table 1 reveals the characteristics of the ITEs and the participants. All 8 ITEs included written test, and 7 of them contained oral tests and high-fidelity

simulation tests. Examinations in February usually involved 35 or 36 EM residents, and examinations in August had 26 to 29 residents each time. The number of residents in each training year ranged from 7 to 10.

Table 1  
Characteristics of the in-training exams (ITEs) and participants

<b>Characteristics of ITEs</b>		<b>count</b>		
Total tests		8		
Containing written test		8		
Containing oral simulation test		7		
Containing High fidelity simulation test		7		
Number of participating residents (lowest-highest)		26–36		
Gender of participants				
Male		55		
Female		13		
Participants' year of initiation	Male	Female	Total	
2012	1	0	1	
2013	8	1	9	
2014	9	1	10	
2015	8	2	10	
2016	8	2	10	
2017	7	3	10	
2018	8	0	8	
2019	6	4	10	

The analysis of ITE content was revealed in Table 2. All the stations of SAQs, oral tests and high-fidelity simulation tests were analyzed. The most discriminative subject is internal medicine (85.71%), followed by pediatrics (80%) and toxicology (60%). Other topics revealed diverse results. The internal consistency of the ITE was 0.95, which indicated a good reliability across different participants.

Table 2  
List of domains of In-Training Exams

Domain of the test	Station Count	(%)
Written test stations	38	
MCQs	7	18.42
SAQs stations	31	81.58
ECG	8	21.05
Image	6	15.79
Ultrasound	6	15.79
Critical care medicine	4	10.53
Gynecology	2	5.26
Others	5	13.16
Oral test stations	15	
Internal medicine	5	33.33
Pediatrics	3	20
Toxicology	4	26.67
Other	3	20
High fidelity simulation stations	14	
Internal medicine	2	14.29
Trauma	6	42.86
Critical care	3	21.43
Pediatric	2	14.29
Toxicology	1	7.14

The written tests usually consisted of MCQs and several stations of SAQs. ECG reading and image tests are commonly covered in the SAQs each time. The most common topics of oral test were internal medicine and toxicology. Trauma and critical care most often appeared in the high-fidelity simulation tests. Table 3 shows the discrimination of the tests according to EM residency training year. Almost all oral and high fidelity simulation tests of each exam were discriminative except one. The discrimination of the written tests was poor. Although oral and high fidelity simulation tests were both discriminative, the scores distribution of these 2 kinds of tests were different. Figure 1 demonstrates the average scores of different EM training year residents in each ITE. The high fidelity simulation scores correlated to the EM training year and R4 had highest scores consistently, followed by R3, R2 and then R1 (Fig. 1c). The oral

test scores (Fig. 1b) had similar results but not as consistent as the high fidelity simulation tests. The written test scores distribution is lack of similarity (Fig. 1a). The high fidelity simulation test had the best discrimination and better correlation between different EM residency training years comparing to the other 2 test forms.

Table 3

Discriminability of in-training examinations, differentiated according to examination forms and test domains.

<b>Discriminability according to different examination forms</b>			
<b>Examination</b>	<b>Written test</b>	<b>Oral simulation test</b>	<b>High-fidelity simulation</b>
Test 1	p = 0.001*	p < 0.001*	p = 0.007*
Test 2	p = 0.967	p = 0.730	p = 0.249
Test 3	p = 0.733	p = 0.001*	p = 0.015*
Test 4	p = 0.384	-	-
Test 5	p = 0.084	p = 0.006*	p = 0.016*
Test 6	p = 0.032	p < 0.001*	p < 0.001*
Test 7	p = 0.261	p = 0.001*	p = 0.016*
Test 8	p < 0.001*	p < 0.001*	p = 0.004*
<b>Discriminability according to test domains</b>			
<b>Test domain</b>	<b>Total stations</b>	<b>Discriminative stations</b>	<b>Discriminative prob (%)</b>
Internal medicine	7	6	85.7
Pediatric	5	4	80.0
Toxicology	5	3	60.0
Mixed content(MCQ)	7	4	57.1
Critical care medicine	7	3	42.9
Trauma	7	3	42.9
Image	6	2	33.3
Sonography	6	1	16.7
ECG	8	1	12.5
Gynecology	2	0	(0)
* p value < 0.05			

## Discussion

To our knowledge, this is the first study to evaluate the utility of high fidelity simulation in the in-training examinations of EM training programs. Our study demonstrated good performance of high-fidelity simulation test to discriminate the difference in competency of EM residents in different training year(s). It may not be surprising that high fidelity simulation tests had better results than written tests. Medical knowledge, as evaluated by written tests, can be obtained by greater efforts in reading textbooks and memorizing core contents. Consequently, as long as he/she studies hard, it is possible for a junior resident to outperform a senior one in a written test. Nonetheless, clinical experience, which was obtained at the bedside gradually with EM residency training years, cannot be rapidly obtained by goal-directed curriculum reading or short-term memory. Previous study also demonstrated improvement in written ITE scores after a structured board review program only in junior residents, but not senior ones.(29) With the accumulation of "real-life" experience, the medical knowledge should be integrated into the clinical scenario, including disease pattern recognition, application of relevant algorithm, and immediate interventions to treat patients properly and timely.(30) The clinical experience, as a basis of competence in patient care, is much more than a simple memory recall. Rodgers' study in an ACLS course concluded that written evaluation is a poor predictor of skill performance.(31) Other studies also fail to use ITE written and oral scores to predict residents' clinical performance.(6, 13, 30)

Oral tests also showed good discrimination in our study. The junior EM residents (R1) almost always had inferior scores comparing to senior residents. Whereas, the results were inconsistent in senior residents, and R2 or R3 physicians sometimes had better scores than R4 physicians in oral tests. One possible explanation is that oral tests may reflect some degree of both clinical experience and medical knowledge, but not as realistic as high fidelity simulation tests. Hence, the scores depended more individually in senior residents as long as they had a certain extent of clinical experience after completing first year EM training program. Alternatively, high fidelity simulation tests created a vivid scenario, which can evoke a physiologic response comparable to real-life clinical situation.(32) Therefore, the performance in high fidelity simulation tests correlated better to the extent of clinical experience than oral tests. According to our result, assessment by high fidelity simulation tests seems to be more appropriate to measure clinical competency of the EM residents.

The high fidelity simulation tests also had several advantages over traditional written and oral tests in assessing certain ACGME core competencies including interpersonal and communication skills, professionalism, patient care, and systems-based practice.(21, 22) The direct evaluation by medical educators through simulation-based assessment provides simultaneous evaluation of knowledge, clinical reasoning, and teamwork.(33)The standardization, fidelity, and reproducibility of medical simulation scenarios make it especially suited to be used in ITEs.(18, 23) With the aforementioned advantages, high fidelity simulation is possibly a better assessment tool for EM training programs.(34)

ITE in EM training program provides not only the evaluation of medical knowledge and clinical competency, but also an opportunity for the residents to know their advantage and deficiency.(10)



Contrary to text-based learning and written tests, EM residents prefer question-based learning, which can be evaluated by simulation tests.(10) Assessment by simulation had higher satisfaction rates when compared to written tests.(35) Although high fidelity simulation test is certainly more expensive than a standard written or oral test, it may be more cost effective than other commonly used methods of assessment like "standardized patient" in Objective Structured Clinical Examination (OSCE). Our study demonstrated the excellent discrimination of the high fidelity simulation tests in ITE of EM training programs. Although further study is needed, it should be considered as a part of EM board certification examination.

## **Limitations**

Our study is a single-center study, the results may mostly reflect local situation. Additionally, the stations of in-training exam were drawn up by our EM faculty, which may not be generalizable to other specialties. Finally, the examiners of each ITE, which were our EM faculty, did not received additional training for each ITE, although the checklist of each station was provided and one station is scored by only one examiner.

## **Conclusion**

High fidelity simulation tests used in ITE had good discriminative ability and well correlated to the EM training year. We suggest high-fidelity simulation should be a part of ITE of EM training programs, and may be used in EM board certification examination in the future.

## **Declarations**

## **Ethics approval and consent to participate:**

The study was approved by the institutional review board of Chang Gung Memorial Hospital (IRB no.202000099B0).

### **Consent for publication:**

Not applicable.

### **Availability of data and materials:**

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

### **Competing interests:**

The authors declare that they have no competing interests.

## Authors' contributions

SYC contributed to the conceptualization, data collection and analysis, and draft writing of this research. CHC contributed to the research design, data acquisition, supervision of methodology, and data analysis. SRY contributed to the project administration, data acquisition, transcript coding, and participated in the analysis of themes. YCC contributed to the supervision of transcript coding and the emergence of themes. CJN contributed to the supervision of project. PL contributed to conceptualization of the research, software supervision and qualitative data analysis. All authors participated and contributed to the critical revision of the manuscript and gave final approval of version submitted for publication.

## Funding:

Not applicable.

## Author contributions:

SYC contributed to the conceptualization, data collection and analysis, and draft writing of this research. CHC contributed to the research design, data acquisition, supervision of methodology, and data analysis. SRY contributed to the project administration, data acquisition, transcript coding, and participated in the analysis of themes. YCC contributed to the supervision of transcript coding and the emergence of themes. CJN contributed to the supervision of project. PL contributed to conceptualization of the research, software supervision and qualitative data analysis. All authors participated and contributed to the critical revision of the manuscript and gave final approval of version submitted for publication.

## Acknowledgements:

Not applicable.

## References

1. World Medical Association. WMA Statement of Medical Education. Adopted by the 57th WMA General Assembly, Pilanesberg, South Africa, October 2006, and revised by the 68th WMA General Assembly, Chicago, United States, October 2017. [Available from: <https://www.wma.net/policies-post/wma-statement-on-medical-education/>].
2. Accreditation Council for Graduate Medical Education. ACGME Outcome Project Glossary 2011 [Available from: <http://www.acgme.org/outcome/project/glossary2.asp#2>].
3. Hobgood C, Promes S, Wang E, Moriarity R, Goyal DG, Council of Emergency Medicine Residency D. Outcome assessment in emergency medicine—a beginning: results of the Council of Emergency

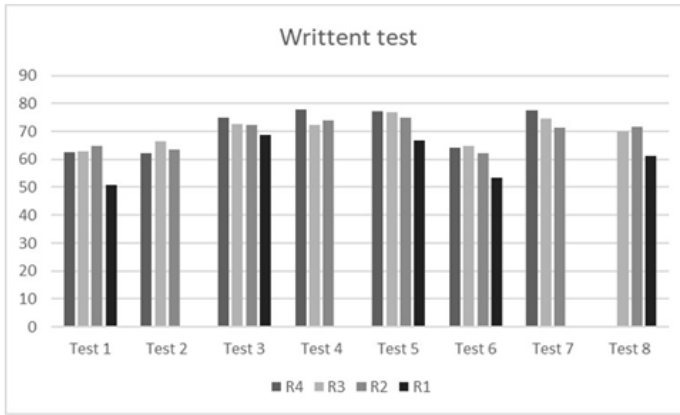
- Medicine Residency Directors (CORD) emergency medicine consensus workgroup on outcome assessment. *Acad Emerg Med*. 2008;15(3):267-77.
4. Getting Ahead of the Curve: the ACGME Best Practices/Core Competencies. Proceedings of the Council of Emergency Medicine Residency Directors (CORD) 2002 Consensus Conference. *Acad Emerg Med*. 2002;9(11):1213-367.
  5. Herndon CH. Quality control of surgeons-in-service examinations. *Bull Am Coll Surg*. 1977;62(3):14-7.
  6. Quattlebaum TG, Darden PM, Sperry JB. In-training examinations as predictors of resident clinical performance. *Pediatrics*. 1989;84(1):165-72.
  7. Hatch SS, Vapiwala N, Rosenthal SA, Plataras JP, Blumberg AL, Small W, Jr., et al. Radiation Oncology Resident In-Training Examination. *Int J Radiat Oncol Biol Phys*. 2015;92(3):532-5.
  8. Joseph JA, Terry CM, Waller EJ, Bortsov AV, Zvara DA, Mayer DC, et al. Enhancement of anesthesiology in-training exam performance with institution of an academic improvement policy. *J Educ Perioper Med*. 2014;16(6):E072.
  9. Godellas CV, Hauge LS, Huang R. Factors affecting improvement on the American Board of Surgery In-Training Exam (ABSITE). *J Surg Res*. 2000;91(1):1-4.
  10. Eastin TR, Bernard AW. Emergency medicine residents' attitudes and opinions of in-training exam preparation. *Adv Med Educ Pract*. 2013;4:145-50.
  11. Hamlin S CT, Frederick R, Hafner J, Thompson L. Correlation of emergency medicine residency training simulated oral Board examination scores with National ABEM Oral Examination scores[abstract]. *Acad Emerg Med*. 2008;15(Suppl 1):s54.
  12. Ledrick D, Fisher S, Thompson J, Sniadanko M. An assessment of emergency medicine residents' ability to perform in a multitasking environment. *Acad Med*. 2009;84(9):1289-94.
  13. Frederick RC, Hafner JW, Schaefer TJ, Aldag JC. Outcome measures for emergency medicine residency graduates: do measures of academic and clinical performance during residency training correlate with American Board of Emergency Medicine test performance? *Acad Emerg Med*. 2011;18 Suppl 2:S59-64.
  14. Grossman RS, Fincher RM, Layne RD, Seelig CB, Berkowitz LR, Levine MA. Validity of the in-training examination for predicting American Board of Internal Medicine certifying examination scores. *J Gen Intern Med*. 1992;7(1):63-7.
  15. Indik JH, Duhigg LM, McDonald FS, Lipner RS, Rubright JD, Haist SA, et al. Performance on the Cardiovascular In-Training Examination in Relation to the ABIM Cardiovascular Disease Certification Examination. *J Am Coll Cardiol*. 2017;69(23):2862-8.
  16. Denson JS, Abrahamson S. A computer-controlled patient simulator. *JAMA*. 1969;208(3):504-8.
  17. Gaba DM, DeAnda A. A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. *Anesthesiology*. 1988;69(3):387-94.

18. Boulet JR, Murray DJ. Simulation-based assessment in anesthesiology: requirements for practical implementation. *Anesthesiology*. 2010;112(4):1041-52.
19. Marco J, Holmes DR, Jr. Simulation: present and future roles. *JACC Cardiovasc Interv*. 2008;1(5):590-2.
20. Binstadt ES, Walls RM, White BA, Nadel ES, Takayesu JK, Barker TD, et al. A comprehensive medical simulation education curriculum for emergency medicine residents. *Ann Emerg Med*. 2007;49(4):495-504, e1-11.
21. Frallicciardi A, Lotterman S, Ledford M, Prenovitz I, Meter RV, Kuo CL, et al. Training for Failure: A Simulation Program for Emergency Medicine Residents to Improve Communication Skills in Service Recovery. *AEM Educ Train*. 2018;2(4):277-87.
22. Wang EE, Vozenilek JA. Addressing the systems-based practice core competency: a simulation-based curriculum. *Acad Emerg Med*. 2005;12(12):1191-4.
23. McLaughlin S, Fitch MT, Goyal DG, Hayden E, Kauh CY, Laack TA, et al. Simulation in graduate medical education 2008: a review for emergency medicine. *Acad Emerg Med*. 2008;15(11):1117-29.
24. Blum RH, Boulet JR, Cooper JB, Muret-Wagstaff SL. Simulation-based assessment to identify critical gaps in safe anesthesia resident performance. *Anesthesiology*. 2014;120(1):129-41.
25. Sidi A, Gravenstein N, Vasilopoulos T, Lampotang S. Simulation-Based Assessment Identifies Longitudinal Changes in Cognitive Skills in an Anesthesiology Residency Training Program. *J Patient Saf*. 2017.
26. Blum RH, Boulet JR, Cooper JB, Muret-Wagstaff SL, Harvard Assessment of Anesthesia Resident Performance Research G. Simulation-based assessment to identify critical gaps in safe anesthesia resident performance. *Anesthesiology*. 2014;120(1):129-41.
27. Crosby E. The role of simulator-based assessments in physician competency evaluations. *Can J Anaesth*. 2010;57(7):627-35.
28. Decker S, Utterback VA, Thomas MB, Mitchell M, Sportsman S. Assessing continued competency through simulation: A call for stringent action. *Nurs Educ Perspect*. 2011;32(2):120-5.
29. Gillen JP. Structured emergency medicine board review and resident in-service examination scores. *Acad Emerg Med*. 1997;4(7):715-7.
30. Ryan JG, Barlas D, Pollack S. The relationship between faculty performance assessment and results on the in-training examination for residents in an emergency medicine training program. *Journal of graduate medical education*. 2013;5(4):582-6.
31. Rodgers DL, Bhanji F, McKee BR. Written evaluation is not a predictor for skills performance in an Advanced Cardiovascular Life Support course. *Resuscitation*. 2010;81(4):453-6.
32. Bong CL, Lightdale JR, Fredette ME, Weinstock P. Effects of simulation versus traditional tutorial-based training on physiologic stress levels among clinicians: a pilot study. *Simul Healthc*. 2010;5(5):272-8.
33. Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356(4):387-96.

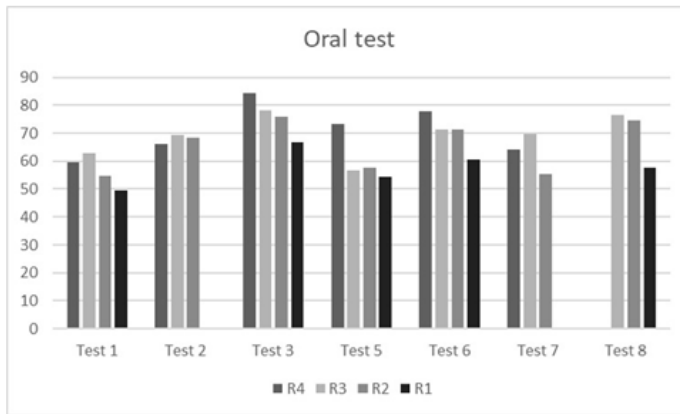
34. Sherbino J, Bandiera G, Frank JR. Assessing competence in emergency medicine trainees: an overview of effective methodologies. *CJEM*. 2008;10(4):365-71.
35. Christian MR, Sergel MJ, Mycyk MB, Aks SE. Comparison of High-Fidelity Medical Simulation to Short-Answer Written Examination in the Assessment of Emergency Medicine Residents in Medical Toxicology. *Mo Med*. 2017;114(5):396-9.

## Figures

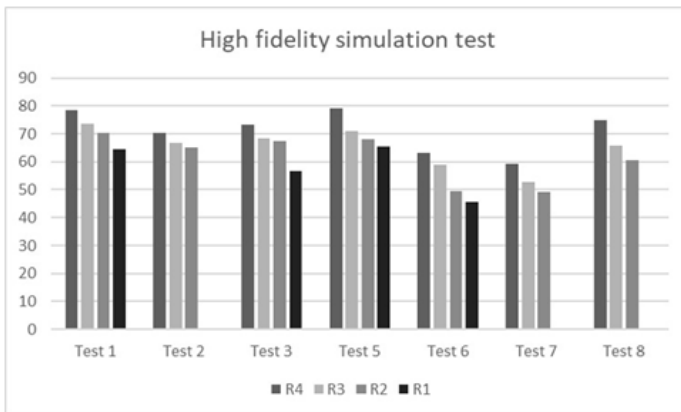
(a)



(b)



(c)



**Figure 1**

Illustration of test scores from (a) the written tests, (b) the oral tests, and (c) the high fidelity simulation tests of residents in different training years.