An Investigation of the Integration of Technology to Enhance Consistency in Grading Clinical Skills

Sabine Dunbar, DNP, NMW, RN

INTRODUCTION

- Current focus on quality and safety in healthcare requires accurate evaluation of performance ensuring safe clinical practice.
- Vital aspect of health professional education is evaluation of clinical competence. Nurse educators are charged with ensuring students become safe and competent practitioners.
- Development of competence begins during prelicensure nursing education.
- Clinical skills laboratories provide learning environments where prelicensure nursing students become proficient in performing skills prior to practicing these on patients in clinical settings.
- Clinical skills laboratories also serve to evaluate student performance of clinical skills, an important component of nursing competence.
- Evaluation commonly accomplished through direct observation by nurse educators - may be flawed with inconsistency in grading.
- Physical assessment skills commonly taught in prelicensure nursing education, with summative evaluation through student demonstration of physical examination on simulated patient.
- PROBLEM - Some students report inconsistency among educators grading summative physical examination in a baccalaureate, prelicensure health assessment course.

PURPOSE

The purpose of this pilot project was to investigate consistency among nursing faculty grading summative physical examinations in a health assessment course through the integration of audio-visual technology to compare live and video review methods of grading.

METHODS

A descriptive, comparative design was used to compare live grading to grading based on a video recording, and to measure agreement among nurse educators.

RESULTS

- Total grades for live review of the physical examination ranged from 39.75-47.5 out of a total of 53 possible points (75.89-89.6%), while grades for video review ranged from 39.25-44.5 (74.06-83.9%).
- There were no significant differences in total scores for live review (M=43.25, SD=2.89) versus total scores for video review (M=42.25, SD=2.02); t(5)=1.26, p=0.26.
- A comparison of total and mean live and video review scores for body systems sections of the checklist also revealed no differences.
- Item specific percent agreement for both live and video review ranged from 33.3-100% (Table 1).

<table>
<thead>
<tr>
<th>Grading Method</th>
<th>33.3%</th>
<th>50%</th>
<th>66.7%</th>
<th>83.3%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live Review</td>
<td>1</td>
<td>9</td>
<td>11</td>
<td>26</td>
<td>35</td>
</tr>
<tr>
<td>Video Review</td>
<td>1</td>
<td>8</td>
<td>13</td>
<td>21</td>
<td>38</td>
</tr>
</tbody>
</table>

- Mean interrater percent agreement for live review = 83.94%; for video review = 84.56%.
- Intrarater percent agreement across faculty evaluators ranged from 64.2-88.9% (Table 2).

<table>
<thead>
<tr>
<th>Participant</th>
<th>% Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>85.10</td>
</tr>
<tr>
<td>B</td>
<td>79.03</td>
</tr>
<tr>
<td>C</td>
<td>82.72</td>
</tr>
<tr>
<td>D</td>
<td>64.20</td>
</tr>
<tr>
<td>E</td>
<td>82.72</td>
</tr>
<tr>
<td>F</td>
<td>88.89</td>
</tr>
</tbody>
</table>

- A discrepancy in pass/fail determination was found with both methods of grading, with one evaluator assigning a failing grade, and five evaluators allocating passing grades.

CONCLUSIONS

- Comparison of live and video reviews for grading physical examinations revealed no significant differences in scores.
- Percent agreement revealed slightly higher interrater agreement with video review (84.56%) vs live review (83.94%).
- May be related to ability to replay recordings during video review, which some participants identified as helpful.
- Interrater agreement with both methods of grading were considered acceptable (80% agreement recommended as minimum (Mohag, 2012)).
- Despite acceptable interrater agreement, discrepancy in pass/fail determinations across evaluators existed, suggesting consistency in grading among evaluators can be improved.
- Analysis of intrarater agreement identified need for enhancing consistent rating of performance by individual evaluators.
- Consistency in evaluation of clinical skills by nurse educators may have implications for clinical settings.
- Interrater reliability promotes clinical competence by nursing students.
- Competent performance of physical assessment skills may be linked to patient safety, improved quality of care, and improved patient outcomes.

FUTURE DIRECTIONS

Results of this pilot project provide several recommendations for related future projects:
- Evaluator training program
- Investigating interrater reliability among nurse educators grading a larger sample of real nursing students
- Evaluating student physical assessment skills during clinical placements
- Checklist refinement
- Developing electronic checklist and pilot study on use of electronic checklist on mobile devices

ACKNOWLEDGMENTS

Thank you to Dr. Rodney Hicks and Dr. Shirley Bristol for their guidance, assistance, and mentoring. Dr. Neil Patel is acknowledged for his kind help with statistical analyses. Many thanks to the nurse educators, simulated student, and simulated patient who participated in this study.

References available upon request.