Towards Investigating Performance Differences in Clinical Reasoning in a Technology Rich Learning Environment

Tenzin Doleck[1], Amanda Jarrell[1(✉)], Eric Poitras[2], and Susanne Lajoie[1]

[1] Department of Educational and Counselling Psychology, McGill University, Montreal, QC, Canada
{tenzin.doleck,amanda.jarrell}@mail.mcgill.ca, susanne.lajoie@mcgill.ca
[2] Department of Educational Psychology, University of Utah, Salt Lake City, UT, USA
eric.poitras@utah.edu

Abstract. Technology Rich Learning Environments (TREs) are increasingly used to support scholastic activities. BioWorld is an example of a TRE designed to support the metacognitive activities of learners tasked with solving virtual patient cases. The present paper aims to examine the performance differences of novice physicians in diagnosing cases in BioWorld. We present an empirically guided line of research concerning the performance differences: (1) across three endocrinology cases, (2) between genders, (3) between goal orientations, and (4) in diagnosis correctness.

Keywords: Technology rich learning environments · Clinical reasoning · Performance differences · Learner outcomes

1 Introduction

It is widely acknowledged that clinical reasoning, a process "that uses formal and informal thinking strategies to gather and analyse patient information, evaluate the significance of this information and weigh alternative actions" [1], is a crucial skill for medical students [2, 3]. However, current teaching and learning approaches have been found to be inadequate in developing the necessary clinical reasoning skills [4]. Thus, there is a need for affording learners opportunities to learn, practice, and improve their clinical reasoning skills. The educational technology literature is rife with studies documenting the efficacy of Technology Rich Learning Environments (TREs) in spawning positive learning outcomes in varied learning contexts [e.g. 5, 6, 7]. Thus, it is worthwhile to consider such learning systems to support and enrich learners. The present research concerns investigations of learner outcomes in a medical TRE designed to support learners in developing expertise in clinical reasoning. This paper lays the foundation for an empirically guided line of research concerning the performance differences: (1) across three endocrinology cases, (2) between genders, (3) between goal orientations, and (4) in diagnosis correctness.

C. Conati et al. (Eds.): AIED 2015, LNAI 9112, pp. 566–569, 2015.
DOI: 10.1007/978-3-319-19773-9_63

Examining performance differences in clinical reasoning can have implications for both instructional design and learner outcomes. In this paper, we outline the general methods, provide details about the data and data analysis, and present preliminary results for the first study, i.e., performance differences across the three patient cases of varying difficulty levels.

2 Methods

2.1 Participants and Study Procedure

Participants were recruited through advertisements (on classified website) and newsletter (via email). A total of 30 medical undergraduate students from a large Northeastern Canadian University volunteered to participate. The convenience sample comprised of 11 men and 19 women, with an average age of 23 (*SD* = 2.60). Participants were compensated $20 at the completion of a 2-hour study session.

Data for this project was collected as part of work that examined factors that influenced attention to feedback in BioWorld [8]. Initially, participants completed a training case, which provided an opportunity to familiarize with the BioWorld system. Upon completion of the training case, participants solved 3 endocrinology cases. The three endocrinology cases (**correct diagnosis for each case indicated in brackets*) participants were asked to solve were: Amy (*diabetes Type 1*), Cynthia (*pheochromocytoma*), and Susan Taylor (*hyperthyroidism*). The order of the cases was counterbalanced to mitigate practice effects. The difficulty levels of the cases were ascertained through a previous study [9]: the anticipated accuracies for the three cases were Amy (94%), Cynthia (33%), and Susan Taylor (78%).

2.2 Learning Environment: BioWorld

BioWorld [10, 11] is a technology rich learning environment (TRE) that supports learners in practicing clinical reasoning skills. BioWorld consists of four main spaces (Problem, Chart, Library, and Consult). The 'Problem' space provides the patients' case history. In solving the patient case, the learner reviews the patient summary and formulates a differential diagnosis (with the help of the Hypothesis Manager Tool), along with updating their level of confidence in relation to the most likely diagnosis (via the Belief Meter). In the Chart space, learners can review patient's vital signs and order lab-tests to confirm or disconfirm specific diagnosis. The Library and Consult serve as help-seeking tools. The final step in problem solving involves submitting a final diagnosis, sorting and prioritizing evidence, and writing a patient case summary.

2.3 Measures: Log-Files

The BioWorld system logs learner-system interactions in log-files. There are three different types of performance metrics in the system logs: *diagnostic efficacy* (e.g., percentage of matches with experts), *efficiency* (e.g., time taken to solve a case), and

affect (e.g., confidence). The specific information logged include: the *attempt identifier* (participant and case ID), a *timestamp*, the BioWorld *space* (e.g., chart), the specific *action taken* (e.g., add test), and *details in relation to the action* (e.g., Thyroid Stimulating Hormone (TSH) Result: 0.2 mU/L).

3 Data Analysis

3.1 Data Cleaning

A box plot analysis was conducted for each of the performance variables in order to identify potential outliers in the dataset; this analysis identified few outliers. To preserve the sample size, these extreme values were replaced using the next most extreme value within their corresponding case.

3.2 Performance Metrics

Performance metrics that characterize diagnostic accuracy and efficiency were extracted from the system logs. Within this study, we operationalize *Accuracy* by the number of evidence matches with the expert solution and define *Efficiency* by the total time taken to solve the case and the number of laboratory tests ordered.

3.3 Preliminary Results

We present preliminary results for performance differences across the three endocrinology cases of varying difficulty levels. A MANOVA analysis was conducted with the three indices of performance as dependent variables (i.e., number of lab tests ordered, number of correct matches with the expert solution, and time taken to solve the case) and the case with three levels as the independent variable. The results from the MANOVA suggest that there is a significant difference in the pattern of means between cases across participant performance indices, $F(6, 168) = 9.474$, $p < .001$, $\eta^2 = .25$. For comprehending the nature of these differences, a series of ANOVA post-hoc comparisons were conducted which revealed that there was a significant difference in number of matches with the expert solution between cases, $F(2, 86) = 18.39$, $p < .001$, $\eta^2 = .30$ and in how many lab tests were ordered, $F(2, 86) = 10.39$, $p < .001$, $\eta^2 = .20$. However, there were no significant differences in elapsed time across cases, $F(2, 86) = 4.18$, $p = .019$, $\eta^2 = .089$. To understand how the cases differed, Tukey HSD post-hoc comparisons were conducted; the results suggest that participants had significantly more matches with the expert solution for the Amy and Susan Taylor cases in comparison to the Cynthia case ($M = 6.39$ $SE = .35$, $M = 7.09$ $SE = .35$ and $M = 4.26$ $SD = .35$, respectively). The results also indicate that participants ordered significantly more lab tests for the Cynthia case in comparison to the Amy and Susan Taylor cases ($M = 13.77$ $SD = .95$, $M = 9.67$ $SD = .95$ and $M = 7.83$ $SD = .95$, respectively).

4 Concluding Remarks

In this paper, we have discussed ongoing research on empirically investigating the performance differences in clinical reasoning in a TRE for medical education. We presented the preliminary results obtained for our first research question: "Are there performance differences in the manner in which novices solve cases with varying levels of complexity?" Our next step is to complete the series of analyses to investigate performance differences (1) between genders, (2) between goal orientations, and (3) in diagnosis correctness.

References

1. Simmons, B.: Clinical reasoning: concept analysis. Journal of Advanced Nursing **66**(5), 1151–1158 (2010)
2. Delany, C., Golding, C.: Teaching clinical reasoning by making thinking visible: an action research project with allied health clinical educators. BMC Medical Education **14**(1), 20 (2014)
3. Norman, G.: Research in clinical reasoning: Past history and current trends. Medical Education **39**(4), 418–427 (2005)
4. Levett-Jones, T., Hoffman, K., Dempsey, J., Jeong, S., Noble, D., Norton, C., Roche, J., Hickey, N.: The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and manage clinically 'at risk' patients. Nurse Education Today **30**(6), 515–520 (2010)
5. Beal, C.R., Walles, R., Arroyo, I., Woolf, B.P.: On-line tutoring for math achievement testing: A controlled evaluation. Journal of Interactive Online Learning **6**(1), 43–55 (2007)
6. Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., William, W.C., Stylianides, G.J., Koedinger, K.R.: Cognitive anatomy of tutor learning: Lessons learned with SimStudent. Journal of Educational Psychology **105**(4), 1152–1163 (2013)
7. Vanlehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., Wintersgill, M.: The Andes physics tutoring system: Lessons learned. International Journal of Artificial Intelligence in Education **15**, 147–204 (2005)
8. Naismith, L.: Examining motivational and emotional influences on medical students' attention to feedback in a technology-rich environment for learning clinical reasoning (Unpublished doctoral dissertation). McGill University, Canada (2013)
9. Gauthier, G., Lajoie, P.S., Naismith, L., Wiseman, J.: Using expert decision maps to promote reflection and self-assessment in medical case-based instruction. In: Proceedings of Workshop on the Assessment & Feedback in Ill-Defined Domains at ITS, Montréal, Canada (2008)
10. Lajoie, S.P.: Developing professional expertise with a cognitive apprenticeship model: examples from avionics and medicine. In: Ericsson, K.A. (ed.) Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments, pp. 61–83. Cambridge University Press (2009)
11. Lajoie, S.P., Poitras, E.G., Doleck, T., Jarrell, A.: modeling metacognitive activities in medical problem-solving with bioworld. In: Peña-Ayala, A. (ed.) Metacognition: Fundamentals, Applications, and Trends, pp. 323–343. Springer (2015)