

January 2017

# Pblcloud Virtual Patient Simulator: Enhancing Immersion Through Natural Language Processing

Pierre Rodricus Martin  
*Yale University*, pierre.martin@yale.edu

Follow this and additional works at: <http://elischolar.library.yale.edu/ymtdl>

---

## Recommended Citation

Martin, Pierre Rodricus, "Pblcloud Virtual Patient Simulator: Enhancing Immersion Through Natural Language Processing" (2017). *Yale Medicine Thesis Digital Library*. 2151.  
<http://elischolar.library.yale.edu/ymtdl/2151>

This Open Access Thesis is brought to you for free and open access by the School of Medicine at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Medicine Thesis Digital Library by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact [elischolar@yale.edu](mailto:elischolar@yale.edu).

PBLCLOUD VIRTUAL PATIENT SIMULATOR: ENHANCING IMMERSION THROUGH NATURAL  
LANGUAGE PROCESSING

A Thesis Submitted to the  
Yale University School of Medicine  
in Partial Fulfillment of the Requirements for the  
Degree of Doctor of Medicine

by

PM

2017

## PBLCLOUD VIRTUAL PATIENT SIMULATOR: ENHANCING IMMERSION THROUGH NATURAL LANGUAGE PROCESSING

PM M.Ed.<sup>1</sup>, LD M.D.<sup>2</sup>, and TW M.D. M.P.H.<sup>2</sup>.

<sup>1</sup>Yale School of Medicine <sup>2</sup>From the Division of Critical Care Medicine, Department of Anesthesia, Perioperative and Pain Management, Boston Children's Hospital and the Department of Anesthesia, Harvard Medical School, Boston, MA, USA. (Sponsored by JH, Department of Pediatrics, Yale School of Medicine)

Virtual patient simulation has been utilized to teach interviewing skills, often employing selection-based methods (e.g., multiple-choice lists and menu-based prompts) to simulate doctor-patient conversations. Users have evaluated these systems as inauthentic, which can diminish user immersion (influenced by control, realism, distraction and sensory factors) and, in turn, negatively affect skill acquisition, mastery and transfer.

Our objectives were to design and develop PBLCloud, a scenario-based and highly interactive platform that uses natural language processing to support a more realistic doctor-patient conversation and create an immersive clinical learning environment.

PBLCloud was developed utilizing an iterative design thinking process and its initial evaluation involved a mixed methods approach. We recruited a convenience sample of 11 participants: three (27%) fourth-year medical students from Harvard Medical School as well as two (18%) residents, four (36%) fellows and two (18%) attendings from Boston Children's Hospital. There were two rounds of formative evaluation testing with eight participants in Round 1 and three participants in Round 2. Each participant completed a semi-structured think-aloud protocol exploring our pilot case, 10-item system usability scale (SUS) and 10-item open-ended questionnaire.

The chat-based functionality provides users with computer-generated context-specific responses during the historical encounter. Users have the opportunity to perform physical examinations, review incorporated multimedia, order and interpret diagnostic investigations, order therapeutic interventions that have appropriate effects on patient vitals and laboratory data, formulate and refine a differential diagnosis, receive just-in-time feedback regarding user-initiated actions and complete embedding learning exercises. 73% of participants strongly agreed that PBLCloud was useful (i.e., it is clinically-oriented, realistic, provides helpful feedback and is widely applicable) and 64% of participants strongly agreed that their experience with the system was enjoyable (i.e., it is relevant with an engaging interface). It was deemed to be more interactive and engaging than other simulators and 82% of participants were very interested in utilizing the system in the future. The average SUS score for Round 1 and 2 were  $79.7 \pm 12.0$  and  $82.5 \pm 19.8$  respectively. Areas of improvement were identified, in particular, the unsatisfactory response accuracy of the chat-based functionality.

Future work will include the investigation of various strategies to optimize the platform's natural language processing algorithm as well as the formal evaluation of the system's validity, reliability, level of induced user immersion and educational impact. We anticipate that PBLCloud will serve as a cost-effective and scalable approach for the instruction and assessment of clinical reasoning.

## **ACKNOWLEDGEMENTS**

First, I would like to thank Dr. TW and Dr. LD for their unwavering support and mentorship throughout the research process. It has been an invigorating learning opportunity to interface with both the physician leadership at OPENPediatrics and the Med2Lab team to design and develop the PBLCloud system. Moreover, this work would not be possible without the tireless efforts of the production specialists, medical animators and office managers at OPENPediatrics.

Secondly, I would like to thank Dr. JH whose insights regarding the research and practice of medical education were instrumental to the success of this project. Her perspectives will undoubtedly serve as an inspiration—a guiding light—throughout my future career in medical education.

Finally, I would like to thank my family, friends and past mentors who have shaped both my personal as well as professional development.

## TABLE OF CONTENTS

I.	INTRODUCTION-----	1
	a. Clinical Reasoning: Pedagogical and Technological Advances-----	1
	b. Virtual Patient Simulation: Definition, Use and Adoption-----	1
	c. Virtual Patient Simulation: Effectiveness-----	2
	d. Problem Statement: Unnatural Interactions and Immersion-----	3
II.	SPECIFIC AIMS-----	5
III.	METHODS-----	6
	a. Collaborators-----	6
	b. Design-based Research-----	6
	c. Design Thinking Process: Empathize-----	6
	d. Design Thinking Process: Define-----	7
	e. Design Thinking Process: Ideate-----	7
	f. Design Thinking Process: Prototype-----	8
	g. Design Thinking Process: Test-----	8
	i. Participants and Rounds-----	8
	ii. Data Collection-----	9
	iii. Qualitative: Thematic Analysis-----	9
	iv. Quantitative: Descriptive Statistics-----	10
	h. Platform Refinement-----	10
	i. Yale Medical Student Contribution-----	11
IV.	RESULTS-----	12
	a. Design and Development: Interface-----	12
	b. Design and Development: Case Assistant-----	15
	c. Design and Development: Historical Interview-----	15
	d. Design and Development: Physical Examination-----	16
	e. Design and Development: View Lab/Study Results-----	18
	f. Design and Development: Assessment - Differential Diagnosis-----	20

g.	Design and Development: Plan – Orders-----	21
h.	Design and Development: Hint System-----	22
i.	Design and Development: Interactive Test Questions-----	23
j.	Design and Development: Summative Feedback-----	24
k.	Usability Testing: Participant Demographics-----	26
l.	Usability Testing: Qualitative Analysis – Emergent Themes-----	28
m.	Usability Testing: Quantitative Analysis - System Usability Score-----	32
n.	Usability Testing: Quantitative Analysis - Bugs, Usability Issues and Content Errors---	32
V.	DISCUSSION-----	33
a.	Introduction to Immersion-----	33
b.	Importance of Immersion-----	33
c.	Factors that Influence Immersion-----	34
d.	PBLCloud: Distraction and Sensory Factors-----	34
e.	PBLCloud: Control and Realism Factors - Linearity and Interactivity-----	35
f.	Interviewing Skills-----	36
g.	Simulation of the Historical Encounter-----	36
h.	Natural Language Processing and Chat-Based Approaches-----	37
i.	Selection-Based Approaches vs. Chat-Based Approaches-----	37
j.	Usability Testing: Demographics and Themes-----	39
	i. Participant Demographics-----	39
	ii. PBLCloud and Usefulness-----	39
	iii. PBLCloud and Enjoyment-----	40
	iv. PBLCloud as Compared to other Virtual Patient Simulators-----	40
	v. PBLCloud and Areas for Improvement-----	41
	vi. PBLCloud and Participant’s Future Use-----	42
k.	Limitations-----	42
l.	Future Directions-----	43
	i. New Features and Psychometric Testing-----	43

ii.	Application of PBLCloud: OPENPediatrics' Mission-----	43
iii.	Application of PBLCloud: Global Community-----	44
iv.	Application of PBLCloud: Clinical Learning Pathways (CLPs) -----	45
v.	New PBLCloud Interfaces: Instructor Interface-----	46
vi.	New PBLCloud Interfaces: Administrator Interface-----	47
m.	Conclusion-----	47
VI.	REFERENECES-----	48
VII.	APPENDIX I: THINK ALOUD PROTOCOL-----	52
VIII.	APPENDIX II: PBLCLOUD SYSTEM USABILITY SCALE QUESTIONAIRRE-----	64
IX.	APPENDIX III: OPEN-ENDED QUESTIONAIRRE -----	66



## INTRODUCTION

### Clinical Reasoning: Pedagogical and Technological Advances

Clinical reasoning refers to the cognitive processes that underlie the collection, evaluation and synthesis of data used to formulate a patient's diagnosis and treatment plan.<sup>1</sup> In undergraduate medical education, the teaching and assessment of clinical reasoning is a Herculean task, due to factors both intrinsic and extrinsic to the individual learner, and its development, from novice to expert, has been rigorously studied for the past 50 years.<sup>2,3</sup> Theoretical considerations and empirical investigations from educational and cognitive psychology have helped medical education transition from teacher-centered and product-focused practices to strategies that are student-centered and process-focused; we now understand that "how" you learn is just as important as "what" you learn. These pedagogical advances have been paralleled by advances in Internet-based technology that enable the construction of innovative digital learning environments that support on-demand, personalized and active learning embodied by the e-, blended and social learning movements. Virtual patient simulation (VPS) is one educational modality borne out of the marriage between these recent pedagogical and technological advances.

### Virtual Patient Simulation: Definition, Use and Adoption

VPS is defined as "a specific type of computer-based program that simulates real life clinical scenarios; users must emulate the role of the healthcare provider in order to obtain a history, conduct a physical exam, order diagnostic investigations and initiate therapeutic interventions".<sup>4</sup>

Shorter lengths of stay, higher patient acuity, increasing patient-to-staff ratios and duty-hour restrictions all contribute to the reduction in clinical exposure of medical students and simultaneously inhibit the ability of clinical educators to fulfill their teaching obligations, in addition to their clinical, research and administrative responsibilities.<sup>5</sup> Thus, clinical educators have turned to VPS, which supports the deliberate practice<sup>6</sup> of routine as well as low-volume, high-risk clinical scenarios in an environment that is psychologically safe for students.<sup>7</sup> The flexibility of these systems allows the structure and content of cases to be adapted to a wide range of specialties, conditions and levels of expertise.

Overall, VPS has generally been well received by students, particularly systems that are comprehensive and well organized.<sup>7</sup> In 2007, Huang et al. demonstrated that approximately 24% of medical schools have implemented

VPS into their curriculum.<sup>8</sup> As such, its adoption has been steadily increasing and it has been adapted for training in other healthcare professions<sup>9-11</sup>, nationally and internationally<sup>8,12</sup>. VPS has been studied for more than 50 years and its status as a viable training tool was underscored by the Liaison Committee on Medical Education, which explained “the medical student should be able to remedy gaps [in clinical exposure] by simulated experience”.<sup>13</sup>

#### Virtual Patient Simulation: Effectiveness

When compared to no interventions, VPS is associated with large positive effects with respect to knowledge outcomes and clinical reasoning.<sup>14</sup> Though when compared to non-computer interventions (e.g., lectures, paper instruction, slide-tape instruction, standardized patients), VPS confers only small positive effects.<sup>14</sup> Importantly, multiple studies have demonstrated that VPS doesn't address the same knowledge, skills and attitudes as other educational modalities. This is evidenced by the minimal correlation between a student's performance in VPS and their performance on multiple-choice examinations, with standardized patients (SP) or in real patient encounters.<sup>15</sup> VPS has been utilized to teach and assess interviewing skills as well as higher-order cognitive skills such as critical thinking, decision-making, and problem solving.<sup>1</sup> With this in mind, Cook et al. emphasized that VPS is ideally suited for the instruction and evaluation of clinical reasoning.<sup>1</sup> Its value as a valid and reliable method for large-scale high-stakes assessment was solidified with its inclusion in the USMLE™ Step 3 in 1999.<sup>16</sup> Undoubtedly, VPS will not replace other educational modalities, but will serve as a powerful complement and its integration with other modalities will require thoughtful coordination.

### Problem Statement: Unnatural Interactions and Immersion

Despite the significant opportunity that VPS systems offer, there exists considerable heterogeneity in the purpose, design, implementation and effectiveness of existing VPS systems. This variation is a byproduct of the paucity of theories, guidelines and standards available to support practitioners in both instructional design and curricular integration.<sup>17</sup>

Another dimension by which VPS systems can be evaluated is the extent to which they induce and sustain user immersion. Immersion is a psychological state in which an individual is engaged and engrossed in another reality<sup>18</sup>, in this case, the virtual learning environment. Studies throughout the education literature suggest that immersion is important for the learning process and any reduction can have significant implications for skill acquisition, mastery and transfer.<sup>19</sup> Witmer et al. provide a theoretical framework that outlines the various factors that affect immersion and categories them into the following factors: distraction, sensory, control and realism.<sup>18</sup> Limited distractions, multimodal stimulation, increased interactivity, compelling narratives and permissible user-initiated actions that are natural, unrestrained and have meaningful/predictable consequences all contribute to a user's sense of immersion.

Existing VPS systems have focused much effort towards designing simple and intuitive interfaces in an attempt to reduce distractions and help users focus their limited mental resources on the most relevant information in the case. This relevant information is typically in the form of rich multimedia demonstrating key features of the patient's presentation. However, many of the widely adopted VPS systems, such as CASUS<sup>20</sup> and CAMPUS<sup>21</sup> (Case-based Training in Medicine as part of Problem-oriented Educational Strategy), are characterized as linear ("string of pearls") systems and remain quite constrained in terms of user interactivity. For example, most VPS systems employ selection-based approaches (e.g., multiple-choice lists, menu-based prompts and forced branching) to simulate the doctor-patient conversation. Users have consistently evaluated these systems as inauthentic due to the unnatural interaction between doctor and patient; this detracts from user immersion. One approach to enhancing the user's immersion is to target control and realism factors, which involve the user having natural interactions with the system/patient that have meaningful and predictable consequences based on experience in the real world.<sup>18</sup>

Here we describe the design, development and initial usability testing of Problem-Based Learning Cloud (PBLCloud), a highly interactive and scenario-based platform for the instruction and assessment of clinical reasoning. We aimed to focus on increasing user interactivity, and by extension user immersion, via a chat-based

approach to the simulation of the historical encounter. This functionality is supported by natural language processing (NLP) that allows the system to 1) understand free-text queries from the user and 2) provide users with accurate context-specific responses, ultimately approximating a realistic conversation between doctor and patient. Interactivity will be further enhanced as users perform physical examinations, order and interpret diagnostic investigations, order therapeutic interventions that have appropriate and observable effect on the patient's vitals and laboratory data, formulate and refine a differential diagnosis, receive just-in-time feedback regarding user-initiated actions, complete embedded learning exercises and are assessed based on specific criteria (i.e., diagnosis, overall management and interactive test questions).

### **SPECIFIC AIMS**

1. To design and develop an interactive scenario-based platform, PBLCloud, that utilizes natural language processing to support a chat-based approach for the simulation of the doctor-patient conversation and enhance user immersion.
2. To conduct, via a mixed methods approach, a formative evaluation of the PBLCloud platform in order to refine the content and functionality of the system for its target audience, medical students.
3. To describe the future application of the PBLCloud platform as a tool to support the instruction and assessment of clinical reasoning.

## METHODS

### Collaborators

PBLCloud represents the product of the collaboration between two organizations, OPENPediatrics and Med2Lab. OPENPediatrics is a knowledge-sharing platform that is Internet-based, open-access, peer-reviewed and not-for-profit.<sup>22</sup> The purpose of the organization is to leverage existing technology for the creation of a global community of practice to improve the care of critically ill children. Med2Lab is a software development team based out of Washington, D.C. that created the proprietary NLP algorithm behind the chat-based functionality of PBLCloud's historical encounter and implements the computer programming for the user interface and overall functionality.

### Design-based Research

Our team used a design thinking process<sup>23</sup> for the development of PBLCloud. This approach helps designers define a problem and implement specific solutions that take into account the needs of the user demographic. It is an iterative process that involves transitioning through the following stages: empathize, define, ideate, prototype, test and refine. The following section provides a summative review of our team's design thinking process.

### Design Thinking Process: Empathize

The first step of the design thinking process involves working to understand the needs of potential end-users. Medical students are the intended user demographic for PBLCloud. Before developing the system, we conducted a review of the VPS literature, focusing on articles detailing the design, development, psychometric analysis and educational impact of published VPS systems. These insights were supplemented by the medical education experience of the authors.

### Design Thinking Process: Define

Broadly, our search revealed that medical students appreciate comprehensive, well-organized and intuitive systems that focus on clinical reasoning, offer just-in-time feedback for user-initiated actions, maintain authenticity (i.e., content and multimedia) and that are relevant in terms of the type, quality and quantity of information provided.<sup>7,24-27</sup> Conversely, many of the current VPS systems are perceived as linear, maintain limited interactivity and include unnatural interactions between doctor and patient during the historical encounter.<sup>28,29</sup> As a result, users evaluate these systems as inauthentic, which negatively impacts the user's immersion into the virtual learning environment and, in turn, their potential educational gains.

### Design Thinking Process: Ideate

Ideation involves an exploration of possible solutions for the problem statement above. Among the many features included in PBLCloud, our unique approach was to increase user interactivity and support more natural interactions between the doctor and patient. Specifically, our main aim was to develop a chat-based approach to simulate a more realistic doctor-patient conversation. Such an approach would allow users to free-text their historical interview questions and Med2Lab's NLP algorithm would provide users with computer generated context-specific responses. This could serve as a resource to instruct and assess history-taking skills around what, when, how and why questions were posited by users. This assessment would be an indirect measure of clinical reasoning to be synthesized along with the user's decisions/actions evaluated throughout the remainder of the case.

In addition, we endeavored to include a comprehensive set of orders (i.e., investigations and interventions) that would have meaningful effects on patient vital signs and laboratory values, parameters that update based on both natural disease progression and user-initiated actions. This interactivity allows users to progress through the case in a free-form manner. A hint system offers just-in-time feedback for these user-initiated actions. Finally, multiple embedded learning exercises (e.g., diagnostic formulation and interactive test questions) were included, along with a scoring mechanism to assess the user's clinical reasoning based on diagnosis formulation, initiated orders and responses to interactive test questions.

Similar to extant systems and in keeping with the needs/priorities of medical students, we endeavored to create a system with limited distractions, by providing minimal text on screen as well as implementing an engaging layout and intuitive functionality. These strategies have been utilized to help users focus on relevant case information and multimedia.

#### Design Thinking Process: Prototype

We utilized a suite of web-based collaboration tools to support system design, development and editing. Initial interface mock-ups were developed in Adobe®Photoshop®, functionality was demonstrated via interactive prototypes created in InVision, and the result was incorporated into HTML code for final evaluation.

#### Design Thinking Process: Test

After completing the initial version of the PBLCloud platform, we conducted usability testing to gather user feedback and optimize the system. The authors utilized a mixed methods approach<sup>30</sup> during this phase of the design thinking process, which involved the collection, analysis and integration of qualitative as well as quantitative data generated by the participants.

#### *Participants and Rounds*

For usability testing, researchers have demonstrated that the recruitment of 5 - 10 participants allows for the identification of approximately 85 – 95% of existing issues with a prototype.<sup>31</sup> And if you are testing multiple groups, recruiting 3 – 4 participants per round is generally sufficient.<sup>32</sup> Ultimately for this study, two rounds of formative evaluation were completed; we recruited eight participants in the first round and three participants in the second round.

Specifically, we recruited a convenience sample of 11 participants that included three (27%) fourth-year medical students from Harvard Medical School, two (18%) pediatric residents from the Boston Combined Residency Program, four (36%) fellows, from the Pediatric Critical Care Medicine Fellowship Program at Boston



Children's Hospital/Harvard Medical School and two (18%) attending pediatric physicians from Boston Children's Hospital.

#### *Data Collection*

Data collection for each participant occurred in three parts and included a 1-hour think-aloud protocol<sup>33</sup> conducted in-person by one of the authors (LD or PM), completion of a 10-item system usability scale (SUS) questionnaire<sup>34,35</sup>, and 10-item open-ended questionnaire. Demographic information was collected from participants including: gender, specialty, years in training, and expertise with computer- and web-based technologies. Afterwards, participants explored our pilot case (i.e., 2-year-old male with croup who presents with fever and respiratory distress). The think-aloud protocol included free exploration of the system by the user as well as directed questions from the facilitator regarding observations, problems experienced and recommendations for the system. After finishing the case, participants completed the SUS. The SUS is a 10-item attitude Likert scale that has been demonstrated to be both a simple and reliable tool that measures the global view of an individual's assessment of a system's usability. More specifically, this questionnaire measures usability, learnability, and user satisfaction. Finally, participants completed a series of 10 open-ended questions designed to gain a better understanding of each participant's experience with the system; this provided critical subjective information. The aforementioned materials are available for review in Appendices I, II and III.

#### *Qualitative Analysis: Thematic Analysis*

After completion of the usability sessions, participant responses taken from the authors' notes during the think-aloud protocol and open-ended questionnaire were investigated via thematic content analysis<sup>36</sup> to pinpoint, examine and record themes that emerged from the data. These themes were identified through a six-phase process: initial review of the data in its entirety, coding of the responses to collapse the data into categories for a more high-level and efficient analysis, synthesis of categories into overarching themes, exploration of the relationship between themes and the theoretical perspective, definition of the relationships and reporting of results. Microsoft® Excel® for Mac was utilized to facilitate coding and thematic exploration.<sup>37</sup> The proportion of participants that expressed

each theme was calculated and reported along with representative excerpts of the participant responses.

### *Quantitative Analysis: Descriptive Statistics*

An important adjunct to the qualitative analysis above was the use of descriptive statistics to characterize and emphasize certain successes and areas of improvement in the system's interface and functionality. For the three Likert-type items (usefulness, user enjoyment and future use) in the open-ended questionnaire the number of participants designating each response level was calculated and reported. The median response for each Likert-type item was reported along with the corresponding proportion of participants for that response level.

Moreover, the results from the SUS questionnaire for each end-user were recorded and the mean SUS score ( $\pm$  standard deviation) was reported for both Round 1 and Round 2.

In addition, problems experienced by users during the think-aloud protocol testing were recorded and coded using a preexisting framework<sup>33</sup>, separating problems into usability issues, bugs or content errors. Usability issues encompass aspects of the system or a demand on the user that make it onerous, inefficient or unpleasant for the user to achieve a specific goal in a typical usage situation.<sup>38</sup> In this context, bugs are a special type of usability issue, which represent a critical issue with the program's source code, design or paired operating system.<sup>39</sup> Lastly, content errors involve identified aspects of the case that misrepresent the clinical environment, diagnostics, management and/or consequences of user-initiated actions; this is based on empirical research, consensus management guidelines and expert knowledge. The number of usability issues, bugs and content errors were reported for both Round 1 and Round 2.

### Platform Refinement

The interface and functionality of the system were continuously updated after each session, based on both the insights gathered during the usability testing as well as expert review. Critical bugs were addressed immediately upon discovery. Usability issues and content errors were typically addressed at the end of each round, in keeping with the spirit of grounded theory<sup>40</sup>, so that the authors could receive comprehensive feedback from all levels of users before proceeding.

More participants are currently being recruited for a third round of usability testing.

#### Yale Medical Student Contribution

From May 2015 to May 2016 PM worked with OPENPediatrics as an Associate Medical Editor and Project Coordinator. PM collaborated with the remaining authors (LD and TW) and Med2Lab to design and develop PBLCloud, including both the content and layout of the platform. With feedback from LD and TW, PM authored the pilot case and adapted the think-aloud protocol, SUS questionnaire and 10-item open-ended questionnaire for usability testing. PM conducted the first 4 usability testing sessions and completed the qualitative and quantitative analysis of all participant responses.

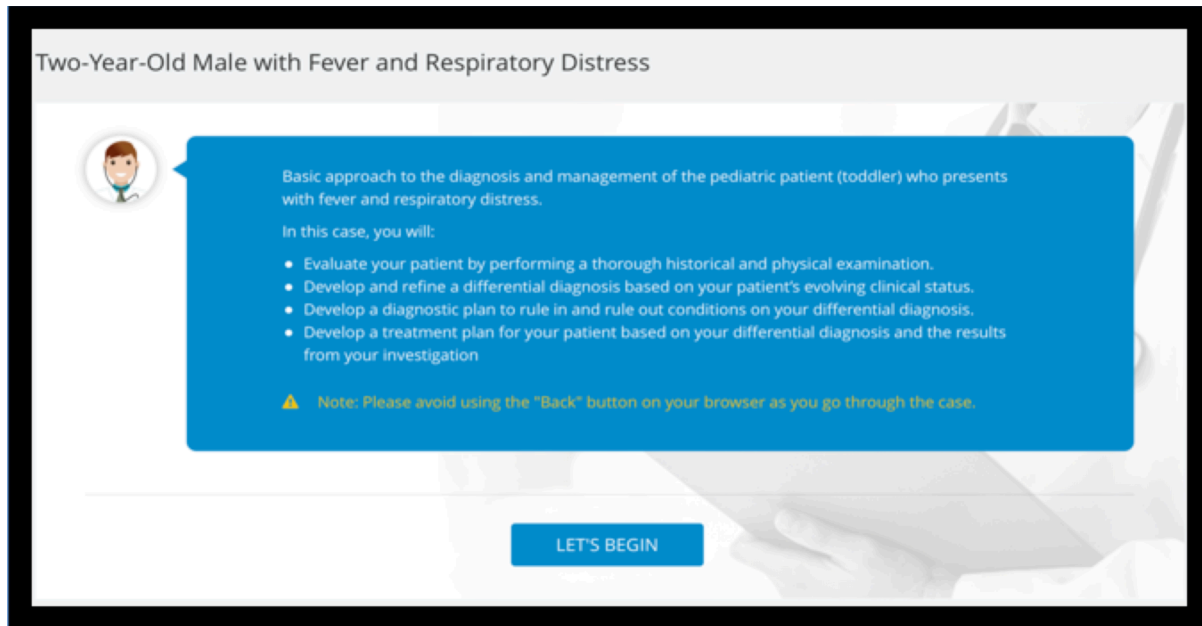
## RESULTS

The following section details 1) the final product of the authors' design thinking process with respect to PBLCloud's interface and functionality and 2) the formative evaluation of the system.


### Design and Development: Interface

PBLCloud is based on the HEIDR model of VPS<sup>41</sup> and allows users to obtain a clinical history, conduct a physical examination, order investigations, view diagnostic results, record a differential diagnosis and order therapies, ultimately mirroring the actions involved in a real patient encounter. Importantly, problem based learning is a student-centered pedagogy, which involves the utilization of carefully designed clinical scenarios that resemble real world problems, often open-ended and ill-structured.<sup>42</sup> As students solve these complex and authentic problems they will learn not only content knowledge, but also hone higher order processes such as critical thinking and problem solving.

At the start of the case, a pop-up window appears, which includes the case title, a short case description and a list of general instructor-derived objectives (Figure 1). Afterwards, the user will enter the Home Screen (Figure 2) that users will return to throughout a case. On the left side of the screen the user can view an image of the patient reflecting his/her current clinical status and review the patient's vital signs. During other sections of the encounter (e.g., physical examination), users can refer to the patient's chart by clicking on the "View Patient's Chart" button. On the right side of the screen users are able to freely navigate between conducting a historical interview, reviewing physical examination results and reviewing lab/study results.




Two-Year-Old Male with Fever and Respiratory Distress

 Basic approach to the diagnosis and management of the pediatric patient (toddler) who presents with fever and respiratory distress.

In this case, you will:

- Evaluate your patient by performing a thorough historical and physical examination.
- Develop and refine a differential diagnosis based on your patient's evolving clinical status.
- Develop a diagnostic plan to rule in and rule out conditions on your differential diagnosis.
- Develop a treatment plan for your patient based on your differential diagnosis and the results from your investigation

 **Note:** Please avoid using the "Back" button on your browser as you go through the case.


LET'S BEGIN

Figure 1: An example of the pop-up window that appears at the beginning of the case to introduce the user to the case title, a short case description and the instructor-derived objectives.

Two-Year-Old Male with Fever and Respiratory Distress Reset the Case

**CASE TIMELINE**

Presentation 20 mins 1 hr 4 hr



[View Patient's Chart](#)

**VITAL SIGNS**

Vital Sign	Value
Temperature	38.9 °C
Heart Rate	140 beats/min
Systolic Blood Pressure	110 mmHg
Diastolic Blood Pressure	70 mmHg
Respiratory Rate	35 breaths/min
O <sub>2</sub> saturation	98 %

**Obtain History** **Physical Exam** **Assessment and Plan** **View Labs and Imaging Results**

**Patient Information** **Ask Mother beta**

**Chief Complaint:** fever and difficulty breathing

**History of Present Illness:**

2-year-old male with fever to 101°F (38.5°C), rhinorrhea, and cough x 2-3 days duration. Over the past 24 hours his mother reports he started having "difficulty breathing." She has also noticed his cough has become harsher and that his breathing sounds very noisy, particularly when he breathes in. When crying, his voice sounds hoarse. His mother has noted coughing fits occurring more frequently at night, and that both his cough and breathing appear worse when he is upset as opposed to when he is calm. His cough is dry and hacking (non-productive of mucus). She has not noticed any drooling. She thinks he also has a sore throat. She has given acetaminophen (Tylenol) to treat the fever and pain but not sure it has helped. She reports he has been eating and drinking less than usual over the past 2 days and has been having a decreased number of wet diapers. She reports some children at the daycare he attends have had similar upper respiratory tract symptoms and fever. She denies any obvious foreign body ingestion.

**Past Medical History:** No prior hospitalizations or surgeries.

**Birth History:** Born at term via spontaneous vaginal delivery. Uncomplicated pregnancy and delivery.

Figure 2: An example of the Home Screen demonstrating the layout of the patient's image and vital signs on the left as well as the case timeline and navigational panel on the top right. The bottom right side of the page is reserved for user-initiated actions throughout the various sections of the clinical encounter. The patient's image and vital signs update appropriately based on the patient's clinical status and user-initiated actions.

### Design and Development: Case Assistant

Secondary to user confusion regarding navigation in the PBLCloud platform, a case assistant was incorporated to guide users throughout an individual case (Figure 3).

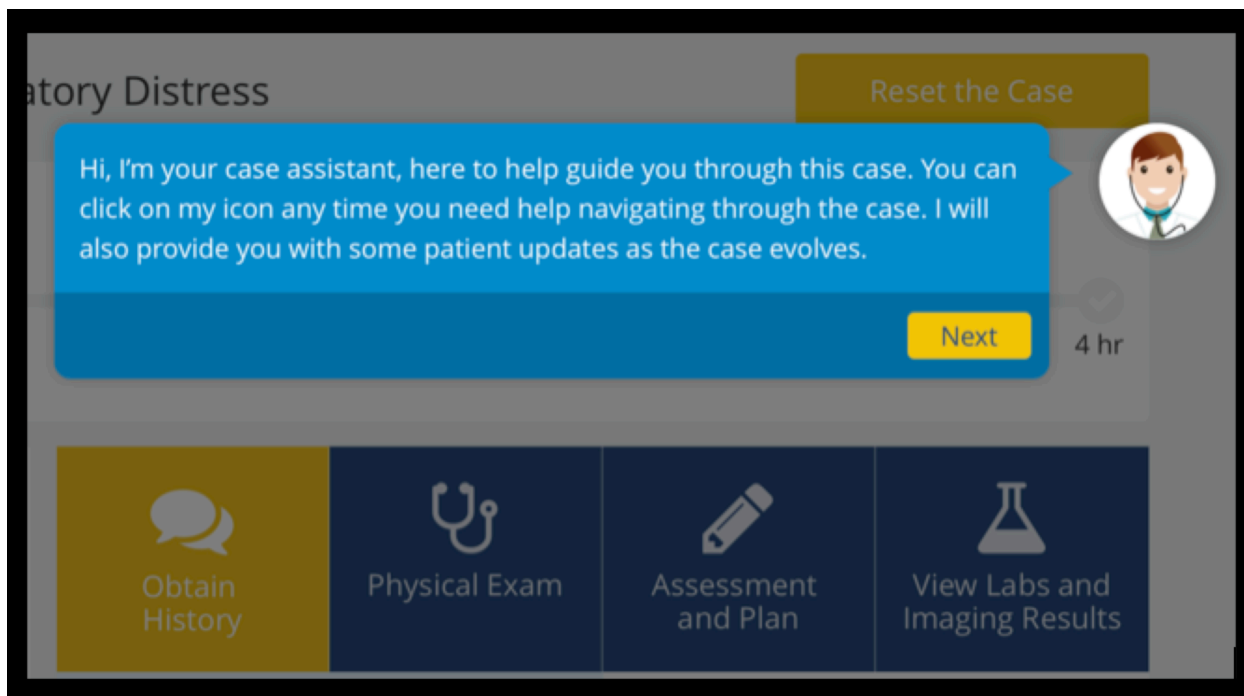


Figure 3: The case assistant provides notifications throughout the case in order to support navigation and provide information to help the user advance through the case.

### Design and Development: Historical Interview

Users are able to enter free-text queries for the patient and the NLP algorithm provides users with context-specific responses (Figure 4). The system will include default responses to circumvent unanticipated questions and basic socio-emotional modeling<sup>43,44</sup> will allow targeted responses for unwanted, unnecessary or repeat questions. The history-taking feature was identified to be frustrating during usability testing (described later), thus the clinical history is now provided under “Patient Information” in the current iteration of the system (Figure 2).

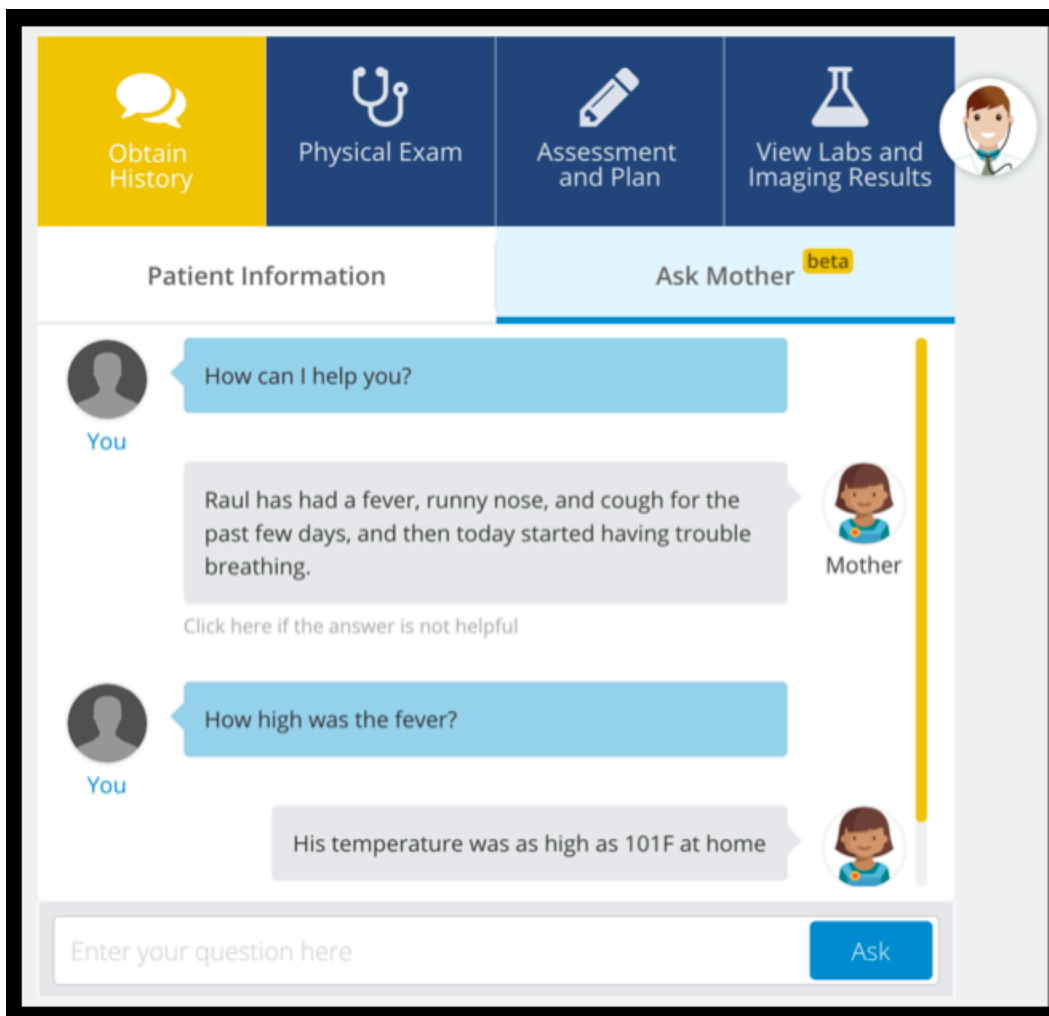


Figure 4: To obtain patient history, users submit free-text inquiries and receive computer-generated context-specific responses using this SMS-like (short messenger service) interface under “Ask Mother<sup>beta</sup>”.

#### Design and Development: Physical Examination

The user is presented with an interactive pediatric body diagram (Figure 5) of appropriate age, gender and ethnicity for the particular case. The user is then prompted to choose from thirteen organ systems to evaluate the patient for normal/abnormal findings. Findings are displayed in either text format or multimedia (e.g., audio, images, animations, live-action video) for interpretation. For example, in the pilot case of a 2-year-old with croup, users can choose the “Pulmonary Exam” which triggers an animation of a toddler in respiratory distress (Figure 6).



**Chief Complaint:** fever and difficulty breathing

**History of Present Illness:**  
 2-year-old male with fever to 101°F (38.5°C), rhinorrhea, and cough x 2-3 days duration. Over the past 24 hours his mother reports he started having "difficulty breathing." She has also noticed his cough has become harsher and that his breathing sounds very noisy, particularly when he breathes in. When crying, his voice sounds hoarse. His mother has noted coughing fits occurring more frequently at night, and that both his cough and breathing appear

View Patient's Image

Obtain History
Physical Exam
Assessment and Plan
View Labs and Imaging Results

CLICK on different parts on the body to view the exam findings



**VITAL SIGNS**

Vital Sign	Value
Temperature	38.9 °C
Heart Rate	140 beats/min
Systolic Blood Pressure	110 mmHg
Diastolic Blood Pressure	70 mmHg
Respiratory Rate	35 breaths/min
O <sub>2</sub> saturation	98 %

Figure 5: Users can explore findings (via text or multimedia) from the patient's physical examination using an interactive body diagram. Each node represents a different organ system.

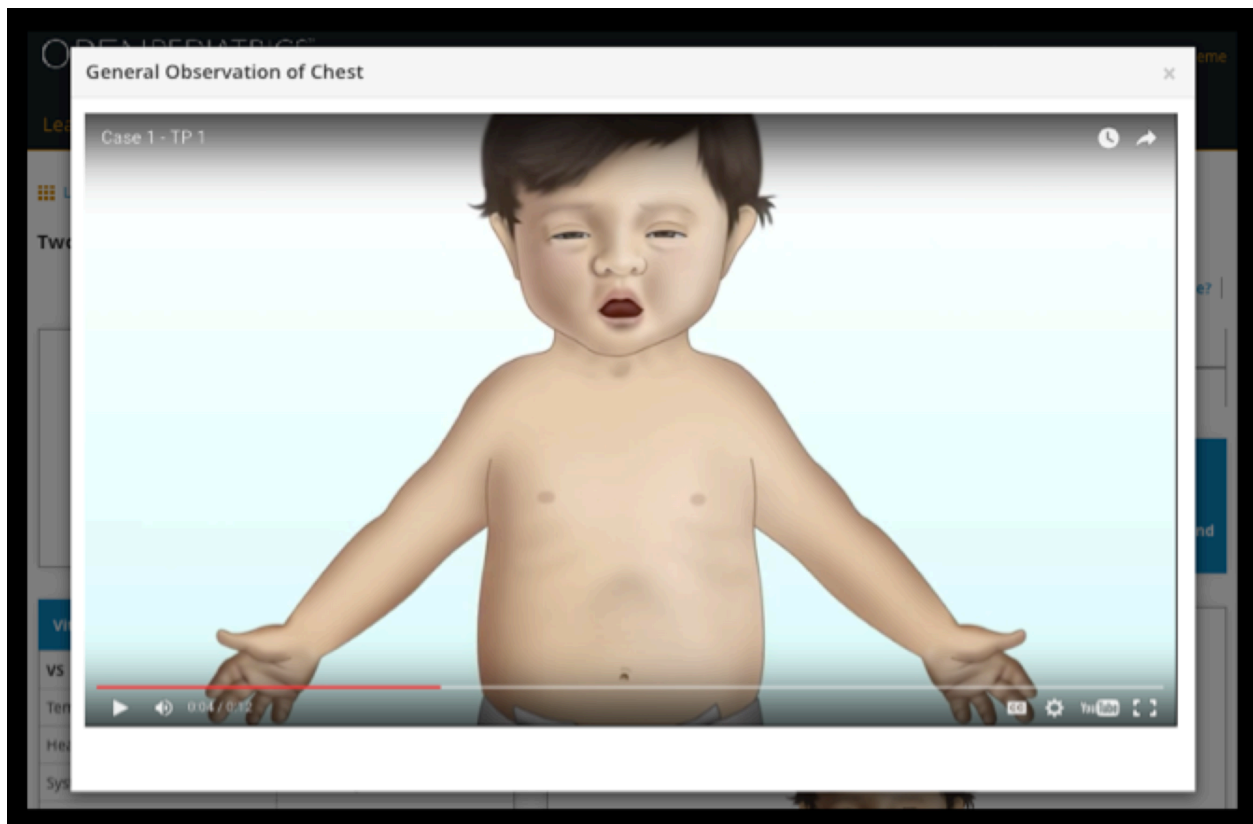


Figure 6: Image from an animation of 2-year-old male who presents with fever and respiratory distress as evidenced by a pale face, retractions, stridor and a brassy cough.

#### Design and Development: View Lab/Study Results

Users can refer to the results of diagnostic investigations (e.g., labs, imaging or others) as they become available (Figure 7). For imaging and other studies, users will be able to first review the data to practice independent interpretation of the findings and then compare their answers to that of an expert (e.g., radiologist's, electrophysiologist's or pathologist's report).

The screenshot displays a clinical interface with a top navigation bar containing four buttons: 'Obtain History', 'Physical Exam', 'Assessment and Plan', and 'View Labs and Imaging Results'. Below the navigation bar, there are two tabs: 'LABS' and 'IMAGING'. The 'LABS' tab is active, showing a 'Basic Metabolic Panel' table. The table lists various laboratory tests and their results. A blue tooltip is shown over the BUN result, displaying the normal range: '5 mg/dL - 18 mg/dL'.

Lab Panel	20 mins	
Na	135 mmol/L	
K	3.6 mEq/L	
Cl	103 mEq/L	
HCO <sub>3</sub>	23 mEq/L	
BUN	16 mg/dL	5 mg/dL - 18 mg/dL
Cr	0.4 mg/dL	
Glucose	95 mg/dL	
Ca	9.1 mg/dL	
Mg	1.8 mEq/L	
Phos	5.1 mg/dL	

Figure 7: Users can review the data from laboratory studies in order to interpret the findings and integrate the information with other clinical data to formulate a patient's assessment. Normal values are given in a blue tab when the user hovers over the results of a laboratory value.

### Design and Development: Assessment - Differential Diagnosis

After gathering all the patient information that they deem appropriate, the user can select “Assessment and Plan” and would then be prompted to record their differential diagnosis and input their orders (i.e., investigations, interventions and/or disposition). During the differential diagnosis activity, users can choose from a list of 15 diagnoses on the extended differential for their patient’s condition and create a prioritized differential diagnosis. Importantly, users can refer to their differential diagnosis when inputting their orders for the patient.

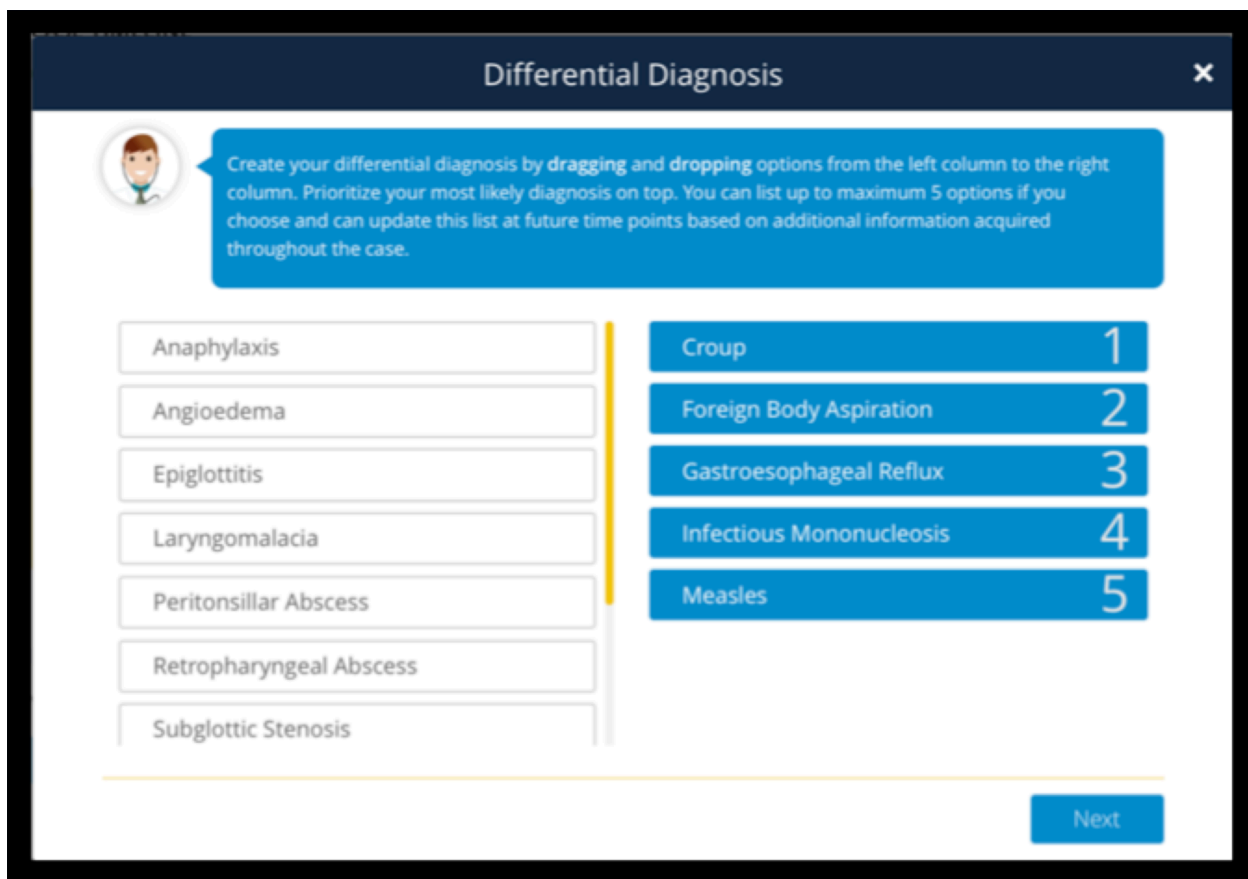


Figure 8: Users can input, prioritize and update their patient’s differential diagnosis based on synthesized clinical data.

### Design and Development: Plan - Orders

Afterwards, users can select either investigations (e.g., labs, imaging, and other studies) and/or interventions (e.g., medications, procedures and disposition) from a list of available orders (Figure 9). Using a rule-based system, each of these interventions are linked to appropriate changes in the patient's vitals signs and laboratory values. Thus, when users initiate certain investigations or interventions patient parameters change accordingly.

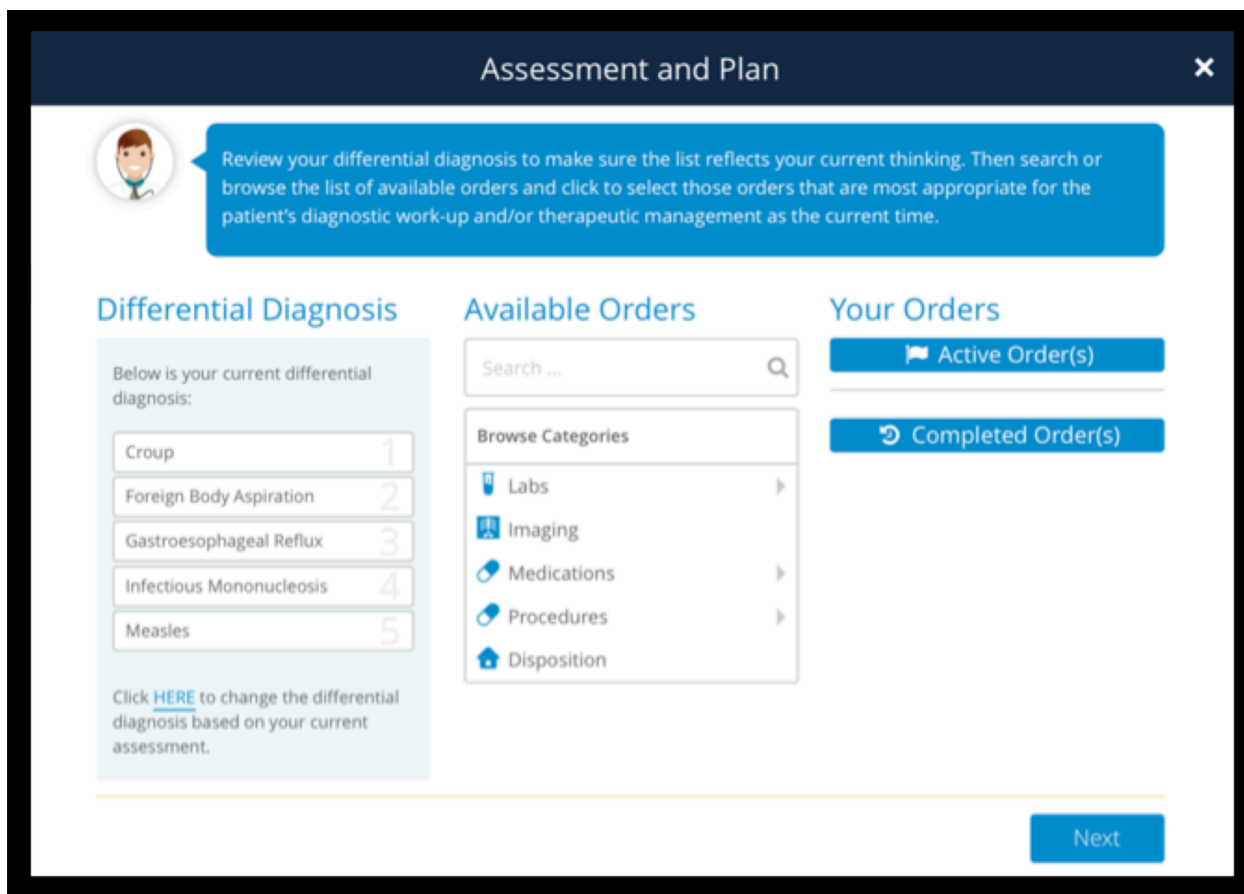


Figure 9: Users can input their orders for investigations and/or interventions (e.g., pharmacological or procedural). The system is able to track a user's order history, recording the specific order, time entered and its current status (i.e., active or completed).

### Design and Development: Hint System

Throughout the case, users receive just-in-time formative feedback relating to their orders. If users do not complete a green (indicated) action by a certain time point or if they initiate a red (non-indicated and harmful) action, then a pop-up window (Figure 10) will appear on the screen to urge the user to reevaluate the patient and reconsider their decisions. Moreover, users will receive notifications if abnormal vitals signs and critical lab values remain unaddressed.

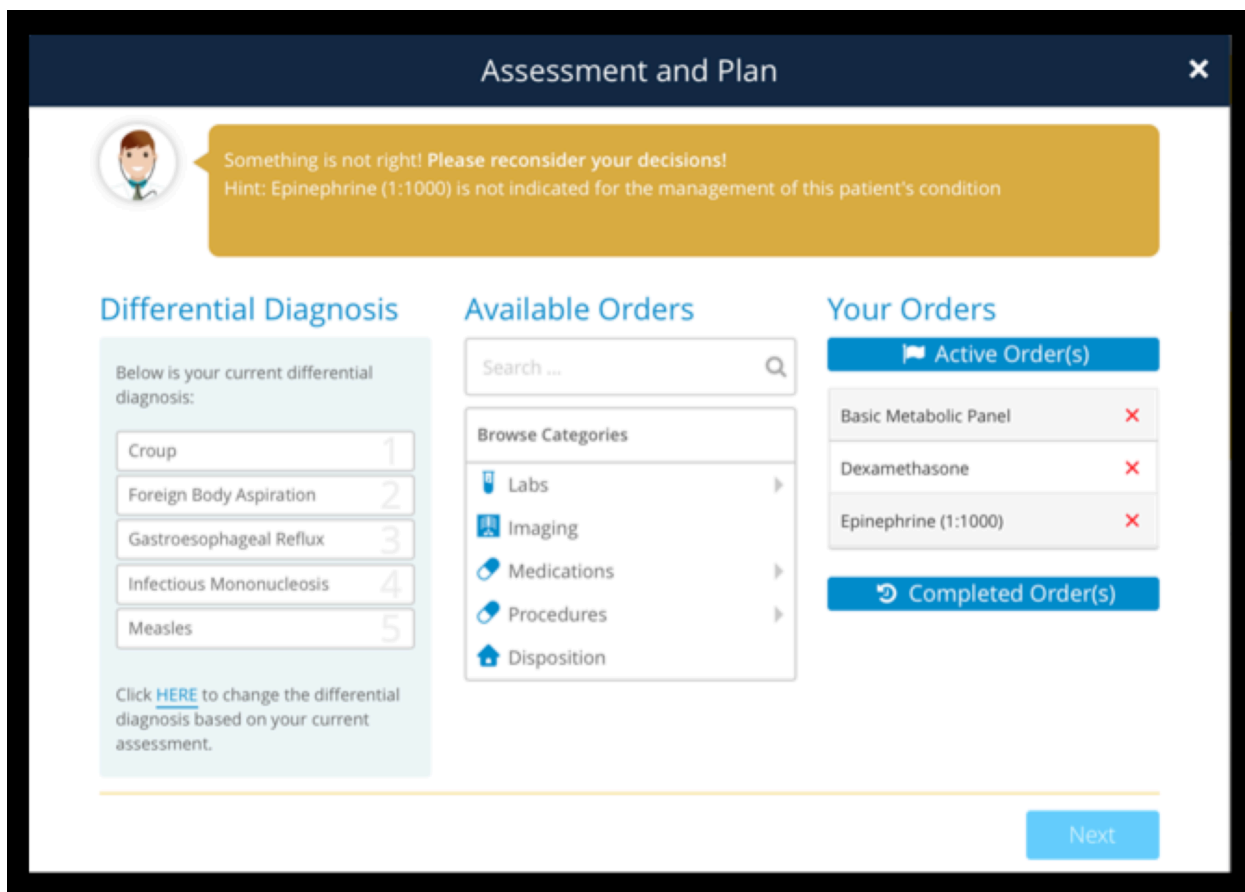


Figure 10: Users receive just-in-time feedback in the form of notifications and explanations about their actions and/or inactions (top-center of the pop-up window).

### Design and Development: Interactive Test Questions

At any point during the case, the user may encounter embedded multiple-choice questions (e.g., single and multi-select) that must be completed before moving forward in the case (Figure 11). Each question is linked to corresponding feedback and multimedia (e.g., video, audio, PDFs, hyperlinks) for self-directed learning. These exercises are designed to be activities to facilitate user reflection on the target knowledge, skills or attitudes of the given case.

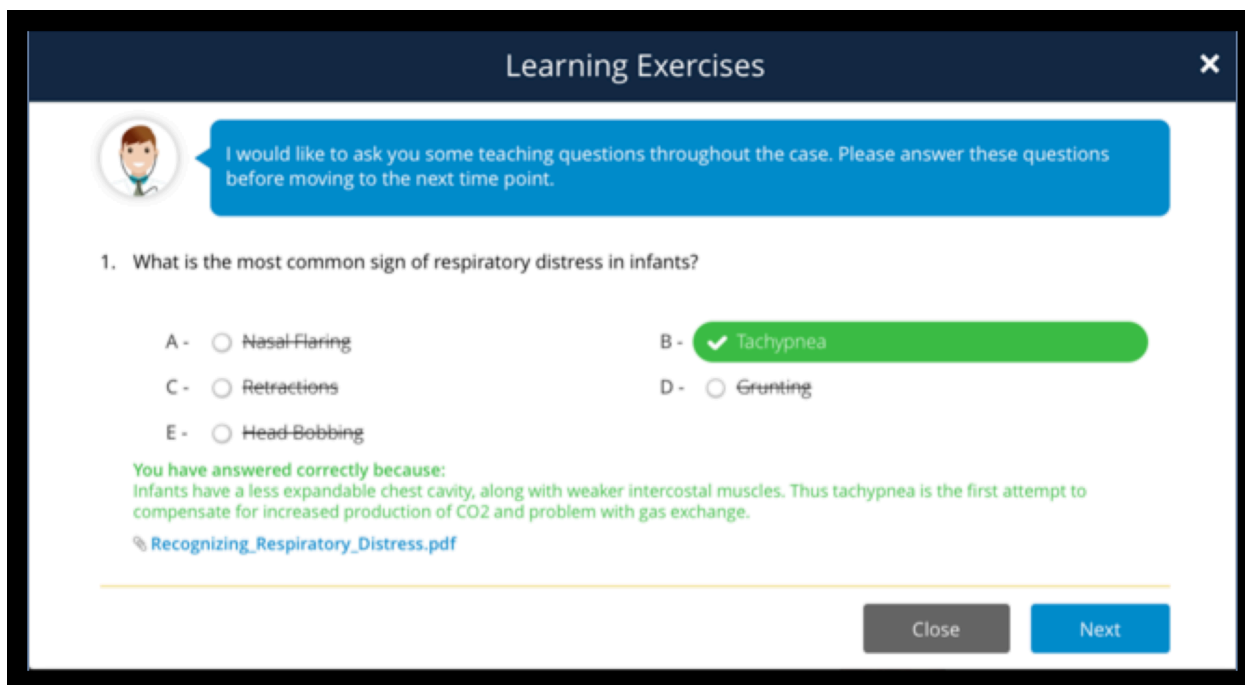


Figure 11: Interactive test questions can be included to further assess learning throughout the case. Immediate feedback is provided for users, along with links to additional resources to encourage self-directed learning.

### Design and Development: Summative Feedback

PBLCloud's cumulative scoring system grades the user on his/her differential diagnosis, orders (i.e., investigations and interventions) and responses to interactive test questions (Figure 11). Scoring for the diagnosis activity is based on the presence and appropriate priority level of the leading diagnosis, which is designated by the author of the case (Figure 12). The authors are considering other scoring mechanisms that would assess the entire 5-item differential at each time point, in order to track diagnosis formulation and refinement. The scoring system for user-initiated orders is based on the scoring system from the Computer Based Case Simulation of the USMLE™ Step 3<sup>45,46</sup> and it separates actions of the user into three categories: green (indicated, beneficial), yellow (non-indicated, non-risky) and red (non-indicated, harmful) actions. These distinctions are determined by the case-author during the case construction process and should be based on a combination of consensus guidelines and local practice (Figure 13). Finally, interactive test questions are scored as either correct or incorrect responses (Figure 14). Depending on the objectives of a particular case, these sub-scores (i.e., diagnosis, orders, interactive test questions) can be weighted differently in order to emphasize certain skills such as history-taking or diagnosis formulation.

In addition to reviewing feedback for their specific actions, users will be provided case-specific objectives, a diagnosis summary and a cumulative case summary to review the takeaways and the reasoning behind key elements of the case.

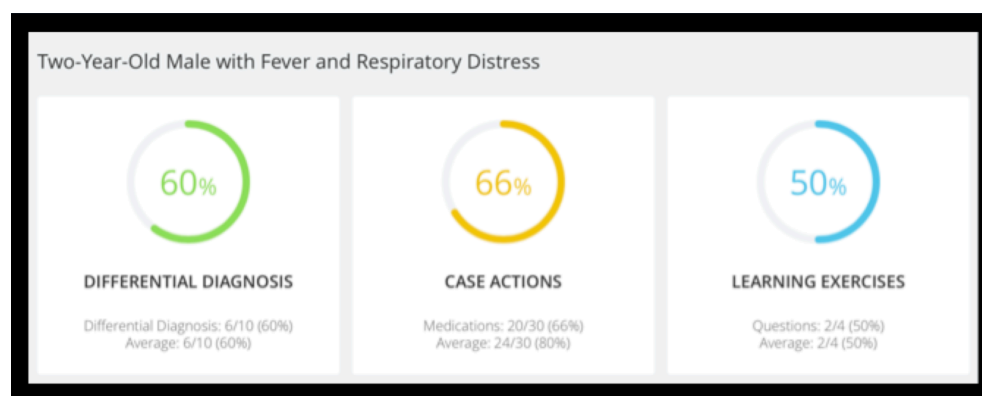


Figure 11: The user's cumulative score is based on the differential diagnosis, case actions and learning exercises sub-scores.



DIFFERENTIAL DIAGNOSIS		
Your Final Differential Diagnosis	Interpretation	Explanation
Croup	Correct Diagnosis!	Your patient presented with a stridorous cough, increased work of breathing after a few days of URI symptoms and improved with racemic epinephrine and corticosteroid, the mainstay of treatment
Gastroesophageal Reflux	Not a top differential diagnosis in this case	
Angioedema	Not a top differential diagnosis in this case	
Measles	Not a top differential diagnosis in this case	
Retropharyngeal Abscess	Not a top differential diagnosis in this case	

Figure 12: The differential diagnosis section is scored based on the presence and appropriate priority of the leading diagnosis, which is identified by the case author.

Case Actions:		
Presentation	Interpretation	Explanation
Dexamethasone	Indicated - beneficial to patient.	Corticosteroids are anti-inflammatory and decrease the laryngeal mucosal edema. These medications have been shown to reduce hospitalization rates, reduce the number of patients returning. A single 0.6 mg/kg dose of dexamethasone is typically used as the first-line corticosteroid in treating patients with croup.
Racemic Epinephrine	Indicated - beneficial to patient.	Racemic epinephrine acts on alpha-adrenergic receptors, causing arteriole vasoconstriction and thereby helpful in improving laryngeal mucosa edema in croup. Racemic epinephrine also acts on beta-adrenergic receptors, resulting in bronchodilation. Onset of action of 30 minutes, with a duration of action for 1.5 - 2 hours. Patients receiving racemic epinephrine should be observed at least 3 hours post-administration to monitor for recurrence of symptoms or worsening respiratory distress. Sinus tachycardia is a side effect of this medication.

Figure 13: User-initiated actions are scored as green (indicated, beneficial), yellow (non-indicated, non-risky) or red (non-indicated, harmful), and are accompanied by the reasoning for their designation.

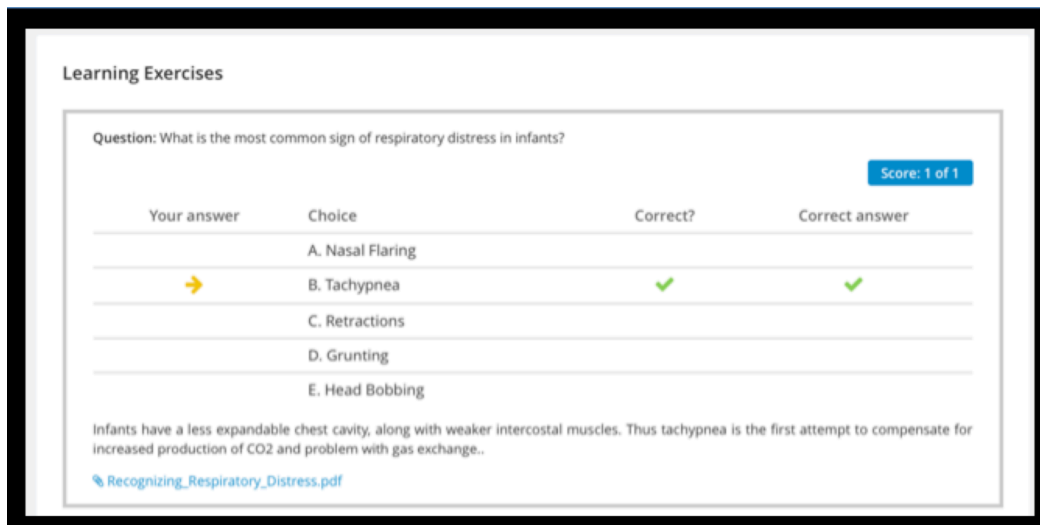


Figure 14: The interactive test questions are scored as either correct or incorrect. Explanations for appropriate responses are provided, along with links to other resources for self-directed learning.

#### Usability Testing: Participant Demographics

Of the 11 participants, five (45%) were male and six (55%) were female. The median level of comfort with computer-based technologies was 5/5, indicating that participants were very comfortable using this type of technology. This confidence was likely a consequence of the average amount of time that participants spent browsing the Internet ( $7.8 \pm 1.8$  hours), on email ( $3.2 \pm 2.8$  hours) or texting ( $1.5 \pm 1.4$  hours) in a given week. With respect to their favorite websites, the most often cited included: [www.google.com](http://www.google.com), [www.espn.com](http://www.espn.com), and [www.nytimes.com](http://www.nytimes.com). Participants emphasized that relevant content, an organized interface, easy-to-use functionality and visually engaging multimedia were all aspects of these websites that they appreciated. In addition, 100% of the participants admitted to using social networking tools (e.g., Facebook, Instagram, or Twitter).

More importantly, the participants each had medical education resources that they utilized either in service of self-directed learning or for clinical decision-making. Amongst this small cohort, Up-to-date and PubMed were the most frequently cited resources used in this vain. As compared to their more senior level colleagues, medical students appeared to be more aware of and employed a more varied list of resources that included: UWorld question banks, Pathoma<sup>47</sup>, Picmonic<sup>®</sup><sup>48</sup>, Piktochart and Sketchy Micro<sup>49</sup>. All participants had been exposed to other virtual patient simulators, ranging from the lowest to the highest fidelity platforms. MedU's CLIPP cases<sup>50</sup>, the USMLE™ Step 3 Case Simulator, and other OPENPediatrics simulators (i.e., Mechanical Ventilator<sup>51</sup> and Peritoneal

Dialysis<sup>52</sup>) were among the main systems previously used by the group. Participants explained that they appreciated systems that were clinically-oriented (i.e., focused on diagnosis and management), realistic (in terms of content and the consequences of user-initiated actions), open-ended, and that provided just-in-time feedback. However, they lamented systems that had irrelevant content, too much text, long completion times and limited interactivity.

Table 1: Demographics of participants and experience with computer-based technology

Variable	n (%)
	n = 11
<b>Gender</b>	
Male	5 (45)
Female	6 (55)
<b>Training</b>	
Medical Student	3 (27)
Resident	2 (18)
Fellow	4 (36)
Attending	2 (18)
<b>Specialty Training (Current or Future)</b>	
General Pediatrics	4 (36)
Pediatric Critical Care	4 (36)
Pediatric Emergency Medicine	1 (9)
Emergency Medicine	1 (9)
Unknown	1 (9)
<b>Previous Online Course Completion</b>	
Yes	7 (64)
No	4 (36)
<b>Exposure to Social Networking Tools</b>	
Yes	11 (100)
No	0 (0)
<b>Experience with Computer-based Technology</b>	
<b>Observations with Units</b>	
Median Comfort with Computer Based Technology	5/5 (Likert Scale)
Mean Hours Internet Browsing	7.8 ± 11.8 hr
Mean Hours on Email	3.2 ± 2.8 hr
Mean Hours Texting	1.5 ± 1.4 hr

### Usability Testing: Emergent Themes

With a median Likert rating of 5/5, 73% of participants strongly agreed that the PBLCloud system is a useful resource (Figure 15). 27% of participants appreciated the simulators grounding in reality and 64% of participants appreciated the focus on clinical diagnosis and management. Moreover, 27% of participants concluded that the just-in-time feedback was another helpful feature of the simulator. Likely a result of the aforementioned, 91% of participants believed that the simulator would be appropriate for a “broad spectrum of learners,” pending “tailoring of the [case] content and complexity” (Table 2).

64% of participants strongly agreed, with a median Likert rating of 5/5, that they enjoyed their experience with the PBLCloud simulator (Figure 15). 27% of the participants attributed their enjoyment to the relevance of the type and amount of information incorporated in the case as well as the overall focus on knowledge application. Importantly, 64% of participants emphasized that they appreciated the interface noting the attractive layout, straightforward functionality and inclusion of multimedia (Table 2).

With a median Likert rating of 5/5, 82% of participants reported that they were very interested in using PBLCloud in the future and this is likely the indirect result of assessing the system as both useful and enjoyable (Table 2). They acknowledged the potential of the PBLCloud simulator for both active self-study (55% of participants) as well as more formalized cased-based clinical instruction (36% of participants) (Table 2).

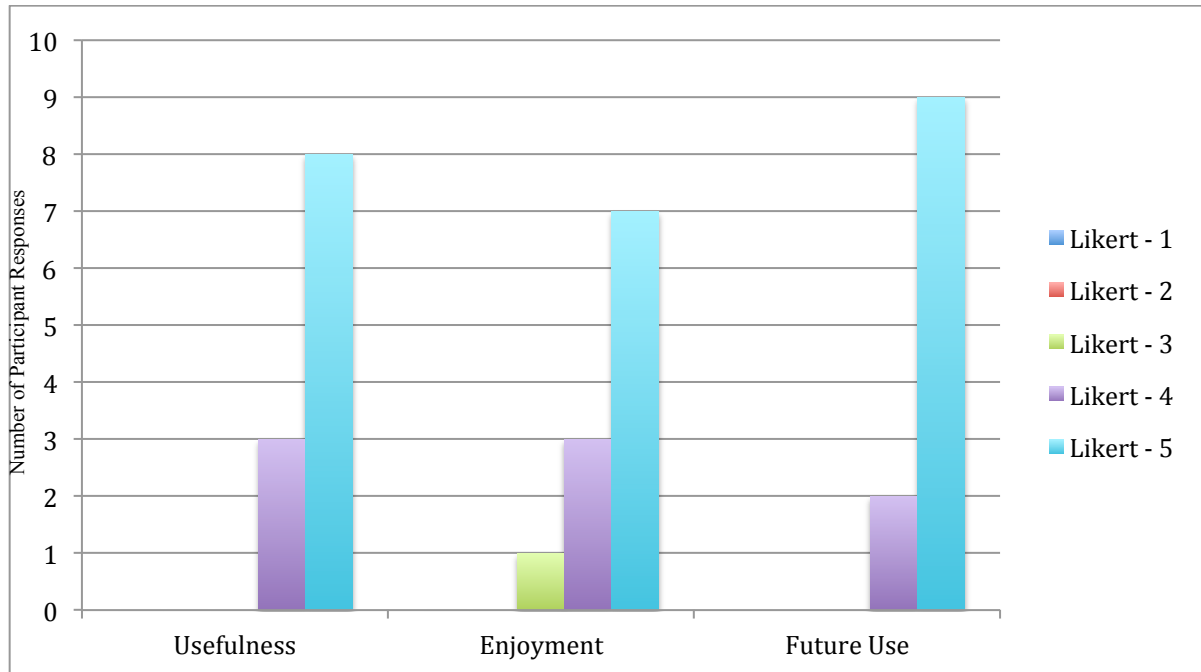


Figure 15: Bar graph demonstrating the number of participant responses (Y-axis) for three Likert-type items (Questions 1, 10 and 5 see Appendix III) from the 10-item open-ended questionnaire (X-axis). N = 11.

Given that all the participants had been exposed to VPS, whether in medical school or during residency, it was important to assess how PBLCloud compared to other simulators. 27% of participants stressed that the PBLCloud case simulator was more interactive than other simulators, allowing for a more free form experience that “[was] like open waters with multiple avenues to do down.” Overall, 27% of participants judged the system to be more engaging than others on the market, highlighting the limited text on screen as well as the incorporation of engaging multimedia on the Home Screen, in the Physical Examination section and in the Review Labs/Study section (Table 2).

There were three areas of improvement that the participants identified for the current version of the PBLCloud simulator. 64% explained that they initially had difficulty understanding how to navigate on and between webpages. They appeared to find the system intuitive “once I figured it out”, with a few participants recommending the addition of a tutorial before starting the first case. The history-taking feature of PBLCloud, enabled by a proprietary NLP algorithm, is the most novel as well as the most challenging aspect of this simulator to develop. 55% of participants described the history-taking function as frustrating as they received incorrect or unexpected responses for their queries. It was this frustration that prompted the authors to deemphasize the NLP

feature for the second round of usability testing and include a “Patient Chart” (Figure 2) for users to review the historical information until a reasonable approach to this challenge could be implemented. Finally, 55% of participants emphasized that more multimedia should be incorporated throughout the simulator. In particular, they suggested the inclusion of more images/animations “for all relevant physical examination findings” as well as potential “tracings” for vital signs, given the ICU setting (Table 2).

Table 2: The emergent themes from participants' responses to the open-ended questionnaire

Question Topic	Thematic Code	Description	Participants Observed In n (%)	Representative Excerpts
Usefulness	Clinically-oriented	Emphasis is placed on diagnosis and management.	7 (64%)	"allows for assessment formation and derivation of plan" Fellow-4
	Realism	Mimics the clinical environment in both look and feel.	3 (27%)	"It is useful that the cases and simulator are grounded in a real encounter, it is like open waters with multiple avenues to go down." Attending-2
	Helpful Feedback	Just-in-time feedback offered for learner-initiated actions.	3 (27%)	"formative feedback in how the patient evolves based on your actions" Resident-1
	Widely-applicable	The platform is appropriate for all levels of clinical training	10 (91%)	"broad spectrum of learners, could also be interprofessional" Fellow-4
Enjoyment	Engaging Interface	Layout, functionality and the inclusion of multimedia captured the users attention.	7 (64%)	"attractive interface" "interface was intuitive and straightforward" "like inclusion of multimedia" Attending-1
	Relevant	Addition of only crucial information and focus on the application of knowledge.	3 (27%)	"felt...relevant" Medical Student-1
Interest in future use	Self-Study	Opportunity for active, asynchronous and self-paced learning.	6 (55%)	"promotion of interactive, active learning" Resident-1
	Trainee Instruction	Option of case-based teaching tool	4 (36%)	"would use it during teaching session in the future" Attending-2
Comparison to other simulators	Robust Interactivity	Learner-initiated actions less restricted.	3 (27%)	"better...as it provides more interactivity" Resident-2
	More Engaging	Limited text on screen and incorporation of multimedia.	3 (27%)	"user interface is more engaging/aesthetically pleasing" "multimedia incorporation made it more engaging" Medical Student-3
Areas for Improvement	Navigation Issues	Page exploration and transitioning was not completely intuitive.	7 (64%)	"didn't know what to do initially" "the flow was intuitive, once I figured it out" Fellow-1
	Frustrating Historical Encounter	Natural language processing algorithm did not allow provide expected or appropriate responses to learner-initiated queries	6 (55%)	"the history taking was challenging" "would have been better to just have a patient chart" Resident-2
	Increase Multimedia	More multimedia should be included throughout the case simulator	6 (55%)	"should use images/animations for all relevant physical examination findings" "more audiovisual feedback" Fellow-2

#### Usability Testing: System Usability Score

The average SUS during Rounds 1 and 2 were  $79.7 \pm 12.0$  and  $82.5 \pm 19.8$ , respectively. A score of 68 indicates that a system is above average.<sup>34</sup>

#### Usability Testing: Bugs, Usability Issues and Content Errors

The number of bugs, usability issues and content errors decreased between the two rounds of testing from three bugs, four usability issues and five content errors identified in Round 1 to two bugs, one usability issue and two content errors identified during Round 2.



## DISCUSSION

### Introduction to Immersion

Our understanding of immersion is informed by both the gaming and virtual reality literatures<sup>18,53-56</sup>, each encompassing multiple conceptualizations of this construct. As described above, immersion can be defined as a psychological state characterized by perceiving oneself as included in and interacting with an environment (real or virtual) that provides a continuous stream of stimuli and experiences.<sup>18</sup> The term describes the degree of one's involvement in that environment and consists of three distinct stages: engagement, engrossment and total immersion (also known as presence). Humans have limited mental resources to help orient themselves to and selectively process environmental stimuli. Thus, one of the enabling conditions for immersion is focusing one's attention towards particular stimuli in an environment. Engagement, then, implies a deeper investment of one's attention, effort and time.<sup>18</sup> And as an individual is drawn further into an environment by the narrative, tasks and visuals he/she becomes emotionally invested and engrossed.<sup>18</sup> Finally, the individual reaches total immersion once the majority of sensory inputs received, interpreted and acted upon originates from that environment. High-fidelity virtual reality environments are not required to induce immersion; many books and movies with compelling narratives can foster a sufficiently immersive experience.<sup>18</sup> Broadly, the level of immersion depends on an individual's degree of isolation from his/her surroundings as well as the coherence and relevance of the stimuli therein.<sup>53</sup>

### Importance of Immersion

It has been demonstrated that immersion has significant implications for skill acquisition, mastery and transfer.<sup>57,58</sup> Gutierrez et al. investigated how the degree of immersion in a virtual reality simulation influenced knowledge acquisition.<sup>59</sup> The authors compared the performance of first-year medical students wearing a head mounted display (full immersion) to those using a standard computer screen (partial immersion). While both groups benefited from the virtual reality simulation, the fully immersed group demonstrated higher educational gains. Some of the highest fidelity VPS systems support voice and gesture recognition and have 3D avatars of patients that express synthesized or pre-recorded responses.<sup>60,61</sup> While these systems represent both a technological as well a pedagogical advancement, low fidelity systems can be utilized to accomplish specific educational objectives.<sup>1</sup> This leads to two important questions: 1) "just how immersive does an educational modality need to be to achieve its

objectives?” and 2) “which elements of a VPS system are the most sensitive and specific for inducing, sustaining and enhancing immersion?”

### Factors that Influence Immersion

Given our limited attentional resources, the experience of immersion is ever fleeting. It is influenced by multiple factors, both human- and system-related, that could be leveraged by game developers and educators alike. This discussion will focus on the system factors that influence immersion. Stanney et al. provide an in depth analysis of the human factors that contribute to the experience of immersion.<sup>62</sup> Witmer et al. offer a useful theoretical framework that categorizes the system factors that influence immersion into the following: distraction, sensory, control and realism factors.<sup>18</sup> Each of these factors influence immersion in their own way, but also interact with one another in ways that are hitherto misunderstood. In the following section, we utilize Witmer et al. framework to highlight specific design features of PBLCloud’s user interface and functionality that were incorporated to increase immersion and contrast these with features from other VPS systems in the literature.

### PBLCloud: Distraction and Sensory Factors

Immersion depends on the user’s ability to focus their selective attention on coherent and meaningful stimuli in an environment. Focusing one’s attention on novel information and selecting for relevant stimuli requires a considerable amount of mental effort. And according to Held and Durlach, this process is impaired when system interfaces are unnatural, clumsy and artifact laden.<sup>63</sup> Thus, we sought to develop an easy-to-use and intuitive interface by minimizing distractions on screen (e.g., excessive text, diagnostic tools, etc.) and highlighting key features of the patient (e.g., time point in the case, patient’s image, vital signs). Our goal was to free up the user’s mental resources to focus on the most clinically relevant features of the patient. In particular, audio, images, animations and live action video can be embedded into PBLCloud to facilitate multimodal sensory stimulation. The inclusion of this multimedia creates a rich virtual environment that is both complete and coherent. For example, our pilot case of pediatric patient with croup features multiple animations of a toddler in respiratory distress, demonstrating pale skin, nasal flaring, retractions, stridor and a “seal-like” cough. Unlike with text-based findings, users have the opportunity to review key features of the patient’s symptomology. This provides an opportunity for deliberate practice, in the hopes of bolstering the user’s retention of information and transfer to the real world

environment. Ultimately, by minimizing distractions and including opportunities for multimodal sensory stimulation, users may become more immersed in the activity and in turn develop a clearer mental model of the patient to manage with greater interactivity.

#### PBLCloud: Control and Realism Factors - Linearity and Interactivity

Witmer et al. describe how a higher degree of control increases user immersion, especially when the user is able to appreciate that their actions have immediate, meaningful and predictable results.<sup>18</sup> The CASUS, CAMPUS and Web-SF<sup>64,65</sup> systems all support linear (“string of pearls”) cases in which the patient’s narrative flows in a predetermined arc. While OpenLabyrinth<sup>66</sup> and DecisionSim™<sup>67</sup> are two systems that support branching narratives and allow users to explore the consequences of their actions. This feature helps to foster “deep learning” and critical thinking around key decision points<sup>68</sup>, especially when accompanied by adaptive feedback. PBLCloud is an example of a scenario-based VPS system that supports branching narratives. If users initiate non-indicated harmful actions or if patient parameters (e.g., vital signs and lab values) fall outside of acceptable normal limits, users receive formative feedback, through the hint system, explaining that their patient’s clinical status is deteriorating and that they should return to reevaluate their patient and management approach.

Immersion is further enhanced if these results are consistent with expectations derived from the real world. When the user initiates specific therapeutic interventions, their actions have appreciable consequences on patient parameters (i.e., clinical severity, vital signs and laboratory values). With respect to our pilot case, consensus guidelines identify the combination of corticosteroids and racemic epinephrine as the mainstay for the management of croup. Specifically, the onset of action of racemic epinephrine is between 10 – 30 minutes and the duration of effect is approximately 90 minutes. Using a rule-based system, the PBLCloud system stores this information and is able to model the physiologic processes that result from natural disease progression as well as user-initiated therapeutic interventions. For example, the administration of racemic epinephrine to a patient with uncomplicated croup will result in a decrease in respiratory rate, increase in heart rate and temporary improvement in symptomology. The combination of increased interactivity and a system that responds immediately and realistically will be a powerful tool to help engage users and facilitate the transfer of their insights from the virtual world back to the real world.

### Interviewing Skills

Osler is credited with saying, “Listen to the patient, he is telling you the diagnosis”. 70 – 75% of diagnoses are correctly identified based on the historical examination alone.<sup>69-71</sup> Despite this fact, until the 1970s, interviewing skills were not a part of medical curricula<sup>72</sup>, even though effective communication can be taught, learned and practiced. Originally, instruction around interviewing skills relied on lectures, textbooks, checklists and video taped examples.<sup>72</sup> Interviewing skills involve both the application of medical knowledge as well as basic communication skills. Medical education research in the 1970s – 1990s highlighted the importance of these skills and now communication is designated as a core competency by the ACGME.<sup>73</sup> These are critical components of the diagnostic process and proper communication has been demonstrated to improve patient satisfaction and outcomes, by both empowering patients and reducing medical errors.<sup>29,74</sup>

### Simulation of the Historical Encounter

VPS has been shown to be an effective educational modality for the teaching of interviewing skills.<sup>75,76</sup> Many systems employ selection-based approaches (e.g., multiple-choice lists, menu-based prompts and forced branching) to simulate the doctor-patient conversation. For example, students using the CASUS and CAMPUS systems must select questions from short preformed question lists. While, students using the Virtual Interactive Case System<sup>77</sup> from Canada have access to a comprehensive and well-organized menu of questions to chosen from. In contrast, other systems utilize chat-based (voice-recognition versus free-text) approaches to simulate a more natural back-and-forth between doctor and patient.<sup>72,78</sup> For example, the high fidelity 3D avatars from USC’s Institute of Creative Technologies utilize natural language processing to support voice recognition and combine this technology with speech production software to simulate the historical encounter.<sup>79</sup> Few other systems have made a successful attempt at using the free-text and response approach.<sup>28,29,80</sup>

### Natural Language Processing and Chat-Based Approaches

Chat-based approaches to simulate the doctor patient conversation require robust NLP algorithms to provide context specific responses to user inquiries. NLP involves a range of computational techniques used for the representation and automatic analysis of human language, based on syntax, semantics and pragmatics. This field of research is informed by computer science, artificial intelligence, machine learning, statistics and computational linguistics. Researchers and practitioners have been tackling this problem of human-computer interactions since the 1950s, with great success embodied by Google's search engine, IBM's Watson and Apple's Siri.<sup>81</sup> Though, in the domain of medical education, there have been a few studies that reported on how the ineffective chat-based approach of typed and spoken questions in early VPS systems distracted from the realism of the user's experience<sup>72,82</sup> and may have even increased the user's cognitive load.<sup>1</sup> With that in mind, Cook et al. suggested that it may be more effective for users to select their questions from a standard bank, ultimately emphasizing that other educational modalities, such as SPs, may be better suited for testing patient interviewing skills<sup>1</sup>.

### Selection-Based Approaches vs. Chat-Based Approaches

Educationally, selection-based approaches are not able to evaluate if students know 1) what questions to ask, 2) when to ask certain questions and 3) how to ask those questions. Many users have described systems utilizing selection-based approaches for the doctor-patient conversation as linear and inauthentic. Halan et al. compared selection-based versus chat-based approaches to simulating the doctor-patient conversation during the training of novice students.<sup>80</sup> Many of the students judged the chat-based interaction to be more realistic. This underscores how more natural interactions facilitate authenticity, which we know enhances immersion through control and realism factors. Some students were frustrated using the chat-based interaction since the response accuracy was lower than the selection-based cases. This is expected given that in selection-based approaches, students know which questions are available and have answers, while in chat-based formats, response accuracy depends on the instructor's ability to properly anticipate questions from users. One study demonstrated that it is difficult for experts to generate a comprehensive list of novice level questions as a result of the qualitative and quantitative difference in their approach to interviewing the patient.<sup>28</sup> Specifically, experts often ask fewer and more sensitive/specific questions than novices, reflecting their extensive experience and deeper understanding of the material.<sup>83,84</sup> Furthermore, chat-based interactions require students to recall what, when, how and why certain

questions should be asked. Accurate recall of information is more cognitively taxing, as compared to the recognition of information.

Cook and Triola explain that high fidelity systems are not always required to achieve educational objectives and can be quite resource and time intensive to develop.<sup>1</sup> This, in turn, requires a careful consideration of how to appropriately utilize lower fidelity systems to address learning goals of individuals at different expertise levels. Based on the feedback from participants, Halan et al. suggest that selection-based approaches may be more appropriate for the preclinical years during which students are still learning the basics behind interviewing patients.<sup>80</sup> These virtual patient simulators would focus on modeling the “what”, “when” and “how” of historical interview questions. While chat-based approaches may be more appropriate during the clinical years when students have developed greater clinical acumen. With a chat-based functionality, medical students during their clinical years would have the ability to deliberately practice their interview skills, such as taking patient’s sexual history<sup>85</sup>, in a psychological safe environment. It would allow students to focus more on recall as opposed to mere recognition, as with the selection-based approach. Ultimately, chat-based interactions better approximate the natural interaction between doctor and patient, establishing a more immersive environment. Greater immersion implies a better approximation of the mental space for clinical reasoning, which the user is practicing to transfer to the real world with real patients.

Cook et al. propose that it may be more effective to utilize SPs to acquire and master these skills, but the use of SPs is itself a resource-intensive process and the consistency of SPs has been called into question.<sup>1</sup> In that regard, we assert that with optimization, PBLCloud’s innovative algorithm for NLP will be robust enough to allow for an unscripted conversation that unfolds naturally between patient and physician, supporting context-specific responses to user queries. The social-emotional dimensions of the physician-patient interaction can only be approximated through NLP models and thus VPS will never be a substitute for the real patient encounter.<sup>1,86</sup> But this modality could prove to be an effective tool for the deliberate practice of these skills, especially in developing countries or resource-limited settings that have no other alternatives.<sup>12</sup>

## Usability Testing: Demographics and Themes

### *Participant Demographics*

Our cohort for usability testing demonstrated a familiarity as well as a comfort with computer-based technology. Their extensive experience with this type of technology has undoubtedly informed their preferences regarding features/aspects of these systems that they deem useful or satisfying. In general, participants acknowledged that they appreciate resources that are well-organized, easy-to-use, maintain relevant content and that incorporate multimedia. Moreover, participants explained that they utilize various resources for self-study and clinical decision-making. Senior level trainees and practitioners relied on a few medical education resources such as uptodate.com and PubMed; medicals students employed a more diverse set of resources, including newer applications such as Pathoma and Picmonic.

With respect to virtual patient simulators, participants had been exposed to MedU's CLIPP cases as well as the USMLE™ Step 3 Case Simulator. Similar to other cohorts reported on in the literature<sup>7</sup>, participants preferred systems that were clinically-oriented, realistic, highly interactive, maintained limited text and that included opportunities for just-in-time feedback. The various themes that emerged during analysis underline how the PBLCloud simulator was able to meet, at least in part, the expectations of the participants around computer-based technology broadly and virtual patient simulators specifically.

### *PBLCloud and Usefulness*

The responses from the post-session interview revealed that participants believed the PBLCloud simulator was a useful resource. The purpose of PBLCloud is to serve as a resource for the instruction and assessment of clinical reasoning, focusing on the formulation/revision of the diagnosis as well as the development and execution of a management plan. These represent skills that healthcare providers must practice and apply each day in the service of patient-care. Audiovisual realism and patient parameters that appropriately update based on user-initiated actions combined to successfully mimic the clinical environment. This realism likely allows for better knowledge consolidation along with enhanced knowledge transfer from the virtual world to the real world. Participants appreciated this focus on knowledge application and found the in-case hint system to be quite helpful as it provided

formative feedback regarding user-initiated actions. Given all this, participants assessed PBLCloud simulator to be widely applicable to all levels of training and could serve as a resource for interprofessional education.

#### *PBLCloud and Participant Enjoyment*

Participants enjoyed their experience with the simulator, highlighting the relevant content, limited text and visual stimulation. The clinical setting is a data-rich environment that requires numerous metacognitive skills to successfully navigate its waters. Students prefer systems that have limited text and an organized interface, likely because these features help to manage the cognitive load of the clinical environment. This allows users to focus on the most crucial elements of the case. In addition, participants concluded that the interface and multimedia were engaging aspects of the system. Many explained that the inclusion of multimedia helped to make the system more engaging than other simulators, namely MedU's CLIPP cases. Multimodal stimulation helps to immerse the user in the sights and sounds of the clinical environment and with respect to the physical examination section, helps provide a substrate to practice the clinical skills of observation, description and interpretation. However, in terms of areas of improvement, a theme of "more multimedia" emerged as participants wanted all the "[organ] systems" in the Physical Examination section to have multimedia or at least all the relevant "[organ] systems" that should be examined for a given case.

#### *PBLCloud as Compared to other Virtual Patient Simulators*

Overall, the average SUS was  $79.7 \pm 12.0$  in Round 1 and  $82.5 \pm 19.8$  in Round 2. Both of these scores indicate the PBLCloud system is an above average system in terms of usability, learnability and user satisfaction. Multiple participants noted that PBLCloud was more interactive than more linear systems like MedU's CLIPP cases. This increase in interactivity is reflected the chat-based history taking functionality, the interactive body diagram for physical examination as well as the ability to order various diagnostics and therapeutic interventions in the "Ordering" section. When combined with the individualized formative feedback from the hint system, this increased interactivity has the potential to be a powerful tool for learning clinical reasoning as students explore the consequences of their actions in this psychologically safe environment.<sup>7</sup>



### *PBLCloud and Areas for Improvement*

The novel aspect of PBLCloud is the history-taking feature that is enabled by a proprietary NLP algorithm. The goal was to increase user immersion by enhancing control and realism factor through, first and foremost, the chat-based historical encounter. While the system is able to provide context-specific responses to user queries, it is not 100% correct and can provide inappropriate or unexpected responses. Participants found this to be frustrating and some admitted to foregoing the history-taking activity altogether. As a result of this frustration, the authors included a “View Patient Chart” feature, which details the patient’s relevant clinical history and deemphasized the chat-based functionality. It is likely that this change accounts for a proportion of the increase in SUS from Round 1 to Round 2. Ultimately, the ideal system to simulate the doctor-patient conversation in a cost-effective and scalable manner has yet to be developed. Though, we hope future iterations of PBLCloud will be able to fill this niche.

Various approaches are being considered in order to address the response accuracy of our NLP algorithm. Some researchers have utilized a strategy in which medical students first formulate a case draft with historical dialogue and then senior level residents and attendings use their experience to successfully tailor the dialogue.<sup>28,80</sup> During our usability testing we learned that participants with higher levels of training had more difficulty with receiving appropriate responses to their queries, which recapitulates the extant literature regarding the qualitative difference between how experts and novices pose questions and ultimately conceptualize the clinical environment. Another approach would be to have samples from each level of training complete the case and record all the various ways questions were posed as a means to inform the NLP algorithm. Similarly, the use of machine learning would be a more efficient approach that would allow the authors to avoid requiring large samples of trainees to construct the historical dialogue. Finally, an intermediate approach would be to have a place on screen for users to click and identify incorrect or inappropriate computer-generated responses during history taking.<sup>28</sup> This would help developers and case authors alike to identify problems with the dialogue and begin to conceive of techniques to tackle these issues.

In addition, participants explained that the system was not fully intuitive, noting navigation issues on and between pages. This likely explains the level of confidence, which trended towards neutral, that participants had while using the PBLCloud system. The Case Assistant was one feature implemented in order to address this issue, and will be tested during the next round of usability testing. Moreover, the authors are entertaining the idea of implementing a tutorial section for users to complete before their first case.

### *PBLCloud and Participant's Future Use*

Despite these challenges and areas of improvement, the overall SUS score is approximately 80, which significantly increases the likelihood that an end-user would recommend the device or application to another individual. This is underscored by the fact that the participants reported a high likelihood of utilizing the PBLCloud simulator in the future for either asynchronous self-paced learning or for case-based instruction, which were among the goals for creating the system.

### Limitations

\* See sub-section PBLCloud and Areas for Improvement

The medical education literature lacks theory-based guidelines for the design, development, implementation and evaluation of VPS systems. As a result, the authors completed a thorough literature search to identify any tips<sup>25,27</sup> and guiding principles<sup>24,26,87</sup> to aid in the design process. These guiding principles were supplemented by the combined experience of the authors. Though, design solutions unavoidably reflect the norms, values, and concepts held by designers. Cook et al. calls into question many features commonly incorporated into VPS systems, including the use of multimedia in VPS cases as well as the incorporation of NLP.<sup>1</sup> These concerns are warranted given that the development of VPS systems and the subsequent authoring of cases can be a resource-intensive process. PBLCloud makes use of these features as a means to improve interactivity, authenticity, student immersion and the potential for learning. To arrive at any meaningful conclusions about whether or not it is appropriate to include such features, it will require clinical educators to begin a branch of research dedicated to the psychometric analysis of specific features incorporated into VPS systems. This will prove to be a Titanic undertaking and thus call for a multi-institutional and interdisciplinary approach to conduct hypothesis testing around the most effective features of VPS system, instructional design and ultimate curricular integration.

With respect to usability testing, one limitation involves the fact that we had a small sample size of users pooled from two academic institutions, most with a background in pediatrics. Future psychometric testing for validity will require a larger sample size to make more generalizable conclusions about the system. The usability testing above focused on our initial test case involving the diagnosis and management of viral croup, but further testing will need to include a variety of cases to ensure reliability. Afterwards, we will need to complete usability testing in other geographic and clinical contexts since we anticipate that this tool will be used by the global

community and especially in resource-limited settings. Given this later aim, we will need to ensure that future versions of the system support multilingual users.

### Future Directions

#### *New Features and Psychometric Testing*

Based on usability testing OPENPediatrics will be considering various features for the inclusion in future iterations of PBLCloud. These examples such as a comprehensive differential diagnosis list of ICD-10 codes organized by specialty, costs for individual interventions in the “Orders” as well as to the development of a cumulative scoring system that accounts for the analytical versus non-analytical thinking of the user, by balancing management efficiency and thoroughness. The final steps for PBLCloud will include conducting a psychometric analysis of the system, focusing on validity, reliability and educational impact (i.e., short-term, intermediate-term, long-term and evidence of transfer to real patient encounters). We also intend to investigate the level of immersion induced by the system and compare that against other published VPS systems. This will require multiple tools from the literature used to measure an individual’s capacity for and experience of immersion.<sup>18,62,88</sup>

#### *Application of PBLCloud: OPENPediatrics’ Mission*

OPENPediatrics is leveraging Internet-based technology and the methods of social learning to build and sustain networks of clinicians, across both space and time, to help clinicians make informed decision and solve problems in ways that were previously infeasible. The three core components of the OPENPediatrics knowledge-sharing platform include 1) facilitating information on-demand, 2) growing a global community of clinicians and 3) establishing evidence-based and peer-reviewed guided clinical learning pathways. In the following we will outline how OPENPediatrics will utilize PBLCloud to tackle the challenges facing healthcare education globally and how it will integrate this modality with its other web-based tools into robust guided learning pathways.

*Application of PBLCloud: Global Community*

Supporting healthcare education in resource-limited settings is inherently challenging due to the unequal distribution of the global healthcare workforce as well as the marginal monetary investment in educational initiatives. More specifically, sub-Saharan Africa struggles with 24% of the world's disease burden, but is equipped with only 3% of the global healthcare workforce.<sup>89</sup> This translates to approximately 2.3 healthcare workers per 1000 individuals, while the United States enjoys 24.8 healthcare workers per 1000 individuals.<sup>90</sup> The etiology of these critical shortages is multifactorial and includes the migration of healthcare workers to other countries due to low wages, unstable working environments, weak public health systems as well as deficits in the infrastructure and funding for training.<sup>91</sup> As a result, the stocks of healthcare workers are decreasing and there is limited capacity to replenish those numbers in a sustainable fashion. Moreover, rural areas and under-populated cities in the United States are facing similar challenges and a 2015 study projects that the United States, as a whole, faces a shortage of between 46,000 – 90,000 physicians.<sup>92</sup>

Thus, there is a mounting need to improve the quantity and quality of the global healthcare workforce by investing in innovative methods for the education of both trainees and graduates.<sup>93</sup> Many institutions in low-to-middle income countries have turned to e-learning and related educational modalities as a potential avenue to improve faculty efficiency and effectiveness as well as mitigate the global migration of healthcare workers.<sup>94</sup> One of the cardinal missions of OPENPediatrics is to use Internet based technology to establish and maintain a global community of clinicians, sharing their knowledge in the service of caring for critically ill children. Currently, the OPENPediatric's educational resources are being utilized in 145 countries worldwide, including such countries as Malawi and Ethiopia, both designated as low-to-middle income countries according to the World Bank.<sup>95</sup> We believe that the PBLCloud can be used to support the training of these healthcare workers in developing countries and other resource-limited settings. Our goal is for the system to be intuitive enough to reduce the labor associated with authoring rich VPS cases and flexible enough for instructors to adapt cases to local realities (e.g., disease burden, culture, language). Furthermore, PBLCloud will be an open access resource and thus prove to be a cost-effective alternative for institutions that do not have the requisite infrastructure and institutional funding. Institutional partnerships represent another means of addressing this problem, in addition to the intentional sequencing of educational modalities to support transformation learning. This latter option seeks to maximize the

utility of current resources without incurring additional costs and the effectiveness of this strategy has been demonstrated by a small sample of studies in the literature<sup>91</sup> and will be touched upon in the following section.

*Application of PBLCloud: Clinical Learning Pathways (CLPs)*

There are many educational modalities (e.g., didactics, VPS and SPs) in the clinical educator's armamentarium, but these modalities do not teach or assess the same cognitive, psychomotor or socio-emotional constructs. As a result, curriculum designers need to think critically about the optimal means of integrating these educational methods and materials. CLPs are a combination of educational methods and materials thoughtfully orchestrated into a string of activities to promote and sustain learning.<sup>96</sup> While VPS can facilitate asynchronous and distance learning, other patterns of activity have been proposed, using VPS as a reference activity, in blended learning, for group collaboration and as a bridging activity.<sup>26</sup> In PBLCloud, users have the ability to select and sequence cases to fill self-assessed knowledge gaps. Similarly, instructors can sequence cases in such a fashion as to underscore key teaching points for their students; this could include constructing a module of cases that all together emphasize the differential diagnosis of respiratory disease in a pediatric patient (Figure 8). While there is a paucity of the theories and practice guidelines to help clinical educators integrate VPS into a larger curriculum, the authors have turned to adult learning theory and specifically, Kolb's Theory of experiential learning, to inform the construction of CLPs that are both meaningful and effective.

Kolb's Theory is a four-stage cyclical theory of experiential learning that includes 1) a concrete experience, 2) reflective observation 3) abstract conceptualization and 4) active experimentation before entering the cycle.<sup>97</sup> The following will detail the sequencing of various educational modalities offered by OPENPediatrics into a clinical learning pathway that medical students, for example, could be assigned by their instructors to complete during their 3<sup>rd</sup> year clinical rotation in pediatrics. More specifically, a trainee may be caring for a pediatric patient that arrives in the ED with signs and symptoms consist with a viral croup infection and they were tasked with the concrete experience of formulating the proper diagnosis for this patient and recommending the appropriate evidence-based management. Afterwards, they could collaborate with the global community on OPENPediatrics message boards to help reflect on that patient encounter and then review the various videos and summaries on the platform in order to consolidate the general approach to pediatric respiratory distress. Trainees could then utilize PBLCloud to actively experiment with their recently acquired medical knowledge and reasoning skills in a psychologically safe

environment. Finally, this could be followed up with a concrete experience of using high-fidelity mannequin simulation in an environment that better approximates the real clinic setting. This cycle could be repeated until reaching the next concrete experience, which could be the care of a real patient. Overall, this represents a progressive and student-centered approach to learning the diagnosis and management of a pediatric patient presenting in respiratory distress. Finally, OPENPediatrics will be utilizing the learning course management system known as OpenEdX, which is a web-based application for the administration, documentation, tracking and reporting for educational programs.<sup>98</sup> We anticipate this will further help instructors develop and implement their own CLPs for students.

#### *New PBLCloud Interfaces: Instructor Interface*

In the midst of the inertia from competing clinical, research and administrative responsibilities, clinicians have less time available to dedicate toward teaching. This trend is likely to worsen and thus alternative approaches, such as VPS, that are evidenced-based, peer-reviewed and easy-to-use will prove valuable in addressing this challenge. For instance, Maldonado et al. demonstrated that a hybrid (VP - paper) PBL case module for physician assistant students resulted in a 40% savings in facilitator time, with a projected potential savings of 92% if VPS were to replace the paper cases entirely.<sup>99</sup> Albeit, many studies have stressed that authoring VPS cases can be labor-intensive, time-intensive and monetarily costly.<sup>1</sup> Huang et al. reported that 61% of VPS cases take more than 6 months to produce and 85% of those cases cost \$10,000 to develop and implement.<sup>8</sup> OPENPediatrics provides open-access of evidenced-based and peer-reviewed resources to its users and PBLCloud will be launched on this knowledge-sharing platform. Our goal would be for PBLCloud to represent a cost-effective alternative that supports the efficient authoring, editing, and sharing of cases as well as individual multimedia resources.

A tree-like template system to guide authors, comprehensive repository for historical questions, physical examination signs and condition-specific lab values as well as a large referatory (collection of metadata such as learning objectives for a case) will help to facilitate both de novo case construction as well as the adaptation of published resources to fit local realities and desired educational outcomes.

*New PBLCloud Interfaces: Administrator Interface*

OPENPediatrics will serve as the administrator for PBLCloud cases and curricula published on our knowledge-sharing platform. The main responsibilities of the administrator is to oversee the peer-review process to ensure that 1) the content of the information in instructor-derived cases are up-to-date and evidence-based, 2) the structure of the cases are educationally sound and 3) that any intellectual property issues have been addressed. Furthermore, the staff at OPENPediatrics will be available to support instructors in the development of cases and sequencing of these cases into larger modules and curricula.

**CONCLUSION**

Physicians must command a large corpse of biomedical knowledge and develop a wide array of skills (i.e., cognitive, psychomotor and socio-emotional) that facilitate their ability to “cure sometimes, to relieve often, to comfort always”.<sup>100</sup> This is further underscored and compounded by the fact that in the 1930s only one or two new therapeutic agents were implemented into practice each year and now we are on the order of over 30 new agents developed and introduced each year.<sup>101</sup> VPS systems broadly and PBLCloud in particular, were developed to address this reality as a means to facilitate the instruction, deliberate practice and assessment of clinical reasoning. Many existing systems suffer from limited interactivity and unnatural interactions, both control and realism factors that affect the user’s immersion and by extension potential educational benefit. Our specific strategy was to increase immersion through a chat-based functionality to simulate the doctor-patient conversation. PBLCloud was able to address certain needs/priorities of users, who assessed the system as useful and enjoyable as well as more interactive and more engaging than other simulators, ultimately seeking to use the system in the future. Users identified areas for improvement and future efforts will be geared towards optimizing PBLCloud’s NLP algorithm so as to improve the doctor-patient conversation. Furthermore, it will be important to investigate the validity, reliability, level of induced user immersion and educational impact of the system. Overall, we anticipate that PBLCloud will be a flexible, cost effective and scalable tool for the learning, teaching and assessment of clinical reasoning.

## REFERENCES

1. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. *Medical education*. 2009;43(4):303-311.
2. Norman G. Research in clinical reasoning: past history and current trends. *Medical education*. 2005;39(4):418-427.
3. Elstein AS. Thinking about diagnostic thinking: a 30-year perspective. *Advances in health sciences education*. 2009;14(1):7-18.
4. Ellaway R, Poulton T, Fors U, McGee JB, Albright S. Building a virtual patient commons. *Medical teacher*. 2008;30(2):170-174.
5. Guise V, Chambers M, VÄLimÄKi M. What can virtual patient simulation offer mental health nursing education? *Journal of Psychiatric and Mental Health Nursing*. 2012;19(5):410-418.
6. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic medicine : journal of the Association of American Medical Colleges*. 2004;79(10 Suppl):S70-81.
7. Botezatu M, Hult H, Fors UG. Virtual patient simulation: what do students make of it? A focus group study. *BMC Med Educ*. 2010;10:91.
8. Huang G, Reynolds R, Candler C. Virtual patient simulation at US and Canadian medical schools. *Academic medicine : journal of the Association of American Medical Colleges*. 2007;82(5):446-451.
9. Forsberg E, Georg C, Ziegert K, Fors U. Virtual patients for assessment of clinical reasoning in nursing -- a pilot study. *Nurse education today*. 2011;31(8):757-762.
10. Zary N, Johnson G, Fors U. Web-based virtual patients in dentistry: factors influencing the use of cases in the Web-SP system. *European journal of dental education : official journal of the Association for Dental Education in Europe*. 2009;13(1):2-9.
11. Conradi E, Kavia S, Burden D, et al. Virtual patients in a virtual world: Training paramedic students for practice. *Medical teacher*. 2009;31(8):713-720.
12. Dewhurst D, Borgstein E, Grant ME, Begg M. Online virtual patients - A driver for change in medical and healthcare professional education in developing countries? *Medical teacher*. 2009;31(8):721-724.
13. Colleges AoAM. Effective Use of Educational Technology in Medical Education: Summary Report 2006 AAMC Colloquium on Educational Technology 2007:7.
14. Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in health professions education: a systematic review and meta-analysis. *Academic medicine : journal of the Association of American Medical Colleges*. 2010;85(10):1589-1602.
15. Fida M, Kassab SE. Do medical students' scores using different assessment instruments predict their scores in clinical reasoning using a computer-based simulation? *Advances in medical education and practice*. 2015;6:135-141.
16. Dillon GF, Clyman SG, Clauser BE, Margolis MJ. The introduction of computer-based case simulations into the United States medical licensing examination. *Academic medicine : journal of the Association of American Medical Colleges*. 2002;77(10 Suppl):S94-96.
17. Cook DA. The research we still are not doing: an agenda for the study of computer-based learning. *Academic medicine : journal of the Association of American Medical Colleges*. 2005;80(6):541-548.
18. Witmer BG, Singer MJ. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*. 1998;7(3):225-240.
19. Bransford JD, Brown AL, Cocking RR. How people learn: Brain, mind, experience, and school. National Academy Press; 1999.
20. Fischer M. CASUS-An authoring and learning tool supporting diagnostic reasoning. *Zeitschrift für Hochschuldidaktik*. 2000;1(1):87-98.
21. Garde S, Bauh M, Haag M, et al. CAMPUS-computer-based training in medicine as part of a problem-oriented educational strategy. *Studies in Learning, Evaluation, Innovation and Development*. 2005;2(1).
22. Wolbrink TA, Kissoon N, Burns JP. The Development of an Internet-Based Knowledge Exchange Platform for Pediatric Critical Care Clinicians Worldwide. *Pediatric Critical Care Medicine*. 2014;15(3):197-205.
23. The Design Thinking Process. 2012; <http://dschool.stanford.edu/redesigningtheater/the-design-thinking-process/>. Accessed 12, 2015.
24. Huwendiek S, Reichert F, Bosse HM, et al. Design principles for virtual patients: a focus group study among students. *Medical education*. 2009;43(6):580-588.



25. Posel N, Mcgee JB, Fleischer DM. Twelve tips to support the development of clinical reasoning skills using virtual patient cases. *Medical teacher*. 2015;37(9):813-818.
26. McGee JB. Designing, Developing and Implementing Branched-Narrative Virtual Patients for Medical Education, Training and Assessment. University of Pittsburgh, School of Medicine.
27. Begg M. Virtual patients: practical advice for clinical authors using Labyrinth. *The clinical teacher*. 2010;7(3):202-205.
28. Rossen B, Lok B. A crowdsourcing method to develop virtual human conversational agents. *International Journal of Human-Computer Studies*. 2012;70(4):301-319.
29. Stevens A, Hernandez J, Johnsen K, et al. The use of virtual patients to teach medical students history taking and communication skills. *The American Journal of Surgery*. 2006;191(6):806-811.
30. Creswell JW. *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications; 2013.
31. Six JMaM, J. How to Determine the Right Number of Participants for Usability Studies. 2016; <http://www.uxmatters.com/mt/archives/2016/01/how-to-determine-the-right-number-of-participants-for-usability-studies.php>. Accessed February 1, 2017.
32. Nielsen J. *Why You Only Need to Test with 5 Users*. 2000.
33. Olszewski AE. *Virtual Simulation and Serious Games for Medical Education: A Review of the Literature and Development of a Virtual Peritoneal Dialysis Simulator*: Harvard Medical School, Harvard University 2016.
34. Sauro J. *Measuring Usability with the System Usability Scale*. 2011. <http://www.measuringu.com/sus.php>.
35. Brooke J. SUS-A quick and dirty usability scale. *Usability evaluation in industry*. 1996;189(194):4-7.
36. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC medical research methodology*. 2008;8(1):1.
37. Microsoft Excel for Mac [computer program]. 2011.
38. Lavery D, Cockton G, Atkinson MP. Comparison of evaluation methods using structured usability problem reports. *Behaviour & Information Technology*. 1997;16(4-5):246-266.
39. Wilson C, Coyne KP. The whiteboard: Tracking usability issues: to bug or not to bug? *interactions*. 2001;8(3):15-19.
40. Glaser BG, Strauss AL. *The discovery of grounded theory: Strategies for qualitative research*. Transaction publishers; 2009.
41. Haluck R, Hoffman H, Mogel T, Phillips R, Robb R, Vosburgh K. *Medicine Meets Virtual Reality 17: NextMed: Design For/the Well Being*. Studies in Health Technology and Informatics. IOS Press; 2009.
42. Donner RS, Bickley H. Problem-based learning in American medical education: an overview. *Bulletin of the Medical Library association*. 1993;81(3):294.
43. Bergin R, Fors U. *Interactive Simulation of Patients—An advanced tool for student activating learning in medicine*. Paper presented at: E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education2002.
44. Jung B, Ahad A, Weber M. *The affective virtual patient: An E-learning tool for social interaction training within the medical field*. Paper presented at: Proceedings of TESI-Training Education & Education International Conference. Nexus2005.
45. Clauser BE, Margolis MJ, Clyman SG, Ross LP. Development of automated scoring algorithms for complex performance assessments: A comparison of two approaches. *Journal of Educational Measurement*. 1997;34(2):141-161.
46. Clauser BE, Subhiyah RG, Nungester RJ, Ripkey DR, Clyman SG, McKinley D. Scoring a Performance-Based Assessment by Modeling the Judgments of Experts. *Journal of Educational Measurement*. 1995;32(4):397-415.
47. Sattar HA. *Fundamentals of pathology: Medical course and step 1 review*. Pathoma. com; 2011.
48. Yang A, Goel H, Bryan M, et al. The Picmonic(®) Learning System: enhancing memory retention of medical sciences, using an audiovisual mnemonic Web-based learning platform. *Advances in medical education and practice*. 2014;5:125-132.
49. Sketchy Group L. *SketchyMicro*. 2013.
50. MedU. *CLIPP Pediatric Cases*. <http://www.med-u.org/clipp>.
51. Wolbrink T, Burns J. Serious gaming in pediatric critical care medicine: the development of a virtual mechanical ventilator. *Pediatric Critical Care Medicine*. 2014;15(4\_suppl):221.
52. Olszewski AE. *Virtual Simulation and Serious Games for Medical Education: A Review of the Literature and Development of a Virtual Peritoneal Dialysis Simulator* 2016.

53. Cheng K, Cairns PA. Behaviour, realism and immersion in games. Paper presented at: CHI'05 extended abstracts on Human factors in computing systems2005.
54. Wilcox-Netepczuk D. Immersion and realism in video games-The confused moniker of video game engrossment. Paper presented at: Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational & Serious Games (CGAMES), 2013 18th International Conference on2013.
55. Brown E, Cairns P. A grounded investigation of game immersion. Paper presented at: CHI'04 extended abstracts on Human factors in computing systems2004.
56. Barros HO, Soares MM, Rolim Filho EL, Correia W, Campos F. Virtual Reality Immersion: An Important Tool for Diagnostic Analysis and Rehabilitation of People with Disabilities. Paper presented at: International Conference of Design, User Experience, and Usability2013.
57. Ragan ED, Sowndararajan A, Kopper R, Bowman DA. The Effects of Higher Levels of Immersion on Procedure Memorization Performance and Implications for Educational Virtual Environments. *Presence: Teleoperators and Virtual Environments*. 2010;19(6):527-543.
58. Hirose M, Schmalstieg D, Wingrave C, Nishimura K. Higher levels of immersion improve procedure memorization performance. 2009.
59. Gutierrez F, Pierce J, Vergara VM, et al. The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. *Studies in health technology and informatics*. 2007;125:155-160.
60. Kenny P, Parsons TD, Gratch J, Rizzo AA. Evaluation of Justina: a virtual patient with PTSD. Paper presented at: International Workshop on Intelligent Virtual Agents2008.
61. Talbot TB, Sagae K, John B, Rizzo AA. Sorting out the virtual patient: how to exploit artificial intelligence, game technology and sound educational practices to create engaging role-playing simulations. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*. 2012;4(3):1-19.
62. Stanney KM, Mourant RR, Kennedy RS. Human factors issues in virtual environments: A review of the literature. *Presence*. 1998;7(4):327-351.
63. Held RM, Durlach NI. Telepresence. *Presence: Teleoperators & Virtual Environments*. 1992;1(1):109-112.
64. Zary N, Fors UG. WASP-a generic web-based, interactive, patient simulation system with built-in case authoring. Paper presented at: EdMedia: World Conference on Educational Media and Technology2003.
65. Zary N, Johnson G, Fors U. Web-based virtual patients in dentistry: factors influencing the use of cases in the Web-SP system. *European Journal of Dental Education*. 2009;13(1):2-9.
66. Calinici T, Muntean V. Open Labyrinth-a Web Application for Medical Education Using Virtual Patients. *Applied Medical Informatics*. 2010;27(4):15.
67. DecisionSim. <https://app.decisionsimulation.com/shell/Login.aspx>.
68. Salminen H, Zary N, Björklund K, Toth-Pal E, Leanderson C. Virtual patients in primary care: developing a reusable model that fosters reflective practice and clinical reasoning. *Journal of medical Internet research*. 2014;16(1):e3.
69. Tsukamoto T, Ohira Y, Noda K, Takada T, Ikusaka M. The contribution of the medical history for the diagnosis of simulated cases by medical students. *Int J Med Educ*. 2012;3:78-82.
70. Peterson MC, Holbrook JH, Von Hales D, Smith NL, Staker LV. Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. *The Western journal of medicine*. 1992;156(2):163-165.
71. Lown B. *The lost art of healing: Practicing compassionate medicine*. New York: Ballantine Books; 1999.
72. Dickerson R, Johnsen K, Rajj A, et al. Evaluating a script-based approach for simulating patient-doctor interaction. Paper presented at: Proceedings of the International Conference of Human-Computer Interface Advances for Modeling and Simulation2005.
73. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. *Health Affairs*. 2002;21(5):103-111.
74. Coulehan JL, Platt FW, Egener B, et al. "Let me see if I have this right...": words that help build empathy. *Annals of Internal Medicine*. 2001;135(3):221-227.
75. Triola M, Feldman H, Kalet A, et al. A randomized trial of teaching clinical skills using virtual and live standardized patients. *Journal of general internal medicine*. 2006;21(5):424-429.
76. Campbell AJ, Amon KL, Nguyen M, et al. Virtual World Interview Skills Training for Students Studying Health Professions. *Journal of Technology in Human Services*. 2015;33(2):156-171.
77. Education PI. Virtual Interactive Case System. <http://pie.med.utoronto.ca/VIC/index.htm>.
78. Hubal RC, Kizakevich PN, Guinn CI, Merino KD, West SL. The virtual standardized patient. Paper presented at: Medicine Meets Virtual Reality2000.

79. Kenny P, Parsons TD, Pataki C, et al. Virtual Justina: A PTSD virtual patient for clinical classroom training. *Annual Review of CyberTherapy and Telemedicine*. 2008;6(1):113-118.
80. Carnell S, Halan S, Crary M, Madhavan A, Lok B. Adapting Virtual Patient Interviews for Interviewing Skills Training of Novice Healthcare Students. Paper presented at: International Conference on Intelligent Virtual Agents2015.
81. Cambria E, White B. Jumping NLP curves: a review of natural language processing research [review article]. *IEEE Computational Intelligence Magazine*. 2014;9(2):48-57.
82. Johnsen K, Dickerson R, Raji A, et al. Experiences in using immersive virtual characters to educate medical communication skills. Paper presented at: IEEE Proceedings. VR 2005. *Virtual Reality*, 2005.2005.
83. Schaffner KF. *Logic of discovery and diagnosis in medicine*. Univ of California Press; 1985.
84. Kenny PG, Parsons TD, Rizzo A. A comparative analysis between experts and novices interacting with a virtual patient with PTSD. *Annu Rev CyberTherapy Telemed*. 2009;7:122-124.
85. Deladisma A, Mack D, Bernard T, et al. Virtual patients reduce anxiety and enhance learning when teaching medical students sexual history-taking skills. *Association for Surgical Education*. 2007:756-760.
86. Deladisma AM, Cohen M, Stevens A, et al. Do medical students respond empathetically to a virtual patient? *The American Journal of Surgery*. 2007;193(6):756-760.
87. Bateman J, Allen ME, Kidd J, Parsons N, Davies D. Virtual patients design and its effect on clinical reasoning and student experience: a protocol for a randomised factorial multi-centre study. *BMC medical education*. 2012;12(1):1.
88. Witmer BG, Jerome CJ, Singer MJ. The factor structure of the presence questionnaire. *Presence*. 2005;14(3):298-312.
89. Taylor AL, Hwenda L, Larsen B-I, Daulaire N. Stemming the brain drain—a WHO global code of practice on international recruitment of health personnel. *New England journal of medicine*. 2011;365(25):2348-2351.
90. Organization WH. *The world health report: 2006: working together for health*. 2006.
91. Frenk J, Chen L, Bhutta ZA, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. *The lancet*. 2010;376(9756):1923-1958.
92. Dall T, West T, Chakrabarti R, Iacobucci W. *The complexities of physician supply and demand: projections from 2013 to 2025*. Washington, DC: Association of American Medical Colleges. 2015.
93. Frenk J, Chen L, Bhutta ZA, et al. [Health professionals for a new century: transforming education to strengthen health systems in an interdependent world]. *Revista peruana de medicina experimental y salud publica*. 2011;28(2):337-341.
94. Frehywot S, Vovides Y, Talib Z, et al. E-learning in medical education in resource constrained low-and middle-income countries. *Human resources for health*. 2013;11(1):1.
95. Bank W. *Data for Low income, Least developed countries: UN Classification*. 2016; <http://data.worldbank.org/?locations=XM-XL>.
96. Chegwiddden WR. A problem-based learning pathway for medical students: improving the process through action research. *ANNALS-ACADEMY OF MEDICINE SINGAPORE*. 2006;35(9):642.
97. Kolb David AE. *Experience as the source of Learning and Development*. New Jersey: Prentice-Hall; 1984.
98. Harvard M. OPEN edX. 2012; <https://open.edx.org/>.
99. Maldonado R. The use of multimedia clinical case scenario software in a problem-based learning course: impact on faculty workload and student learning outcomes. *The journal of physician assistant education : the official journal of the Physician Assistant Education Association*. 2011;22(3):51-55.
100. Trudeau E. 1800s.
101. Kinch MS, Haynesworth A, Kinch SL, Hoyer D. An overview of FDA-approved new molecular entities: 1827-2013. *Drug discovery today*. 2014;19(8):1033-1039.

## APPENDIX I

### PBLCloud: Think Aloud Protocol

#### Introduction:

#### Facilitator's Dialogue

Thank you for agreeing to provide OPENPediatrics with feedback on our new virtual patient simulator, PBLCloud. PBLCloud simulates the patient encounter from beginning to end as a means to teach and assess clinical reasoning. Before we begin, I would like to provide you with some background and I will be reading from a script in order to make sure I address all aspects of the testing process.

First off, we are testing the interface and functionality of PBLCloud, and not you. You can stop the testing at any time, so please just tell me and we will stop. We will be using a think-aloud protocol and as you are interacting with the simulator I will be asking you to talk me through your thought processes. At certain times in the protocol I will ask you some directed questions and give you some specific tasks to accomplish. Please just verbalize what you are doing, as you do it. This may feel strange and I would like to warn you in advance that I will ask you to keep talking if you fall silent. Do you have questions for me?

#### Demographics:

#### Facilitator's Dialogue

To begin, I would like to ask you few background questions.

#### Basic Demographics

Name:

Gender (Male/Female/Unidentified/Other):

Position (Medical Student/Resident/Fellow):

Time at BCH:

Specialty:

#### Expertise with Computer-based and Web-based Technology:

1. Please rate your comfort level with computer-based technologies:
  1. Very uncomfortable
  2. Uncomfortable
  3. Neutral
  4. Comfortable
  5. Very comfortable

2. How many hours a day would you say you spend...
  - a. browsing the web?
  - b. on email?
  - c. texting?
  - d. Do you have a favorite website you like to use?
    - i. Why do you like it?
3. Have you ever taken an online course? Yes/No
  - a. If yes, what was it?
  - b. When did you take it?
  - c. What did you like about it?
4. Do you use online social networking tools (e.g., Facebook, Twitter)? Yes/No
  - a. If so, which ones?
  - b. If not, why not?
5. What is(are) your favorite online resource(s) for medical education (e.g., Up-to-date, question banks for boards studying, Pathoma) ?
  - a. Why do you like it(them)?
6. Have you used any online simulations (e.g., CLIPP cases)?
  - a. If so, which ones?
  - b. What did you like about it (interface/functionality)?
  - c. What did you not like about it (interface/functionality)?

Think Aloud Protocol:

## Facilitator's Dialogue:

Thank you. Now I would like you to take a look at the program and tell me what you think as you go through it. Again, I am not testing you. We just want to understand your experience with the program. I am interested in all your reactions, positive and negative. Please think out loud. The more you talk as you go, the more we will understand about what works and what needs to be changed.

Tasks and Questions:

## PRECASE

## Facilitator's Dialogue:

After logging into PBLCloud, the following screen appears.

1. Please explain what comes to mind when you see this screen.
2. From here, how would you start the test case?
3. What did you find confusing or frustrating?
4. Do you have any other comments about this page?

Introduction to the Case

1. Please explain what comes to mind when you see this pop-up window.
2. Do you have any other comments about this page?



### Physical Examination

1. Now that you have completed taking a history, what would you do next to evaluate this patient?
2. Please explain what you see on this new screen?
3. If the tester doesn't focus on the Pulmonary Exam you can explain the following: You are concerned that this patient has a respiratory problem and thus would like to perform a Pulmonary Exam. Please locate the findings for your patient's "Pulmonary Exam."
4. What did you find confusing or frustrating?
5. Is there anything that you think should be included in the system?

### Review Labs/Study Results

1. Now that you have completed your historical and physical examination of the patient, what would you do next to evaluate this patient?
2. Please explain what you see on this new screen?



Assessment and Plan

1. Now that we have viewed the “View Lab/Study Results” section, what would you do next to in the evaluation of your patient?
  
  
  
  
  
  
  
  
  
  
2. Please explain what you expect to happen if you clicked the “Assessment and Plans” radio button.
  - a. Do you find this confusing?

Differential Diagnosis Activity

1. Please explain what you see this new pop-up screen.
  
  
  
  
  
  
  
  
  
  
2. Please complete the activity.
  
  
  
  
  
  
  
  
  
  
3. What did you find confusing or frustrating?
  
  
  
  
  
  
  
  
  
  
4. Do you have any other comments about this page?

Orders

1. Please explain what you see on this new pop-up screen.
  
2. Please order the necessary investigations or interventions for your patient. If the tester doesn't know what to order provide the following prompt: Please order a CBC and AP Neck Radiograph for your patient.
  
3. What did you find confusing or frustrating?
  
4. Is there anything that you think should be included in the system?
  - a. Is there another way of organizing this information that you think would be helpful?

Learning Exercises

## Facilitator's Dialogue

This pop-up window contains a couple of multiple-choice questions that learners have to complete before moving forward with the case.

1. Please go through and answer the multiple-choice questions. When you have completed all the questions please click "Save Answer".
2. Please explain what you see on the pop-up window after saving your answer.
3. Do you have any other comments about this page?

## TIME POINT 1: COMPLETED

1. What did you think about the pop-up windows that you just clicked through during the "Assessment and Plan" section?
2. Were there too many pop-up windows?
3. Do you have any other comments about this page?

## TIME POINT 2: START

## Facilitator's Dialogue

We have now come to the end of the first time point of this case. You will be taken to the next time point. Before entering back into the case please review the pop-up window on screen.

### New Time Point Introduction

1. Please explain what you see on the screen.
  
2. Do you think this pop-up window is useful?

### Facilitator's Dialogue

For this initial version of this case, the content of the historical interview is the same as in time point 1, though is you have more questions that you would like to ask you may do so now.

### Physical Examination

1. Please review the Pulmonary Exam for your patient.

### View Lab/Study Results

1. If the tester is having difficulty with this task you can provide the following prompts. The New Time Point Introduction explained that we have lab and imaging results back for review
  - a. Please locate your the available labs for your patient.
  
  - b. Please locate the reference range for those labs.
  
  - c. Please locate the available imaging for your patient.
    - i. What do you think of the two separate tabs for "Media" and "View Report".

Orders

1. Before we order those interventions let's go forward with the case without addressing the patient's respiratory distress.
  - a. What do you expect will happen to the patient?

Prompt System

1. Do you think that this hint feature is useful?
2. Is there anything else that you think should be included in the hint?

## Orders continued

2. Based on your differential diagnosis please enter in the appropriate orders to manage your patient's condition.
3. If the tester is having a difficult time you can provide the following prompt: There is a high index of suspicion that your patient has croup and the mainstay of treatment involves the administration of dexamethasone PO and racemic epinephrine.
4. Please order dexamethasone PO and racemic epinephrine for your patient.
  - a. Was this confusing or frustrating? If so, why?
  - b. Do you have any other comments about this page?

TIME POINT 2: COMPLETED



## TIME POINT 3: COMPLETED

## TIME POINT 4: START

## Facilitator's Dialogue

Now we will go through the next time point in the same manner as we did Time points 1 and 2.

## 1. Comments:

- a. Historical Exam
  
- b. Physical Exam
  
- c. View Lab/Study Results
  
- d. Differential Diagnosis
  
- e. Orders
  
- f. Learning Exercises

## TIME POINT 4: COMPLETED

## Facilitator's Dialogue:

You have now completed the croup case I have a short survey for you to complete and a few opened questions for you to answer out aloud.

**APPENDIX II**System Usability Scale:

Directions: When a SUS is used, participants are asked to score the following 10 items with one of five responses that range from Strongly Agree to Strongly Disagree:

1. I think that I would like to use the PBLCloud patient case simulator frequently.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree
2. I found the PBLCloud patient case simulator unnecessarily complex.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree
3. I thought the PBLCloud patient case simulator was easy to use.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree
4. I think that I would need the support of a technical person to be able to use the PBLCloud patient case simulator.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree
5. I found the various functions in the PBLCloud patient case simulator were well integrated.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree
6. I thought there was too much inconsistency in the PBLCloud patient case simulator.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree
7. I would imagine that most people would learn to use the PBLCloud patient case simulator very quickly.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree
8. I found the PBLCloud patient case simulator very cumbersome to use.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree



9. I felt very confident using the PBLCloud patient case simulator.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree
10. I needed to learn a lot of things before I could get going with the PBLCloud patient case simulator.
  - a. Strongly Disagree
  - b. Disagree
  - c. Neutral
  - d. Agree
  - e. Strongly Agree

Scoring: For odd items: subtract one from the user response. For even-numbered items: subtract the user responses from 5. This scales all values from 0 to 4 (with four being the most positive response). Add up the converted responses for each user and multiply that total by 2.5. This converts the range of possible values from 0 to 100 instead of from 0 to 40. Average across all users.

**APPENDIX III**Open-Ended Questionnaire:

1. On a scale from 1 to 5, how useful do you think the PBL Cloud patient case simulator is?
  1. not at all useful
  2. not useful
  3. neutral
  4. useful
  5. very useful
2. What type of medical trainee do you think the PBLCloud patient case simulator would be appropriate for?
3. Which part of the PBLCloud patient case simulator do you think is most useful? Why?
4. What part of the PBLCloud patient case simulator is least useful? Why?
5. Rate how interested you would be in using PBLCloud patient case simulator in the future?
  - 1 Very uninterested
  - 2 Uninterested
  - 3 Neutral
  - 4 Interested
  - 5 Very interested
  - a. Please explain.
6. Please name 1 – 2 aspects/features that you liked about the PBLCloud case simulator. Please explain.
7. Please name 1 – 2 aspects/features that you disliked about the PBLCloud case simulator. Please explain.
8. If you have used other online patient case simulators, how did this compare to the ones you have used?
9. Was anything missing from the PBLCloud patient case simulator or that you would like to see included?

10. On a scale from 1 to 5, how would you rate your enjoyment of the case simulation?

1 not at all enjoyable

2 not enjoyable

3 neutral

4 enjoyable

5 very enjoyable

a. Please explain