In search of a method to assess dispositional behaviours: The case of Otago Virtual Hospital

Swee-Kin Loke, Phil Blyth and Judith Swan
University of Otago

While the potentials of virtual worlds to support experiential learning in medical education are well documented, assessment of student learning within these environments is relatively scarce and often incongruent. In this article, a conceptual framework is proposed for formatively assessing dispositional behaviours in scenario-based learning within a virtual world. The framework was devised for use with medical students playing the roles of junior doctors as they solve open-ended clinical cases within an environment called the *Otago Virtual Hospital*. Drawing upon Perkins, Jay and Tishman's (1993) dispositional theory of thinking, it is proposed that the assessment of dispositional behaviours in scenario-based activities can be carried out by measuring the number of times students either seize or miss an opportunity to engage in a particular dispositional behaviour. The approach can potentially also be used for assessing scenario-based learning in other disciplines (e.g. law, business, military).

Introduction

Debates around what makes a good doctor typically acknowledge the importance of medical knowledge, and then point out how medical education has neglected essential humanistic qualities such as responsibility and compassion (Borleffs, 2009; Fones, Hoek & Gan, 1998). Indeed, the move to revise the medical curriculum at the University of Otago in New Zealand stemmed from a need to address an overemphasis on factual content and an inadequate focus on professional attitudes and communication skills (Schwartz, Loten & Miller, 1999). Such tensions between 'hard' and 'soft' skills are also apparent in other disciplines such as engineering (Kumar & Hsiao, 2007).

The imbalance in emphasis is, in a way, understandable given the relative ease of teaching and assessing factual content. How, then, do medical students learn to become knowledgeable and skilful as well as compassionate? In other words, how do they develop their holistic professional identities? For example, the Medical Council of New Zealand (2008) recommends that doctors keep their professional knowledge and skills up to date while upholding values such as respect and integrity. In order to increase its focus on professional values and attitudes, the curriculum at the Otago School of Medicine was revised in 2008 so that students are exposed to professional practice (e.g. interviewing real patients) at an earlier stage of their program (University of Otago, 2011). Lave and Wenger (1991) also suggest that participation in professional activities may be a means by which novices can become legitimate members of professional communities. Adopting a similar orientation, our work involves student

participation in the form of role-playing as junior doctors/house surgeons within the *Otago Virtual Hospital* (OVH).

The OVH is a three-dimensional virtual environment in the form of a hospital in which medical students solve realistic, open-ended clinical cases as junior doctors. These cases were written by us, the authors of the present article, who have backgrounds in education, medicine and nursing, and are all currently practising within our fields. Each scenario is drawn from real-life events and aims to reflect the actual practices of a New Zealand emergency department. Through their avatars, students are able to move around the hospital, communicate with patients and peers via text chat (see Figure 1 and the two-minute video at http://bit.ly/xZ0Net), examine patients, order laboratory and radiology tests, check the results of those tests, prescribe from a range of medicines, and write patient admission/ discharge/ handover notes. We chose to use text chat because, from our experience, turn taking using voice chat (lacking the 'is typing' feature signaling that an interlocutor is about to send a message) typically resulted in too many overlapping utterances. The OVH is built on the *OpenSim*-based New Zealand Virtual World Grid (http://www.nzvwg.org/).

Figure 1: Medical students role-playing junior doctors taking patient history (Group 4)

Such student performance, even though situated in a virtual world, has the potential to support the development of a real-world professional identity (Gee, 2007; Shaffer, 2006). The identity formation is driven by students' "embodied practice" (Taylor, 2002, p. 60) mediated by their avatars.

In this article, we focus on the assessment of students' humanistic qualities. While the potentials of virtual worlds to support experiential learning in medical education are well documented (Boulos, Hetherington & Wheeler, 2007; Wiecha, Heyden, Sternthal & Merialdi, 2010), assessment of student learning within these environments is

relatively scarce and often incongruent (Hew & Cheung, 2010). Currently, in the OVH the measurement of the scenario's outcomes (see the Appendix to this article) is carried out using a post-scenario self-appraisal form (http://bit.ly/nF5BxC) and a patient feedback form (http://bit.ly/qmxVe4). The former focuses on confidence in treatment and discrete items in medical care and patient care. It is timely to focus on students' in-world performance of compassion and other dispositional behaviours that are currently disregarded in our assessment. Broadening our assessment is important because therein lies the distinct benefit of virtual worlds in education - in post-scenario interviews, three of our students identified 'global decision making' (Group 1), 'looking at the big picture' (Group 2) and how the experience 'ties it all together' (Group 4) as key learning points.

We will first outline current trends in the assessment of scenario-based learning before making a case to re-conceptualise medical education as the fostering of doctorly dispositions. We then describe our conceptual framework for assessing dispositional behaviours followed by an empirical illustration based on a case study in the OVH. Lastly, limitations of the framework and future work will be discussed.

Assessment of scenario-based learning

The learning design underpinning the OVH can best be categorised as scenario-based learning (SBL). Naidu (2007) defines SBL as "a pedagogical design in which an authentic or contrived scenario forms the basis of all learning, teaching and assessment activities" (p. 251). SBL, grounded in situative and constructivist learning theories, is similar to problem-based learning (PBL) and case-based learning (CBL) (see Naidu, 2007 for more details on how the three instructional designs differ). Our choice of SBL was motivated by our desire to foreground the centrality of our scenarios, their real time co-construction by participants, the dynamic follow-through on students' ongoing decisions, and students taking on their professional identities.

The introduction of SBL and similar instructional designs in medical education has evoked at least two negative responses. Firstly, because such instructional designs foreground clinical practice over scientific knowledge (perceived to be lacking utility), departments of basic science have often viewed them with anxiety and antipathy (Sweeney, 1999). Secondly, Kirschner, Sweller and Clark (2006) contend that such "minimally guided approach[es]" (p. 75) are effective only when learners have a sufficiently high prior knowledge. In response, we position the learning experiences offered by the OVH as complementary to existing lectures and tutorials, supporting the belief that clinical practice should be guided by a sound understanding of basic science. Also, we target medical students with a higher prior knowledge of clinical practice (namely fourth- and fifth-year students).

The assessment of SBL (and similar designs) in medical education poses significant challenges because such complex and holistic learning experiences are extremely difficult to describe and evaluate (Webster-Wright, 2009). Indeed, the assessment of clinical competence (demonstrated in realistic scenarios) has been thought to be "one of the most difficult tasks facing medical education" (Howley, 2004, p. 286).

Many studies on SBL lack validity in terms of their assessment because they do not measure the qualities that students are meant to develop. In a literature review on CBL in medical education, Williams (2005) found that although the integration of

knowledge and practice was identified as a key learning outcome in the studies reviewed, all but one measured only student attitudes towards CBL models. Another literature review on virtual patients (computer software agents simulating the symptoms, behaviours and reactions of real patients) reported an unexpectedly low number of studies measuring clinical reasoning as an outcome (Cook & Triola, 2009). One example of such a study was that done by Naidu, Oliver and Koronios (1999), in which although the learning outcome was proclaimed to be clinical decision making in nursing practice, only the students' perceived utility of the multimedia courseware was measured.

When studies on SBL do take clinical competence into account, typically they focus on basic and narrow clinical skills (Howley, 2004; McGaghie, Issenberg, Petrusa & Scalese, 2010). For example, the objective structured clinical examination (OSCE) is a form of scenario-based assessment that is widely used in medical education, and educators have been exploiting virtual reality technologies to test OSCE-specific skills such as interviewing (Johnsen, Raij, Stevens, Lind & Lok, 2007). However, OSCEs typically test students only on "isolated aspects of the clinical encounter" (Smee, 2003, p. 705), and the task-specific checklists used for assessment reduce holistic doctor-patient interactions into a list of discrete actions.

Through our own literature searches, we located two projects that put into practice an assessment method of SBL which is more valid. In the first of these, Kneebone et al. (2005) based their assessment (e.g. of patient communication) on real time observations and recordings of students' in-scenario performance. In the second, Youngblood et al. (2008) designed a comprehensive learning experience that simulated crisis management in a virtual emergency department from start to finish (not a part-task performance of a discrete clinical procedure such as wound closure). These two features were incorporated into our proposed framework.

In terms of efforts in making assessment more holistic, we agree with Rogers' (2011) orientation towards developing global team-oriented decision-making skills. However, while he demonstrated that students can productively learn such collaborative clinical problem solving in virtual worlds, he did not suggest how to go about measuring the preferred or targeted forms of decision making.

In light of these gaps, we seek to design a formative assessment method that: (i) features a higher level of validity in terms of alignment between intended learning outcomes and assessment; (ii) gives students a sense of their own progress following iterative runs in the OVH; and (iii) takes students' in-scenario performance into consideration. The importance of taking into account students' in-world performance - their 'lived experiences' - is threefold. Firstly, adopting the theory of situated cognition, we believe that the ways in which human beings think and act are inherently coupled with their environment (Brown, Collins & Duguid, 1989; Clancey, 1997). Secondly, decision making in medical practice is a continuous process (Hatala, Norman & Brooks, 1999), and the process of decision making is perhaps as important as the actions taken. The OVH offers us a rare opportunity to observe students' decision making - potentially developing over iterative cycles - in fine detail. Thirdly, students' performance of compassion and other dispositional behaviours (our learning outcomes) can most fruitfully be observed and understood from their in-world participation. But how do we assess such holistic learning experiences meaningfully?

Fostering dispositions through enculturation

To answer the question above, more clarity is needed regarding what we intend for our students to learn, and how we expect them to develop these qualities. While we strive to develop medical students' professional identities more fully (as alluded to in the Introduction), we struggle to get a handle on such a holistic construct: how do we ascertain that a student has become a more compassionate doctor? More fundamentally, how do we even determine that a student's action in a particular instance is 'compassionate'? After all, there is no single and easily definable act performed in an emergency department that would be considered 'compassionate' at *all* times, and whose frequency we could reliably count in order to determine a doctor's degree of compassion. For example, a doctor who draws the curtains between beds (e.g. to protect patient privacy) more frequently than another should not automatically be considered a more compassionate doctor. The context-appropriateness of the behaviour is key to the meaning we attribute to it.

Conceptualising 'compassion' as a disposition offers a way forward. Perkins, Jay and Tishman's (1993) dispositional theory of thinking is a particularly suitable lens through which to examine SBL in the OVH. We sought to describe good thinking in real-world situations (as compared to controlled laboratory settings). Perkins et al.'s theory holds that human intelligence is underpinned by dispositions comprising three subcomponents: *sensitivity* (e.g. likelihood of noticing the need to draw curtains), *inclination* (e.g. motivation to draw curtains) and *ability* (e.g. physical capability to draw curtains). It provides a suitable lens because good thinking in a real-world emergency department is often more dependent on dispositions (is this a *good time* to draw the curtains?) than on pure ability:

> Everyday contexts present a wilderness of vaguely marked and ill-defined occasions
> for thoughtful engagement. Opportunities for investing one's intelligence must be
> detected. (Perkins, Tishman, Ritchhart, Donis & Andrade, 2000, p. 270)

Whether a medical student detects an opportunity for showing compassion depends on what the student values, because values "influence what we notice within the total range of the presented data" (Ferré, 1996, p. 14). Indeed, the lack of sensitivity is often the "principal bottleneck" (Perkins et al., 2000, p. 276) in intellectual performance, and has been suggested to occur in medical student practice (Coderre, Wright & McLaughlin, 2010; Wilkinson & Harris, 2002). In our case study, we also found some evidence of the lack of inclination resulting in poorer performance in the OVH: in a post-scenario interview, one of our medical students (from Group 1) expressed that while he did notice the patient's thirst during the scenario (i.e. presence of sensitivity), he regretted not pursuing the possible link between dehydration and urinary tract infection at that point in time (i.e. lack of inclination). We therefore believe a dispositional view of human intelligence offers a more accurate explanation of the intellectual performance required in an emergency department than an abilities-centric view.

The theory is also suitable for our purposes because the humanistic qualities that we intend for students to develop are dispositional in nature. Using Ryle's (1949) analogy, a doctor is compassionate just as glass is brittle - a compassionate doctor has the tendency to behave in a compassionate manner when the opportunity arises; brittle glass has the tendency to shatter when struck. Both qualities are *propensities* that are likely to manifest themselves when certain conditions are present. Just as we do not

expect glass to manifest its brittleness under all conditions, we do not expect a compassionate doctor to perform compassionate acts all of the time.

Lastly, the theory is suitable because our identity or self is underpinned by our dispositions: "Our doings and undergoings (... dispositions, habits, inclinations, and tendencies to action and inaction) are, after all, what characteristically constitutes us as individuals" (Rescher, 1996, p. 107). In our work, we seek to develop our medical students' professional identities, and the notion of dispositions offer a useful handle with which to make sense of their developing identities.

Framed within the dispositonal theory of thinking, medical education can be effectively reconceptualised as the fostering of doctorly dispositions. Moving beyond pure cognitive abilities, learning to become a doctor can be thought of as the 'tuning' of dispositions such as doctorly compassion and responsibility. We use the verb 'tuning' to stress that the performance of dispositional behaviours should be optimal, rather than maximal - making small talk with the patient may be considered compassionate in most cases, but possibly not to the extent of consuming two hours of everybody's time! From their study, Perkins et al. (2000) reported that dispositions are "stable traits" (p. 269) that can help explain intellectual performance in everyday circumstances, a stability which holds promise for the transfer of learning from university to workplace settings.

In this article, we focus on and use as an example the assessment of two dispositions that are important for doctors: compassion and open-mindedness. A range of other dispositions are relevant in the medical profession (e.g. integrity, cooperation/teamwork), but we chose to limit the scope of the present study by including one disposition related to the patient's general well-being (compassion) and one directly related to medical diagnosis and treatment (open-mindedness). In terms of the former, *sensitivity* could mean alertness to a possible situation of personal distress or discomfort according to the patient; *inclination* the motivation to show concern and humanity; and *ability* the actual capacity to demonstrate such care and concern. Cahill (2004) highlights and illustrates the importance of compassion in clinical practice by using real-life examples of doctors lacking compassion, including one who asked a patient personal questions in a busy waiting room. In the case of open-mindedness, *sensitivity* could refer to alertness to other reasons for the patient's presentation and awareness of rushed decisions; *inclination* the motivation to ask beyond the obvious and to explore a range of options and ideas; *ability* the actual capacity to identify and question one's assumptions and to probe further. Conti (2005) highlights the importance of open-mindedness in clinical practice by pointing out how a good doctor should "search for secondary medical problems even if there is an obvious cause of the patient's illness" (p. 496).

In fact, we gained a sense of the validity of dispositions in our work when one medical student (from Group 1) identified the disposition of 'gentle assertiveness' as a key learning point during a post-scenario interview. The student had to wake the patient up to perform a physical examination, all the while managing the demands of time pressure yet simultaneously being compassionate.

With the emphasis on dispositional knowledge here, we need also to account for the relative importance of propositional knowledge (e.g. the knowledge that bacteria entering the urethra can cause infection in the urinary tract). The debate around this

dichotomy is vast, and the philosopher Damschen (2009) argues that propositional knowledge-that is a sub-species of dispositional knowledge-how. Distinguishing know-that from know-how, Tishman, Jay and Perkins (1993) also state that "Sensitivity requires not just having relevant guidelines in [mental] storage but *acting* on them in relatively uncued conditions" (p. 149, emphasis in original). The tuning of medical students' dispositions to act in a doctorly manner in relatively uncued scenarios forms the basis of our work, and we previously published early evidence that solving an open-ended case in the OVH required students to notice, record and integrate significant elements of the case by themselves (Blyth, Loke & Swan, 2010).

Even if we accept the fostering of dispositions as an educational goal, how does one learn the desired dispositions? Tishman et al. (1993) believe that thinking dispositions are most effectively learnt through enculturation rather than transmission. This view is resonant with that of Lave and Wenger's (1991), who see professional selves as being formed by partaking in professional activities. We believe that, by role-playing as junior doctors within authentic scenarios in the OVH, medical students are being enculturated into their professional practice.

Conceptual framework for assessing dispositional behaviours

We now present our proposed conceptual framework for assessing dispositional behaviours within the OVH, drawing inspiration from Perkins et al. (2000), who maintain that in real-world situations, "Opportunities for investing one's intelligence must be detected" (p. 270). The main thrust of this article is that dispositional behaviours in scenario-based activities can be assessed by retrospectively examining the number of times students either seized or missed an opportunity to engage in a particular dispositional behaviour (e.g. the open-mindedness to check for heart attack upon learning about the patient's irregular heartbeat). It is important to add that students in the OVH are not purely passive observers of an unfolding scenario, and that they can also *create* opportunities for themselves (e.g. by initiating alternative lines of questioning to search for secondary medical problems related to irregular heartbeat). This is characteristic of the co-construction of the scenario by both the patient and doctors in the OVH.

Given that our goal is to assess the performance of medical students in realistic professional settings, we will emphasise group over individual assessment. Rarely do doctors act alone or are solely responsible for patient outcomes; Hays et al. (2002) identified teamwork as a gap in current performance assessment, recommending that the overall performance of a team be used as an assessment indicator. In fact, cognition in a real-world emergency department - "cognition in the wild" (Hutchins, 1996, p. xiii) - can be more accurately described as cognition that is *distributed* across the whole environment, including doctors, patients, instruments, and hospital policies and protocols.

Applying the framework to the OVH

All fourth- and fifth-year medical students (from a six-year undergraduate program) were invited via email and an in-class presentation to participate in the project. To take part, students had to register their interest on a website (http://ovh.otago.ac.nz/), read the information sheet and give their consent for the results of the project to be published. Student anonymity is maintained by reference to the avatar name only

when reporting data and results. These documents and processes were part of the successful application for ethical approval by the University's Faculty of Medicine.

A case study was conducted with 11 medical students (including 5 fourth-year students, 5 fifth-year students and 1 sixth-year student), who worked on the same clinical case in five groups (Groups 1-5) during five separate runs, each lasting between 70 and 90 minutes. The clinical case involved a female patient by the name of "Mrs Gertrude Macfarlane", in her mid-70s, whose neighbour had found to be increasingly forgetful (e.g. not feeding the cat) and generally unwell (e.g. feverish).

Each group of students took on the role of junior doctors, while Mrs Macfarlane was played by the same faculty member across all of the groups. This faculty member worked from a pre-written script (http://bit.ly/o8vUDQ) incorporating some elements that had to be shared accurately (e.g. being forgetful and confused) but allowing some flexibility as to how much confusion to portray, for example.

The in-world activity of all groups was recorded using a screen-recording software program. All students' in-scenario text chat and written notes were captured. Upon conclusion of the scenario, the students and faculty participated in a debriefing session (to discuss decisions made and points highlighted), followed by a post-scenario interview.

The empirical illustration of our conceptual framework that follows involves in-world data from Groups 4 and 5 only. Not all data are included because the purpose here is mainly illustrative, with our aim being to demonstrate how two in-world performances can be compared in terms of a given disposition. The contrast is meant to show readers what a desired in-world performance looks like. It is noteworthy that although we are comparing two different groups' performances here, our proposed assessment is criterion-referenced (determining whether a group performed well or poorly on a given task) and not norm-referenced (determining whether a group performed better or worse than other groups). We would have preferred to compare two performances by the *same* group for our empirical illustration, but all 11 students have only engaged in the OVH once to date. In Tables 1 to 3, we have identified comparable excerpts illustrating how the two groups behaved differently in terms of compassion and open-mindedness (when a similar opportunity arose during the scenario). We have listed excerpts with the most apparent contrast between groups in Table 1 (compassion) and Table 2 (open-mindedness).

In the excerpts shown in Table 1, compassion was exhibited more readily in Group 4's performance and less readily in Group 5's, given similar opportunities that were available to both groups. Another example of an instance in which compassion was shown in differing degrees, but that does not appear in the table, was when Group 3 discharged the patient without adequately ascertaining the level of social support she had at home (captured in the video clip at http://bit.ly/au7jWH). All four of the other groups admitted the patient into hospital for various reasons, including inadequate social support at home.

During the post-scenario debrief, one student in Group 5 did express regret for not having shown compassion to his patient who made repeated requests to lie on the bed: "If you miss the big cues, there's no excuse for that". This statement of his suggests that the selected excerpt was valid in gauging his group's performance in terms of

compassion, and also that students should be held accountable only to the more obvious opportunities (i.e. the "big cues").

Table 1: In-scenario performance of compassion

	Group 4	Group 5
Characters	• *Junior doctors:* Smudge Williams and Craig Abode • *Patient:* Gertrude Macfarlane • *Patient's daughter:* Leigh Macfarlane	• *Junior doctors:* James King and Dr J. Yellow • *Patient:* Gertrude Macfarlane • *Patient's daughter:* Leigh Macfarlane
Compassion regarding patient's modesty/ privacy (during physical examin-ation)	[14:40] Smudge: **Would you like Leigh to stay in here with you**? [14:40] Gertrude: I don't mind. [14:40] Craig: Dr. Williams her JVP is 0, tongue looks dry, and no ankle oedema [14:40] Smudge: Ok **I will shut the curtain** [14:40] Gertrude: If Leigh needs to get away to work I'll be ok with you. Thank you for doing that	[14:09] James: Mrs Macfarlane, I'd like to take a look at your chest and your tummy, **is that ok**? [14:09] Gertrude: oh yes Ok, that's fine (doctor proceeds to examine patient **without drawing curtains**) [14:10] Gertrude: as long as your hands are warm! and you close the curtain. please [14:10] James: Usually? II. sure [14:10] Gertrude: Thank you
Compassion regarding patient's fatigue	[14:45] Gertrude: I'm feeling a bit tired. could i have a lie down please [14:45] Smudge: **Yes sure there is a bed here**, we will get you moved to the ward as soon as we can [14:45] Gertrude: oh thank you	[13:45] Gertrude: could i have a lie down [13:45] DrJ: **Dr James was asking about your breathing**? [13:45] Gertrude: i am feeling a bit tired just now [13:45] Gertrude: oh my breathing. sorry i missed your question [13:45] James: Have you had any problems with your breathing? (...) [13:46] James: What about any other problems? any chest pain? [13:46] Leigh: Doc, could my mum have a lie down? [13:46] Gertrude: no NO chest pain. I know about chest pain [13:46] DrJ: **I noticed you had a drink** and you were still thirsty? Have you been quite thirsty recently? [13:47] Gertrude: yes, i am still thirsty (...) [13:48] Leigh: Doc, my mum asked for a lie down minutes ago. Could she, please? [13:48] DrJ: **ok. And how have your waterworks been**? [13:48] James: **Would you like to pop on the bed, we will need to have a look at you soon**

Looking at the first pair of excerpts in Table 2, both groups displayed open-mindedness towards the patient's reason for being in hospital, with both starting off wide in their questioning, though Group 4 narrowed the discussion down more quickly (on the fever) than Group 5 did. In the second pair of excerpts in Table 2, both groups again showed evidence of open-mindedness, this time towards the range of

possible causes of the symptoms the patient was experiencing. They both enquired about a number of possible causes within the conversation; however, Group 4 questioned the patient on a slightly broader set of causes. Yet another example (not depicted in the table) in which open-mindedness was shown in differing degrees was when Group 1 asked whether the patient had had a fall. None of the other four groups asked this question to exclude the possibility of trauma to the head being one of the causes of the patient's confused state.

Table 2: In-scenario performance of open-mindedness

	Group 4	Group 5
Characters	• *Junior doctors:* Smudge Williams and Craig Abode • *Patient:* Gertrude Macfarlane • *Patient's daughter:* Leigh Macfarlane	• *Junior doctors:* James King and DrJ Yellow • *Patient:* Gertrude Macfarlane • *Patient's daughter:* Leigh Macfarlane
Open-mindedness towards reason for being in hospital	[13:23] Craig: So Mrs. Macfarlane **what's brought you in today**? [13:23] Smudge: **How can we help you today**? [13:24] Gertrude: Well, my neighbour brought me here, because she said i was not well. I'm not sure, I feel fine at the moment, but i'm a bit thirsty. I think i missed breakfast [13:25] Craig: Okay. **Do you feel like you might have a fever**?	[13:35] DrJ: So **can you tell us why you think you are in here this morning**? [13:36] Gertrude: This morning, well, my neighbour came over to see me, as usual and she told me i called the wrong name, and hadn't fed the cat; my poor ossie, and then she brought me here. Is this the hospital? [13:37] James: yes [13:37] Gertrude: Oh good (...) [13:37] James: Dr Yellow and I will see if we can help you out [13:37] DrJ: **Have you been feeling unwell in any way**? (...) [13:42] Leigh: She HAS been feverish [13:42] DrJ: **Can you tell us how long she has been feverish for**?
Open-mindedness to variety of causes	[13:25] Smudge: How long have you felt thirsty for? (...) [13:26] Smudge: When did you last have something to eat or drink? (...) [13:33] Smudge: What do you mean by a bit puffed? Did you feel breathless? (...) [13:35] Smudge: How have you been sleeping? (...) [13:37] Smudge: Have you noticed any changes in your weight? (...) [13:39] Smudge: What pills do you normally take in the morning? (...) [13:45] Craig: Mrs. Macfarlane have you been going to the toilet more than normal recently?	[13:44] James: Have you had any other problems? how has your breathing been? (...) [13:46] James: What about any other problems? any chest pain? (...) [13:48] DrJ: ok. And how have your waterworks been? (...) [13:52] DrJ: ok. have you had any tummy pain? (...) [13:54] DrJ: ok. Have you had a cough at all? (...) [13:57] James: you mentioned you knew all about chest pain, what medical probs have you had in the past? (...) [13:59] DrJ: What pills are you taking? (...)

(...) [13:46] Craig: Okay. Dr. Williams have you seen the glucose meter? (...) [14:26] Craig: Mrs Macfarlane you mentioned earlier that when you forget to take your betaloc you get palpitations, is that right? (...) [14:28] Craig: Do you ever feel chest pain? (...) [14:32] Smudge: ok, Gertrude. Have you noticed any stomach pain, or back pain? (...) [14:35] Craig: Do you drink alcohol?	[14:02] DrJ: We can tell you after we have looked at your, mum. I think **I'm finished with qns**.

To conclude our empirical illustration of the proposed conceptual framework, we present in Table 3 our complete assessment of Group 5's in-scenario performance in terms of compassion and open-mindedness. Referring to the scenario outcomes (Appendix 1), each of us (authors) independently reviewed the entire text chat log (http://bit.ly/uKhIXb), identified excerpts where the two dispositions were exhibited (in specific actions, statements or interactions) and allocated a score (+ or -) to each of the excerpts. Because our proposed assessment is criterion-referenced, we judged Group 5's performance using a binary scale (+/-) rather than a numeric rating scale (e.g. 1-5). We then discussed our individual scores and reached an agreement on the final score. Our discussion to this end typically revolved around the appropriateness or inappropriateness of different professional behaviours. Our challenge in assessing professional appropriateness is similar to that presently faced by OSCE examiners (Hodges, 2003).

Table 3: Assessment of Group 5's performance in terms of desired dispositions

Compassion	Score	Open-mindedness	Score
[13:45] Failed to respond to patient's requests to have a lie down on three occasions	-	[13:35], [13:37], [13:38] and [13:39] Wide starting point to interaction, with focus on patient perspective	+
[13:50] Showed empathy by building on patient's joke regarding urine	+	[13:35], [13:37], [13:41] and [13:59] Allowed ongoing intake of oral fluids, when they could be a contributing cause	-
[13:54] and [14:28] Acknowledged relative's hurry to leave for work and explained need to be thorough in their examination	+	[13:40], [13:41] and [13:58] Used diverse sources/strategies for information	+
[14:00] Reassured patient that the hospital could give her the medications she had missed	+	[13:44] and [13:46] Wide question ("any other problems") but narrowed by cueing (e.g. "breathing")	-
[14:00] Did not empathise with patient's inability to use blister packs	-	[13:48], [13:52], [13:53] and [13:54] Enquired about other symptoms which may affect diagnosis	+
[14:07] Acknowledged patient's thirst while insisting on her drinking water instead of coffee	+	[13:57] Wide question ("What other medical probs have you had in the past?") without narrowing qualifier	+

[14:10] Forgot to close curtains before examining patient	-	[14:03], [14:04] and [14:05] Discussed diverse investigations and examinations	+
[14:25] and [14:31] Left patient abruptly without taking leave	-	[14:42], [15:09] and [15:13] Recognised elements of case that have been forgotten or overlooked	+
[14:41] Decided to admit patient to hospital because patient cannot take care of herself at home	+	Did not enquire about medication allergies	-
		Conducted multiple observations (pulse, blood pressure)	+

Discussion

In devising our conceptual framework and applying it to the OVH, we attempted to maximise what Brown and Knight (1994) deem to be the three key characteristics or properties of sound assessment: validity, reliability and efficiency. Each of these is discussed in turn below.

Validity

Our proposed assessment method is valid at two levels. Firstly, our conceptual framework targets doctorly dispositions that are valued in medical practice (Cahill, 2004; Conti, 2005). Secondly, the framework assesses medical students based explicitly on their dispositional behaviours within a simulated scenario (e.g. compared to a multiple-choice question [MCQ] test on desired dispositions, which would lack validity). It is also noteworthy that the proposed assessment method draws on students' actual in-scenario performance (e.g. not self-reports on their perceived level of performance). Lastly, the clinical case and the virtual hospital environment were designed by medical professionals and judged to be 'realistic' by all 11 students and two other clinical medical educators (qualified doctors who helped us trial the scenario by role-playing as junior doctors).

One area in which the assessment's validity may be questioned is in the differences between performing as a doctor in a virtual world and in real life. For example, the two students in Group 5 expressed that showing compassion in a virtual world was challenging because facial expressions and voice intonations are not as intuitively performed. We draw once again on Gee's (2007) work on role-playing video games, and reiterate that the main premise of learning to become a doctor in a virtual world is that the OVH is a place "where people can learn to situate meanings through embodied experiences in a complex semiotic domain" (p. 26). The students' experience in a virtual world will always suffer from some loss in fidelity compared to their workplace experience; the OVH and current clinical placements should hence be positioned as complementary learning experiences.

Reliability

Reliability of assessment refers to the extent to which the assessment is consistent - for instance, the likelihood that a student retaking the assessment will obtain similar results. The main threat to the reliability of our conceptual framework stems from the

scenario being open-ended and co-constructed 'on the fly' by the doctors and the patient. The only constants across the runs are the admission/triage form (http://bit.ly/rbxI50) and the patient's script (http://bit.ly/o8vUDQ). Hence there is no guarantee that if the same group of students were to re-engage in the same scenario they would yield similar results, because of the many confounding variables that exist (e.g. number and types of opportunities that might emerge from possible lines of questioning). This is the main reason why our framework is more suitable for ongoing formative assessment of students than for end of course summative assessment.

Standardising all the opportunities given for the doctors (students) to engage in desired dispositional behaviours (e.g. making sure that the patient mentions her existing medications at a fixed time regardless of whether the doctor asks for the information) would be one method of increasing reliability. However, doing this might compromise validity because we aim to teach students to make decisions in "relatively uncued conditions" (Tishman et al., 1993, p. 149). One way in which we are addressing this issue is by featuring an unconscious multi-trauma patient in our next scenario (thus reducing the variability of patient-doctor conversation), but we nevertheless accept the interminable trade-off between validity and reliability.

Related to reliability is our desire to give students a sense of their own progress after iterative runs in the OVH. The relatively low reliability implies that a group that performed well in the first run may perform poorly in the second and/or subsequent runs of the same scenario (and vice versa). Drawing on Kolb's (1984) model of experiential learning, we believe that students' repeated participation and reflection (via post-scenario questionnaire and debrief) have the potential to enculturate them into the values and dispositions of medical practice. The reflection phase is particularly important to develop congruence between students' theories-in-use (worldview and values underpinning their actions) and their espoused theories (worldview and values they *believe* their actions are based on) (Argyris, 1980; Schön, 1987). However, we currently have no evidence to support (or refute) this claim because the 11 students have only engaged in the presented scenario once.

Efficiency

Compared to more commonly used assessment methods (e.g. MCQ tests, standardised observation checklists), our conceptual framework is also not as efficient. At the end of the scenario, assessors have to firstly identify the opportunities that arose in-scenario for the performance of particular dispositions (by examining the text chat log), and secondly judge the appropriateness of the students' behaviour on each occasion (i.e. whether they seized or missed the opportunity). Using these steps, we took an estimated average of 20 minutes to assess a single group for compassion. To reduce the workload imposed on staff and to potentially deepen student learning, we intend to engage students in peer assessment of other groups' performance in the future. For example, a group of students could compare another group's text chat log with the scenario's outcomes of medical care (e.g. see Appendix 1) and assess the latter's performance in terms of the desired dispositions (e.g. see Table 3).

Conclusion

As educators, we focus on the fostering of dispositions because changing ways of thinking and acting are fundamental goals of education:

> The idea is that if we manage to foster the relevant dispositions in students, their thoughts, judgments, and actions will be *changed*. ... If dispositions are not understood as having the potential to cause, to 'bring about', particular sorts of thought and action, what reason would there be to think it educationally important to foster them? (Siegel, 1999, p. 214, emphasis in original)

In an attempt to foster and assess desired dispositional behaviours *validly* within the OVH, we designed a conceptual framework in which medical students are retrospectively assessed based on the number of times they either seized or missed an opportunity to engage in a particular dispositional behaviour. This conceptual framework can be usefully applied to scenario-based learning in virtual worlds in other professional areas and disciplines - for example, in legal education to assess ethical attributes (Maharg & Owen, 2007); in war gaming to assess military leadership (e.g. in maintaining troop morale) (Oswalt, 1993); and in management education to assess the cost-efficiency of a student's factory design (Pidd, 2004).

In this article, we have presented an empirical illustration of our conceptual framework. Further empirical research is needed to determine its construct validity - for example, by comparing medical students' and trained doctors' in-world performances, and by comparing successive runs in the OVH by identical students. We will also continue to assess other doctorly dispositions (e.g. responsibility, teamwork) using the conceptual framework, and attempt to apply it to a range of different scenario types.

References

Argyris, C. (1980). *Inner contradictions of rigorous research*. New York: Academic.

Blyth, P., Loke, S.-K. & Swan, J. (2010). Otago Virtual Hospital: Medical students learning to notice clinically salient features. In *Curriculum, technology & transformation for an unknown future. Proceedings ascilite Sydney 2010*. (pp. 108-112). http://www.ascilite.org.au/conferences/sydney10/procs/Blyth-concise.pdf

Borleffs, J. (2009). What is a good doctor? The impact of 'fruitful irrelevance' in medical education. *Medical Teacher*, 31(5), 447-448. http://dx.doi.org/10.1080/01421590802642545

Boulos, M. N. K., Hetherington, L. & Wheeler, S. (2007). *Second Life*: An overview of the potential of 3-D virtual worlds in medical and health education. *Health Information and Libraries Journal*, 24(4), 233-245. http://dx.doi.org/10.1111/j.1471-1842.2007.00733.x

Brown, J. S., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42. http://dx.doi.org/10.3102/0013189X018001032

Brown, S. & Knight, P. (1994). *Assessing learners in higher education*. London: Kogan Page.

Cahill, A. (2004). A tale of a few hospitals. *Australian Health Review*, 27(2), 1-4. http://dx.doi.org/10.1071/AH042720001

Clancey, W. J. (1997). *Situated cognition: On human knowledge and computer representations*. Cambridge, UK: Cambridge University Press.

Coderre, S., Wright, B. & McLaughlin, K. (2010). To think is good: Querying an initial hypothesis reduces diagnostic error in medical students. *Academic Medicine*, 85(7), 1125-1129. http://dx.doi.org/10.1097/ACM.0b013e3181e1b229

Conti, C. R. (2005). What makes a good doctor? *Clinical Cardiology*, 28(11), 496-498. http://dx.doi.org/10.1002/clc.4960281102

Cook, D. A. & Triola, M. M. (2009). Virtual patients: A critical literature review and proposed next steps. *Medical Education*, 43(4), 303-311. http://dx.doi.org/10.1111/j.1365-2923.2008.03286.x

Damschen, G. (2009). Dispositional knowledge-how versus propositional knowledge-that. In G. Damschen, R. Schnepf & K. R. Stuber (Eds), *Debating dispositions: Issues in metaphysics, epistemology and philosophy of mind* (pp. 278-295). Berlin: Walter de Gruyter. http://dx.doi.org/10.1515/9783110211825.278

Ferré, F. (1996). *Being and value: Toward a constructive postmodern metaphysics*. New York: SUNY Press.

Fones, C., Hoek, K. E. & Gan, G. L. (1998). What makes a good doctor: Defining the ideal end-product of medical education. *Academic Medicine*, 73(5), 571-572. http://dx.doi.org/10.1097/00001888-199805000-00029

Gee, J. P. (2007). *What video games have to teach us about learning and literacy* (2nd ed.). New York: Palgrave MacMillan.

Hatala, R., Norman, G. R. & Brooks, L. R. (1999). Impact of a clinical scenario on accuracy of electrocardiogram interpretation. *Journal of General Internal Medicine*, 14(2), 126-129. http://dx.doi.org/10.1046/j.1525-1497.1999.00298.x

Hays, R. B., Davies, H. A., Beard, J. D., Caldon, L. J. M., Farmer, E. A., Finucane, P. M., ... Sibbald, G. R. (2002). Selecting performance assessment methods for experienced physicians. *Medical Education*, 36(10), 910-917. http://dx.doi.org/10.1046/j.1365-2923.2002.01307.x

Hew, K. F. & Cheung, W. S. (2010). Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research. *British Journal of Educational Technology*, 41(1), 33-55. http://dx.doi.org/10.1111/j.1467-8535.2008.00900.x

Hodges, B. (2003). OSCE! Variations on a theme by Harden. *Medical Education*, 37(12), 1134-1140. http://dx.doi.org/10.1111/j.1365-2923.2003.01717.x

Howley, L. D. (2004). Performance assessment in medical education: Where we've been and where we're going. *Evaluation & the Health Professions*, 27(3), 285-303. http://dx.doi.org/10.1177/0163278704267044

Hutchins, E. (1996). *Cognition in the wild*. Cambridge, MA: MIT Press.

Johnsen, K., Raij, A., Stevens, A., Lind, D. S. & Lok, B. (2007). The validity of a virtual human experience for interpersonal skills education. In M. B. Rosson & D. J. Gilmore (Eds), *Proceedings of the 2007 SIGCHI Conference on Human Factors in Computing Systems* (pp. 1049-1058). New York: Association for Computing Machinery. http://dx.doi.org/10.1145/1240624.1240784

Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75-86. http://dx.doi.org/10.1207/s15326985ep4102_1

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Kneebone, R. L., Kidd, J., Nestel, D., Barnet, A., Lo, B., King, R., ... Brown, R. (2005). Blurring the boundaries: Scenario-based simulation in a clinical setting. *Medical Education*, 39(6), 580-587. http://dx.doi.org/10.1111/j.1365-2929.2005.02110.x

Kumar, S. & Hsiao, J. K. (2007). Engineers learn 'soft skills the hard way': Planting a seed of leadership in engineering classes. *Leadership and Management in Engineering*, 7(1), 18-23. http://dx.doi.org/10.1061/(ASCE)1532-6748(2007)7:1(18)

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Maharg, P. & Owen, M. (2007). Simulations, learning and the metaverse: Changing cultures in legal education. *Journal of Information, Law & Technology*, 2007(1). http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2007_1/maharg_owen

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R. & Scalese, R. J. (2010). A critical review of simulation-based medical education research: 2003-2009. *Medical Education*, 44(1), 50-63. http://dx.doi.org/10.1111/j.1365-2923.2009.03547.x

Medical Council of New Zealand (2008). *Good medical practice: A guide for doctors*. Wellington, New Zealand: MCNZ. [viewed 12 Dec 2011] http://www.mcnz.org.nz/portals/0/guidance/goodmedpractice.pdf

Naidu, S. (2007). Instructional designs for optimal learning. In M. G. Moore (Ed.), *Handbook of distance education* (2nd ed.) (pp. 247-258). Mahwah, NJ: Erlbaum.

Naidu, S., Oliver, M. & Koronios, A. (1999). Approaching clinical decision making in nursing practice with interactive multimedia and case-based reasoning. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, 1(2), 1-11. http://imej.wfu.edu/articles/1999/2/03/index.asp

Oswalt, I. (1993). Current applications, trends, and organizations in U.S. military simulation and gaming. S*imulation & Gaming*, 24(2), 153-189. http://dx.doi.org/10.1177/1046878193242003

Perkins, D. N., Jay, E. & Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking. *Merrill-Palmer Quarterly*, 39(1), 1-21.

Perkins, D. N., Tishman, S., Ritchhart, R., Donis, K. & Andrade, A. (2000). Intelligence in the wild: A dispositional view of intellectual traits. *Educational Psychology Review*, 12(3), 269-293. http://dx.doi.org/10.1023/A:1009031605464

Pidd, M. (2004). *Computer simulation in management science* (5th ed.). Chichester, UK: Wiley.

Rescher, N. (1996). *Process metaphysics: An introduction to process philosophy*. Albany, NY: SUNY Press.

Rogers, L. (2011). Developing simulations in multi-user virtual environments to enhance healthcare education. *British Journal of Educational Technology*, 42(4), 608-615. http://dx.doi.org/10.1111/j.1467-8535.2010.01057.x

Ryle, G. (1949). *The concept of mind*. London: Hutchinson.

Schön, D. A. (1987). *Toward a new design for teaching and learning in the professions*. San Francisco: Jossey-Bass.

Schwartz, P. L., Loten, E. G. & Miller, A. P. (1999). Curriculum reform at the University of Otago Medical School. *Academic Medicine*, 74(6), 675-679. http://dx.doi.org/10.1097/00001888-199906000-00014

Shaffer, D. W. (2006). *How computer games help children learn*. New York: Palgrave Macmillan. http://dx.doi.org/10.1057/9780230601994

Siegel, H. (1999). What (good) are thinking dispositions? *Educational Theory*, 49(2), 207-221. http://dx.doi.org/10.1111/j.1741-5446.1999.00207.x

Smee, S. (2003). ABC of learning and teaching in medicine: Skill based assessment. *British Medical Journal*, 326(7391), 703-706. http://dx.doi.org/10.1136/bmj.326.7391.703

Sweeney, G. (1999). The challenge for basic science education in problem-based medical curricula. *Clinical and Investigative Medicine*, 22(1), 15-22.

Taylor, T. L. (2002). Living digitally: Embodiment in virtual worlds. In R. Schroeder (Ed.), *The social life of avatars: Presence and interaction in shared virtual environments* (pp. 40-62). London: Springer-Verlag.

Tishman, S., Jay, E. & Perkins, D. N. (1993). Teaching thinking dispositions: From transmission to enculturation. *Theory into Practice*, 32(3), 147-153. http://dx.doi.org/10.1080/00405849309543590

University of Otago. (2011). Studying in medicine. http://micn.otago.ac.nz/courses-and-subjects/studying-medicine [viewed 27 Sep 2011]

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research*, 79(2), 702-739. http://dx.doi.org/10.3102/0034654308330970

Wiecha, J., Heyden, R., Sternthal, E. & Merialdi, M. (2010). Learning in a virtual world: Experience with using *Second Life* for medical education. *Journal of Medical Internet Research*, 12(1). http://dx.doi.org/10.2196/jmir.1337

Wilkinson, T. J. & Harris, P. (2002). The transition out of medical school - a qualitative study of descriptions of borderline trainee interns. *Medical Education*, 36(5), 466-471. http://dx.doi.org/10.1046/j.1365-2923.2002.01209.x

Williams, B. (2005). Case based learning - a review of the literature: Is there scope for this educational paradigm in prehospital education? *Emergency Medicine Journal*, 22(8), 577-581. http://dx.doi.org/10.1136/emj.2004.022707

Youngblood, P., Harter, P. M., Srivastava, S., Moffett, S., Heinrichs, W. L. & Dev, P. (2008). Design, development, and evaluation of an online virtual emergency department for training trauma teams. *Simulation in Healthcare*, 3(3), 146-153. http://dx.doi.org/10.1097/SIH.0b013e31817bedf7

Appendix 1: House surgeon outcomes for Gertrude

(C = compassion focus; OM = open-mindedness focus)

1. (OM) Diagnosis of Urinary Tract Infection (UTI) is made AND documented.
2. (OM) Diagnosis of Atrial Fibrillation is made AND documented.
3. (OM) Diagnosis of Dehydration is made AND documented.
4. (OM) Onset of NEW confusion is noted.
5. (C) Each diagnosis is discussed with Gertrude.
6. Plan for UTI includes a legal prescription for an appropriate antimicrobial, which may be intravenous (IV) or oral.
7. Plan for Atrial Fibrillation includes:
    a. finding out usual dose of metoprolol (Betaloc) and aspirin (Cartia);
    b. restarting usual metoprolol; and
    c. restarting usual aspirin.
    d. MAY include initial IV rate control.
8. Plan for dehydration is documented (maybe IV or oral).
9. IF any IV therapy is planned, IV access has first been secured and access is documented.
10. (OM) Clear statement of plan for admission or discharge from Emergency Department.
11. (C) All plans are discussed with Gertrude.

Authors: Swee-Kin Loke, Professional Practice Fellow - Educational Design
Higher Education Development Centre, University of Otago
65 Union Place West, Dunedin 9013, New Zealand
Email: swee.kin.loke@otago.ac.nz

Dr Phil Blyth, Senior Lecturer in E-Learning, Faculty of Medicine
University of Otago, PO Box 913, Dunedin 9054, New Zealand
Email: phil.blyth@otago.ac.nz

Judith Swan, Associate Research Fellow, Faculty of Medicine
University of Otago, PO Box 913, Dunedin 9054, New Zealand
Email: judith.swan@otago.ac.nz