Chapter 8 Conclusions and Further Work

The goal of this thesis has been to explore the support of learning through reflection, in the context of medical students and practitioners, working through a series of simulated consultations involving the diagnosis and management of chronic illness.

In order to achieve this, a web-based virtual patient that enables the user to interact with the patient over a number of consultations has been developed. While a simulation with perfect fidelity to a real patient encounter is not possible with current technology, the feedback provided by the users, and the fact that the medical student, general practitioners, and medical student were more focussed on managing the patient than providing feedback on interface elements, suggests that there was considerable success in producing an environment that enables the user to undertake the core activities of a medical consultation. In particular, the examination, investigations, and management selection interfaces enabled users to effectively interact with the virtual patient. The question and answer interface does need to be improved to enable a smoother interaction with the patient. However, until significant improvements in natural language processing and understanding become available, and this is likely in the not too distant future, questioning from category-based question lists will need to remain a part of simulations, such as SIMPRAC. Notwithstanding, it is quite possible that category-based selections will remain the preferred option of some users. The management component of the software also needs to be improved to make a wider range of options available to the learner. However, this represents a limitation in the case that was authored rather than a fundamental problem with the design of SIMPRAC.

With respect to authoring of cases, important qualitative differences were found in the approach taken by the different medical groups. General Practitioners tended to have a much broader consideration of the patient than either the medical students or the specialist medical practitioners. For example, many of the general practitioners believed an authentic case should include additional elements, such as opportunistic screening, and social considerations, in addition to the primary medical disorder. Given these differences, the educational goals of each of these groups needs to be discussed in depth before further cases can be authored. This needs to take place to ensure that all elements considered important for the consultation can be incorporated. Depending on the nature of the case, different cohorts might then be expected to interact with the patient differently. What is important for one group may not be relevant, or may even be inappropriate, for another group. That is, users can only compare all their activity to their own peer group. Inter-cohort comparisons of performance can only be reliably made on common core elements.

The use of multiple consultations, and the dynamic outcome of the virtual patient resulting from user management decisions, was thought to enhance the experience, as well as provide a form of “natural” feedback. Using a consultation-based progression of the patient did not detract from the authenticity of the patient interaction, at least in the context of chronic illness. Certainly, there was no evidence to suggest that SIMPRAC should model, more complex, real-time changes, that might be considered if the focus had been on the diagnosis and management of acute medical emergencies.

Even though modelling the patient on the basis of discrete consultations, with known states and transitions between states, is simpler than modelling real-time changes, it was

not without its difficulties. As a result of the large number of possible outcomes that can be generated from a relatively small number of interacting variables, it is difficult for case authors to anticipate all important outcomes. The development of an authoring tool might reduce some of the burden of producing new cases. However, the necessary complexity that comes from having a series of finite states, dependant on management that evolves over time, means the number of serial consultations must be limited.

Use of a deeper model, based on an understanding of the pathophysiology of the disease process, could lead to more realistic patient responses to interventions. However, a commensurate amount of effort, possibly more than that required to generate appropriate outcomes under SIMPRAC, would be required to generate these disease models. Moreover, the very simplicity of the SIMPRAC disease model means that it can be readily applied to learning about a number of different chronic illnesses. With the current disease model, patient responses and findings are linked to the current state of the patient. This is highly effective. However, ensuring all of these responses are appropriate, when there are a large number of possible states, entails a lot of work by the case author. On the other hand, the evaluations demonstrated that many of the patient states were never encountered by users. Therefore, with appropriate testing, unnecessary effort could be avoided by excluding those states from further refinement.

Taking this one step further, a further alternative, would be to make the introduction to each consultation independent of the management chosen in the previous consultation. In this way, each consultation is effectively treated as a separate case. Feedback on the patient outcome, based on the management decisions, at the end of each consultation, could still be provided but this would not influence the introduction to the subsequent

consultations. In this way, encouragement for reflection, based on information on patient outcome, might still be supported. With such an approach, not only could the user be provided with feedback on their current management decisions, but they could also repeat a consultation and observe the effect of using a different management strategy on the patient. On the other hand, such an approach could detract from the realism of the case, as the patient may no longer follow a logical progression in their illness. Furthermore, the evaluation found that the users believed the multiple consultations and natural feedback, provided by having different pathways, added to the value of SIMPRAC.

Most users covered the important items necessary for the diagnosis and appropriate management of the patient. With reference to the expected performance of the users (see page 46, and Appendix K on page 277), all but one user identified that the patient had raised triglycerides and cholesterol secondary to Type 1B Glycogen Storage Disease. However, three users failed to measure the triglycerides at all. The majority of users excluded other secondary causes of hyperlipidaemia. However, three failed to consider excessive ethanol use as a secondary cause for elevated serum lipids and only the specialists ordered thyroid function tests to exclude thyroid disease as a secondary cause. Cardiovascular disease is not a feature of Glycogen Storage Disease, despite the elevated cholesterol and triglycerides. It is likely that it is for this reason that only the specialists failed to ask about a history of cigarette smoking. Patient management was less consistent. Only nine of seventeen users prescribed uncooked cornstarch to maintain the patient's blood glucose, despite the desirability of this being mentioned in the information sheet. Effective lipid-lowering treatment was commenced by all users. From the think-aloud sessions, it was clear that users were concerned about side-effects

of medication. However, only five students monitored the muscle inflammation and injury by measuring serum creatine kinase. The others, quite reasonably, simply asked about symptoms of myalgia (muscle pain).

Reflection as a mode of learning is probably used by most people to some degree, although they may not be able to name the process. Whether an individual uses reflection depends on their personal goals and the depth of learning that is desired. The development of SIMPRAC involved the definition of a model of the consultative process, which was implemented as an environment for interacting with a virtual patient. A separate model that overlays the consultation model, was also defined to explore the support for multiple points of reflection. Elements designed to support reflection were included as part of the simulation itself, as well as separate components that were used at the end of each consultation. Those elements within the simulation were the hypotheses, the medical record, and the user notes screens.

The hypothesis screen was used by all but one user, and its presence encourages the hypothetico-deductive approach taught to the medical students. Users found that having to enter their hypotheses did interrupt the patient interaction. However, as suggested by one user, future versions of SIMPRAC could limit these interruptions to those consultations where there was a large diagnostic focus. Support for reflection might be enhanced by getting the learner to explicitly associate their hypotheses with their actions. Although it would be difficult to implement this within the patient interaction without detracting from that interaction, it would be worthwhile exploring whether this process could be included during the review phase, at the end of each consultation. In

this way *reflection-on-action* and reflection on reflection (*reflection-for-action*) could be encouraged.

While the medical record and user notes screen were included, and could be used to support reflection and further learning, they were seldom used by any of the users. It would be useful to see if users would be more inclined to use these facilities if they were undertaking a case over multiple sessions, rather than the artificial situation where a single session is undertaken in the presence of a researcher. Further instruction to the users on the purpose of the tools and the value of reflection may also have enhanced the use of these components. These are all issues that can be investigated in the future.

Points of reflection outside of the patient interaction included the action review screen, as well as the graphical comparison screens. Both screens were presented to the learner at the end of each consultation. The action review processes was not seen as overly tedious, and users indicated that it helped them reflect on their actions. The use of just three levels of importance, made the interface quick and easy to use, but did make many users feels as they were being graded as right or wrong when they reviewed how the case author had classified the actions. In the future, it might be worth while exploring whether there is benefit from having more levels of classification, and whether this would encourage a more positive view of being graded, and as being non-judgemental in nature. On the other hand, the medico-legal aspects of medical practice inevitably carry an implication of being negatively judged by society as incompetent or negligent in the event of error. This ultimately contributes to the strong, competitive drive to “get it right” in most practitioners.

While the users stated they thought the graphical review screens helped them reflect on their actions, no one systematically reviewed all the information that was available, to help them in their reflection. In particular, at least three users appeared to fail to consider questions they had failed to ask, which the case author had regarded as critical. This suggests that improvements are required in this part of the interface. While a tendency to be overly thorough will not generally cause major problems, provided the key issues have been identified, omission of critical questions, examinations, investigations, and treatment can lead to adverse patient outcomes. Currently, SIMPRAC displays to the user, those items that the case author considered critical but have been omitted by the user. However, it does not highlight these. Future versions of SIMPRAC should ensure that these issues are highlighted to the learner. In addition to just providing feedback, reflection could be enhanced by asking the user to indicate whether they agree with the author's classification, and why they hold this belief. Another approach that might be taken, where a number of users have overlooked the same issue, would be to add additional states where the patient either presents with excessive ethanol intake, as an additional precipitating event, or where it becomes an issue later in the patient's disease process. In this way, natural feedback can be used to reinforce the message.

Based on feedback from the users, and disagreements about the classification of various actions, there was evidence of a degree of competitiveness among the users. It could be argued that this competitiveness may have been further encouraged by enabling users to compare themselves to their peers. Despite this, the think-aloud sessions suggested that some user's found these peer comparisons more valuable than being able to compare themselves to an expert. It was not clear from the think-aloud sessions, or the

questionnaire comments, whether the users understood the distinction between the case author and an expert clinician. The concern over competitiveness, and the fear of being wrong, is that it may impede willingness to engage with learning for fear of failure or judgement. However, if SIMPRAC is used anonymously over the Internet, without the presence of others, this fear of failure and ridicule may be avoided, and learning enhanced. Under these circumstances it may compare quite favourably to other more authentic and stressful experiences, such as bedside tutorials with an expert clinician.

Overall, the qualitative data from the user evaluations suggests that the software was able to support reflection. However, users did vary in the degree to which they participated in *reflection-on-action* and reflection on reflection, and it was clear that the tool, while supporting reflection, could not ensure reflection was taking place. To make optimum use of an application such as this one, it should be used in an environment where learners are aware of the reflective process, motivated to use the process, and therefore more likely to use the support tools in an appropriate manner. As part of a reflective learning process, this simulation could then be used as a formative assessment tool to enable users to obtain feedback on their current understanding, as evidenced by their actions. Proof of the efficacy of this process could be demonstrated to users by summative performance parameters, which in turn, could encourage users to use this process more. Feedback could also be provided to teachers on common misconceptions and omissions, so that this can then inform future teaching.

Based on the research undertaken, a number of recommendations can be made for supporting reflection in medical simulations.

1) A layer of reflection should be used that is congruent with the consultation process used in medical problem-solving.
2) Components should be available to enable reflection to take place during a patient interaction, but perhaps more importantly, reflection should be encouraged at the end of a defined interaction.
3) Multiple consultations provide an authentic learning activity, and the patient's response to treatment provides natural feedback on which to base reflection. The number of possible states does, however, need to be constrained.
4) Critical omissions need to be highlighted to the user, so that these can be the focus of reflection.
5) Peer and Expert comparison provide a useful feedback as a source for reflection, and may provide additional information back to the case authors to enable further case refinement.
6) Cases should be contextualized for the user population. For example, cases developed for medical students may have a different focus to those developed for general practitioners.

There is a currently a focus on problem-based learning in medical education and this software could be used as a presentation vehicle for this type of medical education paradigm. However, according to Knowles' (1980) androgogical framework, adult learners are self-directed and tend to base their learning around problems that are currently of concern. Therefore, as part of a continuing professional development resource, this software might be used for auditing one's own understanding of a domain in a reflective process. Thus, after reading and obtaining the required abstract technical

knowledge, a case could then be undertaken to put the knowledge into practice. Comparison of activity against an ideal or peer group could then be undertaken to stimulate reflection on an individual's understanding. In this way, a combination of abstract technical knowledge and situation-specific learning can be integrated, as suggested by Anderson et al. (1995b), with the goal of maximizing transfer of knowledge to multiple situations.

Another role for this software might be as a summative evaluation tool. Use within such a stressful environment may even enhance reflection, especially if this is rewarded. For example, if evidence of good reflection compensates for an initial omission. However, this effect is likely to be dependant on the individual. Issues relating to usability and inter-individual experience with this type of technology would need to be addressed, and validation studies would also need to be undertaken, to ensure that performance on the simulation equates to clinical performance.

Further work needs to be done in a number of areas. Firstly, work needs to be undertaken to improve the software usability. The two key areas are the history-taking interface, as well as the management options available to the user. As indicated previously, the number of available questions and responses needs to be expanded, although care needs to be taken to ensure that concepts are not represented more than once. Improvements are required in the ability of the software to recognize a user's questions and alternative software tools need to be considered. However, full natural language understanding is probably not required, as the domain is restricted and most of the users in the evaluation did seem to be able to work within some constraints. The management options need to be expanded. However, as is the case for the history

component, this still needs to be constrained. Input from proposed user or professional groups would be useful in ensuring appropriate options are available, and that these reflect the expected practice in these groups.

Consideration of input from professional groups raises the interesting possibility of collaborative case construction. Using the original case author's work as a base, the case could be "tuned" for particular professional groups. This could even go so far as allowing "errors" to be entered, so that they could be found by others. If such an approach were taken, care would need to be taken to ensure critical elements were not reclassified incorrectly. While critical elements should not be reclassified as "not relevant", it may be fair to reclassify a non-relevant item as relevant to see if users detect these "mistakes" during the reflective process. This purposeful introduction of errors would also reduce the "seriousness" of the interaction, more in keeping with a computer game, where competitive activity does not carry the same negative implications as it does in medical practice. This might well enhance the reflective learning possible with this simulation.

Another aspect of usability that needs to be addressed is the ease with which cases can be authored. If this software is to be used in the future, then case-authoring tools will also need to be developed.

Assuming the usability issues have been addressed, field-testing in a more open, "real-life" environment need to be undertaken, without the user having immediate access to the author and developer. These evaluations need to occur with different groups and at different levels, and should include at least:

1) Medical students as part of their medical course,
2) General practitioners as part of continuing professional development, and
3) Longitudinal evaluations of changes in learning outcome, and changes in reflective behaviour.

Field testing by medical students could be done as individuals or as groups. Although very interesting, the use of SIMPRAC by groups needs to be done in an environment where users have received prior education on the reflective processes, so individual users are able to get the most benefit from the available tools. One might expect small groups to take longer to complete a case than individuals, and to consider more issues. On the other hand, reflection in a group process does need a trusting environment and individual group dynamics might significantly influence the outcome.

Testing by general practitioners could also be undertaken in a couple of different ways. Firstly, it could be used where there have been new developments in the management of particular conditions, or where a practitioner is going to enter a new area of practice. This might be the case if practitioner moves to a new location to serve a different population of patients, with distinct medical needs. The second situation would be to provide training on conditions that are rare, but where early identification and appropriate management can lead to improved patient outcome.

Longitudinal studies are important for assessing impact on physician behaviour, as well as the sustainability of such changes. Care must be taken in the design of the testing instruments, as test performance may or may not correlate with actual practice performance and behaviour. Direct assessment of physician behaviour may not be possible, and alternative indicators of behaviour might need to be developed. For

example, with the introduction of, and education about a new approach to the prevention of some disease, one might expect the incidence of that disease in the population to be reduced if clinician education has been effective.

During the simulator evaluations, there was some variation in the way in which users proceeded through each consultation. Most students took a fairly linear path, by first taking a history, then examining the patient before ordering investigations, and then devising a management plan. Other users, such as ST09 (see page 119), took a less linear path through each consultation. Such observations suggest it might be interesting to undertake longitudinal studies with the simulation software, to see if use of the simulator correlates, or is able to predict, future medical practice and behaviour. This observation also highlights a weakness of this type of simulation, in that this type of simulated problem can only be undertaken one step at a time. In real practice, especially during emergencies, multiple activities are taken at once. For example, while examining the patient, one can also be asking questions about the presenting features of the medical problem. On the other hand, much of this is just an interface issue, which will eventually be solved by advances in technology. At the core of SIMPRAC is its consideration of the role, and need to support, reflection. This will still need to be considered, even as simulations improve.

Another area that needs further investigation is the use of alternative representations to provide user feedback. Currently, feedback is provided as text and a series of simple charts. It was noted that users did not explore all the possible combinations of information available to them, and it should be possible to structure charts to display this information in a more succinct manner. For example, the current charts display a

single stage and a single comparison group. It may be more useful to provide information on multiple groups on a single chart. At the current time, information on the number of critical, or relevant questions that were not asked, is only available in text. Alternative representations need to be explored that enable a user to get an overview of their actions, while also focussing their attention on core issues that may have been done better. There also needs to be some interactive mechanism that allows the system to "know" that the learner has engaged and is aware of the relevant issues.

In closing, learning through reflection is a method that holds much promise for primary medical education and continuing professional development, but until now has been underutilized. To explore the application of reflection to medical education and training, a web-based virtual patient has been developed that enables the user to interact with the patient over a number of consultations. In addition to just simulating the medical consultative process and providing feedback, as has been done with other computer based learning environments, SIMPRAC has explored the use of an additional layer of reflection over the consultative process. The evaluation of SIMPRAC by medical students and medical practitioners has shown that, while the simulation environment could be improved to provide a more realistic experience, reflection by users can be supported and encouraged by providing appropriate tools, as well as by judiciously interrupting the consultative process and providing time for reflection to take place. Further work is required to improve the simulation environment, improve the interfaces for supporting reflection, and further define the benefits of using this approach for medical education and professional development with respect to learning outcomes and behavioural change.