Conversation sim manual

Robert Nelson & Phillip Dawson

Office of the Pro Vice-Chancellor Learning & Teaching
Monash University

Try a conversation sim or contact the authors at:

http://conversationsim.org

Contents
1. Rationale
2. Building the words
3. Configurations
4. Assigning the numbers
5. Analytical advice
6. Discipline- and context-specificity
7. Theory and evidence
8. Exemplars

Synopsis of the chapters

1. Rationale

This manual explains how to create a conversation simulator or conversation sim for short using standard Learning Management System software. The first chapter outlines some of the theoretical aspirations of the project, which are broadly: how to achieve deep learning by efficient digital means.

2. Building the words

In this chapter, we provide a template in a word-processing document that explains what kind of content is needed for each window in the software.

3. Configurations

In this chapter, we reveal how educational content can be deposited within the structure of classical multiple choice assessment. We plot the typical test architecture and reveal the settings needed for the question-answer-feedback loop.

4. Assigning the numbers

In this chapter, we propose the logic by which the values to the three answers might be established.

5. Analytical advice

This chapter explains what kinds of wit or rhetorical voice might work best in each window, given the nature of the content. We outline the characteristics of 'good' questions, provocations, solutions, and feedback.

6. Discipline- and context-specificity

We outline the conventions of dialectical or arguable structures and suggest the limits of applicability of the system to different types of knowledge. We describe where it is most useful and where it is not useful.

7. Theory and evidence

We point to the development of the system in furthering a philosophy of education being about higher-order thought and conversations, not about rules. We emphasize that while a conversation sim creates optimum assessment-as-learning, it should not be used to replace real-time teaching.

8. Exemplars

We provide exemplary (and non-exemplary) samples of conversation sims alongside a commentary on their efficacy.

1. Rationale

This manual explains how to create a conversation simulator or conversation sim for short using standard Learning Management System software. The first chapter outlines some of the theoretical aspirations of the project, which are broadly: how to achieve deep learning by efficient digital means.

A conversation sim is a new educational genre that folds assessment into learning. It is strikingly efficient, effortlessly scales up to handle large numbers of students and is capable of yielding intellectual stimulation in high-order cognitive challenges.

Assessment-as-learning is much more than a convenience and is arguably the only enlightened destiny of assessment design, as suggested by many scholars, *e.g.* Boud, D. and associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Sydney: Australian Learning and Teaching Council. We hope that the conversation sim can contribute to this convergence, helping to deliver assessment from a somewhat traumatic artificial proof-of-learning at the end of a period of study, where it is most likely detached from the process of learning itself and seen in somewhat negative terms.

The concept of the conversation sim did not primarily arise from a desire to assess but a desire to teach in a flexible-delivery mode. Its origins lay in an attempt to replicate digitally the stimulating and joyful exchange that takes place in a tutorial where the content involves analysis and interpretation. In any discursive discipline, the tutorial is likely to take the following form:

- · somebody asks a question
- · somebody else makes a suggestion toward an answer
- · somebody expresses discomfort or agreement or extrapolates
- · somebody argues either way with reason
- · an air of making a decision attends.

- · The question is put again
- · another suggestion is made toward an answer
- · argument follows
- · the thought of making a decision prompts further responses

and so on, for which there is a tantalizing sense of open-endedness, almost as if the extension of the class could form the basis of future philosophical directions.

Tutorials of this nature permeate the humanities and social sciences but may equally arise in science and engineering, as with: 'How does a bicycle stay up when you pedal it?' or 'Why do two identical balls, one filled with air and the other with lead, fall to the earth at the same rate?' to which contrasting responses may be forthcoming.

Infused with a dialectical tenor, good tutorials are highly participative and often exciting for students to attend. They have a vivacious conversational character, where even an incorrect response (according to rigorous academic criteria) presents a beautiful challenge; because it has to be offset both with a critique and with ideas of a more plausible nature. Sensing that a rich intellectual interchange is within reach, students want to chime in and are highly engaged with the subject matter.

Theoretically, the tutorial could be replicated digitally without any form of computer modelling but simply by the live tutor conducting the discussion through a forum or a discussion board, whether in real time or asynchronously. However, it seems much more difficult to create the necessary mood in an online forum than in a live classroom, with its rich body-language and intonation and of course its constant invitation to participate.

This Socratic educational motif has inspired the conversation sim, which in

the first instance was conceived as replicating the structure of a tutorial discussion in a digital format, but automated and requiring no tutor, in the same way that reading a book does not require a tutor. We posited that the typical 'line' of the tutorial discussion could be expressed thus:

problem → proposal → decision → reflexion

which could recur any number of times on the basis of the same problem—given that most problems are unlikely to afford such rapid closure—bearing diverse proposals for an answer.

The basic line could be translated into a digital format in many ways, all of which centre on the third term, the decision. The decision is the key element requiring interactivity, because the decision in question cannot be made by the person who poses the problem or puts forward a proposal for an answer. Rather, the decision has to be fulfilled by the participant; and in an educational case, that means being made by the student. In the sequence of our basic line, the student progresses from being an audience who wonders, a listener who absorbs, a judge who evaluates and then returns to a listener who can ponder a reaction to the decision that he or she has just ventured.

By a happy and fateful convergence, it emerged that one of the best structures on an LMS to serve the argumentative line is the assessment function. All of the LMS packages have multiple choice functions, which seemed, in effect, the perfect game-engine for the automated tutorial. Almost uncannily, multiple-choice questions provide for

- a question
- a suite of possibilities
- a moment of student choice
- a chance to get feedback on the topic.

Exploiting the software for multiple-choice questions at first seemed counterintuitive, because the pedagogical associations with the tired assessment genre are greatly at variance with what we were trying to achieve. Unfortunately, multiple-choice questions are mostly configured around correct and incorrect answers, where a teacher asks a question, provides a correct answer and three or four dummies. For example: which of the following words has a Latin origin?

1. erg
2. meteorology
3. science
4. mathematics
5. analysis

in which the Latin *scientia* lies buried among four words of Greek origin, suitably misleading by their similarly noble antiquity and equally learned associations. By making the dummy questions somewhat credible as a response, I attempt to trick the student. If, for example, the dummies around 'science' had been words like 'star-gazing', 'fog', 'tosh' and 'bilge', the impeccable Latinity of 'science' would be easier to recognize. We make the choice harder by making the wrong answers more believable; the assessment is more rigorous in the measure by which it is more misleading. In educational terms, the genre of multiple-choice questions lies close to deception; and scholarly literature suggests that students in fact learn mistakes through the proliferation of falsehoods, set up as a trap with the risk of giving the student an enduring misconception. An entire body of literature has emerged exploring the ideal combination of distractors and truth (for example, Kublinger, et al. 2010; Lin, et al. 2010), and the potential harm those distractors may have on learning (Marsh, et al. 2007).

Further, the deficiency of the multiple-choice question model is its dependence upon fact rather than opinion. One of the answers has to

come to 100%, and the others are wrong, unless some clemency encourages the teacher to award consolation marks. The idea that multiple-choice questions might broach subject matter of an open-ended or ambiguous nature, where right and wrong are matters of opinion, has never—as far as we know—been contemplated.

The breakthrough that the *conversation sim* represents is to see the multiple-choice questions (since we still have to use that word as part of the LMS architecture) as expressions of positionality rather than truth. In the face of any opinion, a person can react with agreement, disagreement, uncertainty or indifference—a 'yes', a 'no' or a 'maybe'—to which fine degrees of plausibility might be accorded. Ultimately we agree or disagree with one another on a basis which is likely to include some doubt, which is very unlike our first list of scientific words, most of which are Greek and only one of which is Latin.

The classical conversation sim thus inverts the normal suppositions of the multiple-choice questionnaire. Instead of choosing between facts the student chooses between attitudes to arguments. The choice is not a name or an object or a number but Yes, No or Maybe. It could also be expressed as 'Agree', 'Disagree' or 'Partly agree and partly disagree'. The judgement is over an argument, even if expressed as a truth claim, and is ideally not an unequivocal fact which is axiomatically correct.

Once set up, the flow of ideas in the rhythm already noted, namely

problem → proposal → decision → reflexion

is capable of (a) introducing a student to complicated material in a sympathetic conversational spirit and (b) simultaneously measuring the attentiveness or shrewdness of the student's intellectual presence. The lecturer can agree with the student's judgement or find fault with it; though there is no saying that a student gives a correct answer for the correct

reason, which is a problem that survives from the genre of multiple-choice question, where we have no idea if picking '3. science' shows etymological skills or pot luck (or cheating via a Google doc shared within a team of subversive students). This imperfection does not prevent multiple-choice questions from being used in contexts where knowledge, rather than luck, is essential.

Although, as will be revealed in detail, the conversation sim uses a multiple-choice engine, it distinguishes itself from multiple-choice questions (MCQs) in many respects. One of these is the concept of a return to the question, which we have contemplated in the schema for any lively tutorial discussion. I like your response, but let's come back to the question... whence another response is elicited. In an MCQ it is pointless or even impossible to restate the question and to come at the content again. Once we have established that the Latin word is science, we cannot ask the question again, because the answer has been given. The curiosity is exhausted by the question. But in any topic which affords multiple perspectives, the question needs to be put many times before anyone can be satisfied that it has been satisfactorily covered; and even then, we would never imagine that it has been exhausted, given that 2,500 years of philosophy have kept us curious so far.

In the classical conversation sim, we recommend approaching the same issue four times. As already noted, the notional sequence 'problem → proposal → decision → reflexion' could recur any number of times, given that most problems are unlikely to afford such rapid closure. Instead, we comb through differing proposals for an answer and each of these proposals leads to the mini-question of 'yes', 'no' and 'maybe'. The number four seems a happy cluster. Much fewer than four would suggest that the topic is not terribly rich and many more than four responses would suggest that we are not getting anywhere very quickly and fatigue is setting in.

However, there are no absolutes, in the same way that we do not need to stick to 'yes', 'no' and 'maybe'. Naturally, other configurations may work, provided that they express choices that the participant might judge to be necessary.

2. Building the words

In this chapter, we provide a template in a word-processing document that explains what kind of content is needed for each window in the software.

We strongly advise writing conversation sims in any text form other than the LMS itself. Turn it off and use a simple word processor or write in a text cloud. The reason is that much of the success of a conversation sim depends on continuities and balance of riposte, where the responses can be compared, the coverage made complete (or as complete as it needs to be) and where the spread of positive and negative responses can be controlled. For these purposes, it is very helpful to have a single continuous page or pair of pages that has everything visible in one glance. Further, the art of writing a conversation sim is somewhat poetic in nature, for which your favourite and most intimate writing platform is ideal. You don't want to feel that you're feeding words into a machine, but setting out beautiful ideas on a page. We think of it as writing a play.

A good way to structure a conversation sim is provided with the following template

Think of a witty title of the conversation:

A situation
Situation text (or problem text)

A response
response a

　　　feedback a

Another response
response b

feedback b

Another response again
response c

feedback c

A final response
response d

feedback d

Before beginning, it is helpful to think of a problem or a situation with a degree of complexity. One might imagine different people having different attitudes, based on personality, ideology, upbringing, sensitivity, language and values. You are about to write a kind of play, where the content will be dramatized somewhat by its potential for extrapolation, misunderstandings, people talking over one another, the typical chaos that riddles any discussion. The template above gives us the places to write things but of course formulating the content is an imaginative act which requires insight, some invention and lots of practice.

In practice, it could be explained thus

Witty title of the conversation
If possible, one would write a title that sums up the paradox, the oddness, the crazy juxtaposition of ideas about which anyone could be unclear. The title can exploit any of the resources of poetry or journalism, alliterations, cheap puns, ironic pomposity: no joke is too silly and you can even be a teensy bit vulgar. But it is also fine to write a subject in prosaic deadpan style. There are no rules, only liberty! Our aim is education, not entertainment. Also, if a clever title does not come immediately, treat it like

the title of your PhD. Write the thesis first and then a wonderful title will probably suggest itself through the words chosen in the content.

A situation

Under this heading, you want to contemplate an uncertainty, something a little obscure and preferably deeply paradoxical, the stuff of quandaries and conundrums in science or ethics, society or art, literature or politics. If you can make it somehow atmospheric, that is a bonus but not necessary. The intrigue that we want to set up is entirely in the capacity to elicit doubt, a conceptual *aporia* that you know could be tackled in lots of different ways. It is important that this situation should end with a question that calls for a response. So you could even call it 'A question' rather than 'A situation'.

A response

Under the first response, you might immediately put forward something outrageous, where you suspect that the student will answer no. This would be a reassuring first experience. However, you could equally 'seize the tease' and make a more probable suggestion which might be likely to get a maybe; and it is also legitimate to state something wise, which is likely to get a Yes.

> **Feedback x 3**
> In developing the feedback, you do not really have to think of three separate responses corresponding to Yes, No and Maybe, which is why the template above does not have the x 3. Rather construct your feedback first around the answer that you prefer. Once you have achieved a justification of your preferred answer, you can use the identical text for the less preferred answers, prefacing it with words like: 'Really? I would have said…' or 'Why do you say that? I would have opted for…' and then include much the same text as you developed for justifying the preferred answer.

Another response

If your first response proceeded from an imaginary stubborn or recalcitrant character, do not let him or her go. Allow the voice to come back at you and make it as credible as you can. Use your rhetorical powers, as if in a debate where you have to argue the indefensible, to persuade the room that the view—which you have already discredited in the feedback—is extremely holy. Pick holes in your own feedback. Act like one of those awful politicians who wants to keep refugees out of Australia.

> **Feedback b**
> But then you will have to rise to the occasion with equal eloquence in the feedback, once again providing what you consider a refutation. We would try to avoid sarcasm. Let the imaginary character be sarcastic or rude; but you, as the voice of feedback, have to remain very decorous and authoritative after all. You are the voice of the teacher but also, in certain cases, the voice of the university. (So when did that ever stop you from being cheeky?)

Another response again

In truth these responses can follow in any order and can have any content, positive or negative, things that you love and despise or things that you ponder with sympathetic doubt in your own mind. But you might also like to show how difficult questions can be handled without an adversarial spirit. We can do vigorous dialectic but that does not mean that the mood has to be unpleasant.

> **Feedback c**
> Accordingly, if you have projected a case with a grumpy or belligerent voice through some imaginary hostile interlocutor, you can use the feedback to mediate the grizzly rhetoric with humour. Or if the interlocutor is wishy-washy or equivocal, you can use your teacherly charm, as if to embolden other responses that might be forthcoming in the imaginary room.

A final response

Often, the final response is a no-brainer. You want the student to leave the argument feeling a degree of contentment if not a sense of total resolution. Sometimes, the exposition of doubt is unsettling, because there is no absolute authority which students occasionally crave. We cannot always supply these universal certainties but we can probably find somewhat agreeable propositions that allow us to bow out graciously.

> **Feedback d**
> So too, the feedback here might aim at clinching this conciliatory tone, allowing that some puzzlement remains but that scholarship in the area is fertile and our qualms are taken seriously.

When you come to upload your draft text in the LMS, consider restating the question each time there is a new response: thus

A situation

So as in the examples above, you write a problem or describe something that seems not quite right or that somehow demands an answer.

A response

response a, as above

> feedback a, as above

That situation again ß

Repeat the situation that you used to spark response a ß

Another response

response b

feedback b

That situation a further time ß
Repeat the same situation as above ß

Another response again
response c

feedback c

That situation for the last time ß
Repeat the same situation as above ß

A final response
response d

feedback d

The main reason for repeating the situation is technical: it reminds the participant of what the responses are seeking to answer. Rather than wasting their working memory with the content of our question it is free to use as they wish. But repeating the situation before each response has the additional benefit of emphasizing that the last response did not exhaust the subject matter: the question keeps bouncing back; and there is an implicit suggestion in the structure that the 'final response' does not exhaust the topic either. The themes that we talk about are huge and invite ongoing speculation.

Anything that suggests this open-endedness is vital to the mood of a good conversation sim. The risk with any computer-assisted learning system is precisely that it resembles a closed system for the automaton, an uninflected air of monkey-see, monkey-do, where the natural organic disruptedness and inspiration of thought is falsified and made to fit a dead

template that ends with misleading closure.

All poems, plays, novels and learned books and articles have to end somewhere. Technically, they are all the same as a conversation sim; and in structural terms, we need have no anxiety of closure that we do not share with authors in any other genre. But because the computer is based on an algorithm—ironically made more conspicuous by its participative potential—it arouses suspicion as some kind of robot, a mechanical brain-press that locks up thought and requires the user to commit keystrokes which have been prefigured by a designer. It is a kind of panic no more rational than a fear that an author expects you to turn the pages of a book when you get to the final words on the bottom right hand corner, an action which you complete in automatic compliance or sympathy with the very fabric of the page. Nevertheless, if these fears exist, we want to mitigate them. Constantly restating the question one way or another emphasizes that we are not about to drain the topic of potential, no matter what we write. There will always be another angle and hence another interpretation to handle.

3. Configurations

In this chapter, we reveal how educational content can be deposited within the structure of classical multiple choice assessment. We plot the typical test architecture and reveal the settings needed for the question-answer-feedback loop.

Once you have written the conversation sim more or less according to the template above, you are ready to install the content in the LMS. In all cases, this process is quick and easy once you have a little practice. In some LMSs, like Moodle, you have a choice of where to activate the

conversation sim. You can either deposit your content in the Assessment option or the Lesson option, which also has a quiz facility. Both will yield a similar outcome and both are connected to the grade book, so that the progress of the student is monitored.

The following advice is generic across all platforms and does not explain exactly where to go for creating a quiz or multiple-choice questions. Each LMS is slightly different. As noted, Moodle has two locations for building quizzes, both of which work perfectly. In essence, however, once you have found where to set up a quiz, you are ready to install. If you are in any doubt about where to find anything or which buttons do what, the help facility will probably explain it better than we can, given that we cannot foresee what the problem is.

The first satisfying stage is to enter the witty title that you have contrived from your text document. Next, a quiz facility will allow you to set the number of responses or answers; some will express this as options or 'maximum number of answers'. Set it to three (3). Note that this number has nothing to do with the number of times we are going to ask a question, which we are suggesting should be four (4). The number of answers is set to **three** because that is the number that 'yes', 'no' and 'maybe' add up to. At this stage, many LMS pages will also ask if you want to allow student review, which we heartily recommend. It lets the student go back and look at the answers once the test is completed.

The next stage in constructing a quiz is normally to say 'yes' to all of the above, which automatically brings up a page with spaces for (a) a title for the individual question, (b) the question, plus your first response to it, (c) a suite of three windows in which you write 'yes', 'no' and 'maybe' in sequence and (d) feedback directly beneath them to apply to each respectively. Your text file contains all the material that you need for entering these items. What we did not add is the numerical values that attach to each of the choices 'yes', 'no' and 'maybe'. One of them ought to

get 100% but the others could be anything. To help determine these values in the best spirit of a conversation sim, there is a discussion below.

4. Assigning the numbers

In this chapter, we propose the logic by which the values to the three answers might be established.

As an author, you have just constructed your sims around the concept of doubt. Now is the time when your genuine respect for that doubt will be put to the test. To honour the mood of the conversation sim the marking of 'yes', 'no' and 'maybe' ought to reflect the possibility that the answers are debatable and a conversation could continue, possibly contesting whatever you think. Ideally, if the sim is constructed well, the following principles apply.

- identify the preferred answer and give it 100%
- give the 'maybe' option some credence when it is not the preferred answer
- give the least preferred answer (or the antithetical answer) some marks.

The reason for giving the least preferred answer some marks is that if you have allowed a 'maybe' to have some credit, then it follows that some of that 'maybe' justifies a small bit of the antithetical answer. If we give zero to the antithetical answer, we have more or less defaulted to standard multiple choice, where maybe makes no sense either. Thus we could say: an example of an acidic compound is salt. The 100% answer is 'no' because salt is pH neutral. But that fact also means not only that 'yes' should get 0% but also 'maybe' should do no better, because salt is not acidic. As noted above, purely factual content is not suitable for conversation sims. A better approach to such content would be to think of a sim about the detection of acidic properties, where different methods might reveal chemical characteristics more or less efficiently, so that doubt can be entertained.

Because our content is always a teensy bit contestable, the 'maybe'—or

whatever counterpart we chose—has considerable tension around it. If 'maybe' is the preferred answer, it possibly follows that 'yes' and 'no' will achieve an equal score; however, it is debatable, because although 'maybe' may score 100%, because there is doubt, it is also possible that a 'no' is nicer than a 'yes' and *vice versa*. For argument's sake, you might propose that alcohol should be classed as a toxin. Given that it kills brain cells, we may be inclined to answer 'yes', especially contemplating rural health or Korsakov syndrome; and yet 'maybe' seems more prudent and culturally sensitive and might therefore attract the 100%. But if we are teaching viticulture or even theology (just thinking about the blessed sacrament here or the miracle of Christ turning water into wine) a 'yes' may seem downright offensive and a 'no' a strong contender for 100% tag.

As can be judged from this example, nothing sparks tension in a conversation sim like semantics, debates about language and categories. But the example is instructive, because it suggests that the marks are not the place to settle the debate. The debate exists, which justifies your 'maybe' at 100%. It then falls to your assessment of your own attitudes whether to give 'yes' and 'no' both 70% or perhaps less, given that it seems obtuse not to recognize that there is a debate, which the choice of either 'yes' or 'no' seems to imply. In one reading, both 'yes' and 'no' appear to deny that the statement is debatable, but especially 'yes', because this choice constitutes an unequivocal agreement that that alcohol is a toxin. A 'no' at least argues for the existence of a debate, because the student is taking issue with the statement. Alas, however, we cannot rely on logic magically rescuing us from this dilemma. Personally, I would not rate 'yes' lower than 'no' just because it seems to foreclose on the debate a little more. It does not appear to be a principle.

On the basis of the somewhat cheeky controversial example above, scoring is likely to be generous. If the lowest possible mark is 70%, no student will achieve a mark below a distinction irrespective of the energy and attention devoted to the task. Without doubt, therefore, there is a

degree of pressure on the kindness with which we contemplate divergent opinion; and, given that we have established that our application deals with doubt, we have no recourse to make the system meaner by pretending that subjectivity can be eliminated.

Against that, the element of doubt does not always justify a preferred answer of 'maybe'. Far from it, and fortunately so! It is quite conceivable for a 'no' to rate 100%, a 'maybe' to rate 70% and a 'yes' to score 30%. Consider the contention: 'We know that Aboriginal education is failing in Australia because the grades that students achieve in Aboriginal schools are well below national benchmarks'. The feedback would say something to the effect that being below national benchmarks is not a sign of failure, given that other marvellous things might be happening which are not benchmarked, like art and narrative and cultural transmission or the confidence-building of Indigenous people.

5. Analytical advice

This chapter explains what kinds of wit or rhetorical voice might work best in each window, given the nature of the content. We outline the characteristics of 'good' questions, provocations, solutions, and feedback.

Before plunging into options of the voice that you might adopt—with rich dramatic issues of imaginary interlocutors, alter-egos, belligerent devils and cavilsome antagonists—it is worth noting that not all students will approach your conversation sim in the spirit of your best expectations. Conversation sims somewhat polarize student audiences. Most students, in our experience, enjoy the motif greatly; but others, admittedly a small minority, find it annoying. You might imagine that you are having great fun by introducing clever ideas through colourful *personae*, but around 5% of your audience will begin with suspicion and some will only become angrier with every rhetorical flourish that you add.

On the basis of early experience, a set of well-written sims wins immediate approval and curiosity from around 50% of a cohort. This proportion divides into sophisticated students who understand how the multiple-choice paradigm has been cleverly subverted, and naive students who think that a whole realm of artificial intelligence has been hatched for their educational experience. We hate to disappoint them when we reveal how simple it is! A further 45% will appreciate that the electronic engineering is efficient but will feel somewhat indifferent, perhaps crediting the benefits but at the same time missing the human element. But this leaves 5% of students who are likely to remain skeptical, unless various steps are taken.

In extreme cases, students have complained that they feel manipulated, that they are asked for their opinion but there is no space for it, that they want to argue but diversity—though rhetorically elicited—is nowhere accommodated, that they are punished for having different opinions.

Above all, they feel that the good old multiple-choice question format, for all its faults, at least has integrity: there is a right answer among sundry dummies and selecting the wrong answer deserves unequivocally to fail. In a conversation sim, where the logic is built around doubt, students are invited to think that their opinion matters. Discussion is invoked but not honoured. They protest that their opinion structurally only counts insofar as it matches the lecturer's preferred answer. Some students complain that they pay more attention to gleaning the opinion of the lecturer rather than expressing their own opinion. They know that they are rewarded by answering what the lecturer wants them to think, so they dutifully (and resentfully) comply and proffer answers that they think will achieve consonance with the lecturer's persuasion rather than what they really believe. They thus pay more attention to the prejudice and mindset of the lecturer, based on previous feedback, than honestly evaluating the issues in an independent or honest spirit. They are trained to patronize rather than educated to learn.

These criticisms, if they arise, are extremely valuable, because they point to something that can be improved. There is no intrinsic reason for the feeling noted. If, for example, the student feels that they are answering what the lecturer wants them to think rather than what they really believe, it could be argued that the lecturer's *persona* is too predictable: he or she seems to project a party line which encourages the participant to second-guess and flatter the lecturer's leftist feminist ideology or whatever. If only one student in 50 feels that he or she is being indoctrinated, the chances are that something in the method can still be improved.

There is only one problem that is properly structural and therefore cannot easily be rectified. It is the fact that students are not given space to argue and debate; and if they really hold a different opinion to that of the preferred answer, they have no recourse within the sim itself. Certainly they get feedback on their putative bad choice; but at the end of this

miniature sermon, they may not feel that they stand corrected but may want to defend their position. Clearly, this is an admirable reaction: we ought to encourage debate and skills in the art of justification and argument. It is extremely important not to foreclose on this potential and to suggest an air of closure through the very method which is intended to foster discussion open up debatable content.

Given this danger, it is important to set up a forum for each conversation sim, where students can complain or argue that they are right after all. Often, the forums are filled with furthering and constructive ideas on the topic; but for those students who remain unconvinced after reading the feedback, the forum is extremely useful to dispel any sense of voicelessness. The lecturer should, in good faith, attend to the forums and contribute as much as possible to maintain the trust.

As lecturers, we need to be candid about the limitations of the conversation sims: they are not built around absolutes but also do not handle diversity of opinion with flexibility or proper responsiveness. We owe it to all students, not just the testy 5%, to explain that the conversation sim is structurally a bit like reading a good book, where you may or may not agree with all contentions but cannot easily argue with the author when these differences of opinion arise.

After all this guardedness and circumspection of method, we hope that you are not going to enter the authorship of your sims with diminished colour! In any class, a 5% subculture of skeptics and hecklers can cruel the good mood and extroversion that you entered the room with. Foreseeing that you might elicit some hostility from students who may feel manipulated should not lessen your rhetorical chirpiness for writing to the other 95% who will engage merrily with the content and may even spring to the defence of the technique. Comments in forums often reveal spirited enthusiasm, even when expressing the opposite view to that suggested by the preferred answer.

When conducting a lecture or tutorial, you really only have your own voice. From time to time, you can adopt another person's voice; but this jocular strategy requires special theatrical charisma, which you may or may not bring off as well as you had hoped. The same tone that makes one student laugh may also be considered chauvinist, patronizing or offensive by some members of the audience, as when the other voice reveals a class or age or ethnic bias. In general, if coaching tutors, we would discourage such entertaining flair as perhaps superfluous to learning and

6. Discipline- and context-specificity

We outline the conventions of dialectical or arguable structures and suggest the limits of applicability of the system to different types of knowledge. We describe where it is most useful and where it is not useful.

Because the talkative methodology behind conversation sims depends on doubt, skepticism is invited from disciplines where doubt is not intrinsic to the syllabus. Especially in science and technology, there is a widely-held belief that learning is based on logic and fact rather than the kind of doubt that inheres in humanities. Certainly, science aims to eliminate subjectivity; and meanwhile, our model has initially been inspired by Socratic method, the tutorial or discussion group that has long been common in the garrulous and interrogative humanities. In the humanities, there is doubt even about the facts: is the work really a Rembrandt? Is it really a self-portrait? Is it the same painting as that mentioned in Saskia van Uylenburgh's inventory? Though often presented as fact, many factual contentions can be reviewed and debated, and that is all before we arrive at the interpretative core of the disciplines in which meaning is projected and a critique of existing beliefs or theory is providd.

This discursive vivacity in the humanities seems far from molecular weight or hardy formulae for calculating the flow of water in pipes of a given girth or the naming of body-parts in anatomy. As noted already, purely factual material is not ideal; and, if handling purely factual material, our sims may end up reverting to the kind of dummy answers known in traditional multiple-choice questions. Faced with the invariable certainty of the boiling point of any liquid, we know that in a list of different chemical, the one that boils at 100º is water. Answering CO_2 or canola oil would be unequivocally wrong, and there seems little benefit insinuating doubt, unless we think that it is a good idea to trick students. Along the lines that were argued earlier, it might be felt that in any account of fact, doubt would be an artificial element by which simple learning is larded with confusion.

While it is true that the laws of science are universal and there may be little point inducing artificial doubt upon them, it is erroneous to see facts as the enemy. There are many kinds of doubt, and many are perfectly compatible with science and engineering. There may be no doubt about an outcome or even a theory (like the inverse square law) but in the mind of the thinker, there may be plenty of doubt as to which theory applies, how one might arrive at it, what reasons one can give to explain phenomena, how the theory can be proved and what its implications are. Most situations in science can be expressed as a narrative; and wherever a story is to be told, there are forks in the road, as it were, at which we can make imaginative decisions and invite others to contemplate the basis of them.

In the examples given in the next chapter, we have included a case from engineering: how does a bicycle remain vertical when everything about a body rolling on only two wheels would suggest that it should tumble? Another example might be from physics: why do two volumes of equal volume but different mass—say a ball filled with lead and the same ball filled with air—fall to the earth at the same rate. Galileo may have worked it out in a way that has satisfied scientists since his century; but the reasons are not immediately obvious and, if you ask senior engineering students, you will get a suite of different answers, and some may even not believe you when you describe the phenomenon. Just because the physics are absolute, it does not mean that there are no doubts about how best to explain phenomena.

We distinguish between objective doubt and subjective doubt. Objective doubt is where there are intrinsic limits to objectivity: doubt is inherent in questions like how the universe began or what a dog thinks or whether light is best described as energy or particles or whether there is a 'ghost in the machine' in the human psyche. Subjective doubt, on the other hand, is where interpretations depend on our perspective, that is, the position of the subject. Typically, subjective doubt arises in anything vaguely political

or moral or aesthetic: how best to spend money vis-à-vis transport, education, defence, the arts and so on. When values are in play and much is at stake, there is a tussle between emotion and logic, where even the logic seems to be premised upon some emotional attachment and arguments are difficult to disentangle from disparate positions. Sometimes such difficulties are expressed as 'wicked problems'.

Conversation sims clearly rest happily with the ethical quandaries of subjective doubt; but while science, at first glance, seems harder—because so remote from objective doubt—it is riddled with subjective doubt. Any discussion of any scientific observation that calls for an explanation is likely to generate lively responses, which is why science classes are great to teach. Who knows why such and such happens? At this point, hypotheses fly through the air and each one deserves to be tested against what we know or might calculate or estimate. These hypotheses are the natural fodder for conversation sims. When seen in these terms, the technique has huge overlap with the physical sciences.

We do not accept that conversation sims are any less applicable to science, economics or PE or any other discipline. Naturally, there will be differences in the language that we use and problems and explanations concerning circuitry will be largely impenetrable to anyone but an electrical engineering student; but learning the language of a field is essential to study, and so we feel that while a certain amount of sophisticated disciplinary language (or jargon) may initially be intimidating, it is not to be banished from the conversation sims. They are an ideal way to transact the necessary vocabularies.

However, conversation sims are not necessarily useful for all kinds of education. You would not entrust the teaching of surgery to the computer, no matter how well conversation sims might articulate the moral dilemmas of certain procedures. Conversation sims are ideal for any teaching where the mind has to make associative leaps of an argumentative kind, but

learning which involves identification is perhaps not so well served. Certainly, we cannot teach anything that involves muscle memory. You will never learn the oboe through a conversation sim, though an encounter with attitudes to music or an exposition of historical or performative styles would make them ideal. Imagine a conversation sim on the fiercely debated topic of *vibrato*! They are not ideal for learning identification, where close visual or (obviously) tactile senses are involved. They are also not useful for teaching bald facts but rather arguments, logic, plausibility, more or less anything capable of doubt.

Still, there are many paradoxes. The conversation sims have grown up in the context of training, perhaps coincidentally, because not-for-credit training (as opposed to jealously-guarded credit-worthy units) is a teensy bit disowned and open for experiment. Often, training is seen as a drag, taught in either a menacing or perfunctory spirit, with people anxious to tick of a requirement and get onto what they do better. But in most cases, the training has huge educational potential, in which conversation sims have a particularly illuminating role. They draw attention to an area of education which we might call 'training for consciousness'.

The border between education and training is somewhat fuzzy, and in most professions we require both. An example is nursing, where a great deal of medical science and psychology are built into the education; but at the same time, you expect that each nurse will be thoroughly trained in certain procedures, like giving an injection or taking blood. These techniques somewhat define the word training, which means equipping people with know-how and correct practical methods.

In other contexts, however, the word training is used to describe something rather different, shading off into general awareness. For example, OHS undoubtedly concerns a technical knowledge of hazards (poisons, dangerous equipment, heights and ladders, wiring and so on) but also an understanding of risk, which is conceptually more challenging.

Further, a good OHS program will not simply present and assess facts about safety but create mindfulness for the values of safety, the ability to countenance inconvenience, an attitude in favour of lower risk, where very often there is no rule but a sense that it would be safer to (say) shift a ladder rather than reach a little further to the side when already mounted on the ladder, which imperceptibly becomes quite dangerous. We might call this discretionary, psychological and cultural element training for consciousness.

Training for consciousness is important in anti-discrimination programs, staff inductions and preparation, student inductions, OHS, almost anything where a warm and encouraging keynote needs to be struck to establish helpful insight into a productive, fair, constructive and positive culture. Alas, this form of training, which is actually educative in the highest degree, is ill-defined and is served by few helpful methodologies. In many institutions it is taught in either a bossy or excessively casual spirit, achieving almost the opposite of the intended outcome, and with poor, if any, alignment between objectives, teaching and assessment. In addition, because such training programs do not fit within the mainstream taxonomy of education (*e.g.* six credit point units) they are sometimes taught in a half-hearted way, which is a pity and a lost opportunity for positive cultural change.

The value of the conversation sim in this context is considerable. An obvious point in its favour is logistical. Given the problem of recruiting enthusiastic staff to run training programs, the conversation sim offers the great convenience of sustainability. Whatever ingenuity and enthusiasm is put into the content of a conversation sim can be reactivated each year in a sustainable way.

However, we cannot say that training is good for conversation sims. It's more that conversation sims are good for training. The area of training opened up to us with an opportunity, rather like low-hanging fruit, because

it is often somewhat disowned and seen as unrewarding. We found the subject matter extremely worthy and necessary and offering good scope to try a new approach to educational technology. It is important for us to note that establishing our success in the area of training does not mean that our techniques are in any way best confined to training. If anything, we are about redefining training toward education; our scope is unequivocally educational.

7. Theory and evidence

We point to the development of the system in furthering a philosophy of education being about higher-order thought and conversations, not about rules. We emphasize that while a conversation sim creates optimum assessment-as-learning, it should not be used to replace real-time teaching.

From the outset, we would be keen to temper the uptake of conversation sims with three caveats. First, the concept is in its infancy and much reflective and analytical research has yet to be done. The developers have three refereed publications on the theory and related cultural traditions of education, identifying what we believe to be the originality of the concept but also elaborating on its links with age-old techniques of reading and conversation itself. But these papers are far from exhaustive, because they do not examine empirical data sets. At this stage, as noted, the spread of experience is local and the excellent results so far may be attributable to local factors. More thorough investigation is clearly indicated.

Second, conversation sims are for augmenting other modes of teaching, not replacing them. Although some syllabus is ideally suited to conversation sims (and can be handled through this convenient technique with much satisfaction to students as well as university) the contexts in which they have arisen are somewhat specific, all centring on an idea of preparation for other tasks. In the normal educational context, a lecturer who works hard at establishing contact with students on a personal level is essential. We do not enjoy the idea that a conversation sim is some kind of MOOC child. The idea was developed independently and had its gestation before the MOOC began its insidious rise in international celebrity. We do not enjoy the concept of automated education for its own sake, either as labour-saving or cost-saving. Rather, we love the idea that students can have a rich set of experiences online that augment the forms of teaching that do not scale up so effortlessly. We love the way that

frequent low-stakes assessment can alleviate anxiety and that assessment can occur simultaneously with learning, without the learning being cruelled by competitive assessment rituals. Rather, the spirit of a conversation prevails.

Third, if the uptake of conversation sims is rapid—not to get ahead of ourselves—it is possible that the method could be equally rapidly devalued. Success at the moment might be explained by novelty. Students have never experienced a conversation sim; and so when they encounter it, the moment is seductive if only because of the surprise. Educational interfaces are nothing if not predictable; so here is a cool diversion that possibly wins favour because of the unexpected. Give the method a couple of years, by which time it may be embedded in various units, and this novelty will have worn off. Instead, fatigue may set in, similar to the boredom that we experience with lock-off recordings of live lectures, pdfs or MCQs on the LMS.

Taking this skeptical glance at the project is helpful in two ways. First, it discourages us from making too much of empirical data sets that justify our swagger. We know that the great majority of our students at Monash like the conversation sims. We do not know exactly what they like about them; nor do we know that they always will like them if the method proliferates. And second, the modesty that this unpredictability encourages also suggests a focus on the ingredients of a conversation sim that are not subject to novelty, especially the joy of conversation and reading.

We do not tire of reading; and if there are more books in the world, the exercise of reading does not pale in comparison to when there were fewer books. Books are boring when they are badly written or when we cannot bestir ourselves to take an interest in the subject matter. Reading itself is sustainable and is hardly diminished in prestige by its ubiquity. So, we imagine, with conversation sims! If they are any good, they will stand up

regardless of their overuse or pervasiveness. The measure of being good is how well the concepts are strategized as a dramatic encounter and thereafter how well they are written.

To date, the student response to the conversation sims has been excellent. The conversation sim has been trialled successfully at Monash University with the most discerning and critical student demographic. It was developed in a unit on research student supervision in the Graduate Certificate of Higher Education among ambitious new academics. We can vouch for the somewhat unforgiving attitude of our students (who mostly have PhDs and are building their publications) by the fact that previous incarnations of the unit were rated very low in the unit evaluations from the same semester in 2010, a frightening 1.38 both for 'Intellectually stimulating' and 'overall satisfaction' in the typical scale from 1–5.

We set about transforming the unit last year (2011) using the conversation sim and immediately achieved a result of 4.43 for 'overall satisfaction'. Further, on the question of whether or not the unit was 'intellectually stimulating', students rated the unit in the very highest bracket at 4.75. Solicited and unsolicited feedback has identified the conversation sim as the determinant of this improvement. The unit's Chief Examiner received a flow of emails throughout the semester complimenting the design for its ease of use but also its engagement with difficult and, at times, insoluble problems of research graduate supervision. On the strength of (a) the student approbation and (b) the close correlation between grades in independent assignments and grades according to the conversation sims, a decision was taken to run the unit entirely by conversation sims in the first semester of 2012, in a grand suite of 72 sims. The experiment yielded similar results to those of the previous semester (all in the 4s) and the following semester's unit evaluation results will be entered when they appear.

Among the many discursive features of the conversation sim, the students

especially valued the integration of assessment within the learning experience. It seemed wonderful that an extraneous process of examination could be entirely avoided and that assessment tasks would be acquitted within the process of learning or attending digital class. Student grades, along with their qualitative feedback are testament to a lengthy and considered engagement with content, contributing to higher-level learning outcomes.

Encouraged by these results, we have also constructed a leadership program (MOHO, Monash Orientation for Higher Objectives) entirely in the LMS using the conversation sim. The subject matter deals with sustainability, wellbeing and social cohesion, all areas that must accommodate doubt and opinion. The number of students taking the program could be in the hundreds or even thousands, as it is open to all students and is non-selective, making it quite the opposite of our research graduate supervision unit. 548 students joined MOHO, though many were clearly tentative, as they undertook none of the tasks but were merely checking it out. Students were perhaps led to believe that this window-shopping is legitimate, because we wanted to make it clear that there were no repercussions for withdrawing or non-completing.

167 students completed MOHO with excellent grades according to Moodle (though as a not-for-credit 'unit', the program ultimately only recognizes PGO) and a similar number filled out an evaluation instrument which was constructed to mirror the SETU. Again, the results are all in the 4s.

Again, encouraged by the success, we built a suite of modules for the Teaching Associate training program at Monash (TACT, Teaching Associate Conversational Thinking) using the same approach. The results were again very good and sufficient to give the university confidence to entrust the teaching of Research Integrity to the new paradigm. Research Integrity is a mandatory part of the new Monash Doctorate and Masters by

research, which have coursework folded into them as well as training. The stakes are high and the clientele discerning. Official evaluation results of the Research Integrity program will not be generated till 2013; however we will be soliciting participation from existing candidates in sufficient number to be able to verify if it is as good as we think it is.

The brief in all of these instances could be considered somewhat mark-insensitive. Many of them are assessed to Pass Grade Only (PGO) and the primary aim is for the institution to provide guidance in best practice, to arouse interest and curiosity in the field but not to credentialize students on a competitive basis. For this purpose, conversation sims have many advantages and apparently no drawbacks. In a highly competitive environment, where career prospects, higher places or scholarships hang upon the outcome, fine degrees of grading are considered important; and conversation sims are arguably not suitable. They are not terribly good to secure an even distribution of marks according to 'the curve'. They are rather ideal for establishing a sound attendance at certain thoughts, providing evidence of attention but not necessarily outstanding performance or diligence in extra effort and reading. For a more in-depth theoretical discussion of the conversation sim approach and its antecedents, see:

Nelson, R., & Dawson, P. (2014). A contribution to the history of assessment: how a conversation simulator redeems Socratic method. Assessment & Evaluation in Higher Education, 39(2), 195-204. doi:10.1080/02602938.2013.798394

8. Exemplars

we provide exemplary (and non-exemplary) samples of conversation sims alongside a commentary on their efficacy. To try a sim live, go to http://conversationsim.org

Let us start, reflexively enough, with a meta-sim, a sim about sims, a sim sim.

A sim sim or Stick this in your VLE and smoke it.

A situation

Did you hear about the two guys at Monash who claim to have developed a new way of teaching with the computer? They reckon that if you frame your syllabus around conundrums, paradoxically you can make it come to life on the computer. Do these guys really believe that they can animate a robot?

A response

The idea is ridiculous. Under the catchy mendacious title of a 'conversation sim', they're pretending that what is essentially a multiple choice quiz—shamelessly exploiting the MCQ facility on any LMS—is something akin to pedagogical wisdom. It's actually nothing but a glorified test.

> Not really. It's true that we use the test structure but it's more as a kind of peg upon which to hang a coat of narrative. While it has the advantage that you can assess the student's participation, this convenience in no way lessens the strength of the storyline nor the sense of participation in the student experience.

Another response

But for goodness' sake, tricking out multiple-choice questions—which are

the least conversational of all technologies and a completely insular Web 1.0 structure—with the air of a conversation is pure pedagogical blarney. It encourages teaching to be duplicitous chicanery, proposing to students that a conversation is to be had but then never listening to their voice.

> Not really. There is nothing fraudulent or underhand about a representation of other people's views when they are not necessarily your own. Plato did it; and even when it's a matter of setting up skittles to knock them down (or, mixing metaphors, creating straw men) I don't think that it speaks for some kind of arrogance or encourages us to see teaching as deception.

Another response again

If you don't listen to the student's voice in this 'conversation', you only hear your own prejudices trotted out and you've constructed the whole dialogue around your belief. Okay, you've tried to imagine what another person might think but no way does this entitle you to assume the authority of an independent voice, which you will never deign to hear. You say that you're teaching curiosity but you're practicing willful deafness.

> Maybe! I agree that we have to guard against our own conceit. There is nothing in a conversation sim to prevent us from sliding into self-satisfaction with our own rhetorical powers; but bear in mind that these dangers were always present with lectures, which could equally be set up to ridicule or satirize methodologies and scholars that the lecturer dislikes. Finally, the invocation of plurality is not entirely in vain, because a conversation sim would normally exist in an LMS with abundant opportunity for critical student chat.

A final response

Perhaps instead the conceit goes along the lines of a play, where the lines have been constructed around personalities whom you represent or conjecture or extrapolate into existence, as if you're a dramaturg. It is true

that we don't exactly listen to anyone but then we don't claim to. There isn't really a pretence but more a representation which is in sympathy with imagined voices.

> Yes, I maybe agree with that. We don't claim to listen to anyone, though it could be argued that we suggest a listening process, rightly or wrongly, when maybe none has transpired. When we project other voices which we might also discredit, I hope that it's with good will and sympathy, though I cannot prove that this will always be the case and that some abuse will never occur.

How does a bike stay up? A problem in engineering

A situation
A bicycle cannot stand up without support, at least while it is still. Like anything that rests on two points, it is extremely unstable and will tumble unless it leans on a third point. So what accounts for the relative stability of the bike when it is in motion. What is the difference between a bike in motion and a bike at rest that accounts for its dynamic stability?

A response
Well, it's a matter of balance, isn't it? It's no different to someone walking on a thin plank or a tight rope. You more or less have to hold both arms out, so that if you feel yourself falling to left, you rapidly pull up your right arm and let your left arm go down, thus imparting a thrust—by reaction—to the right and hence rectifying your less than upright disposition. In other words, you correct your imbalance using your body and so prevent a fall.

> Not really! Admittedly, it's a good description of balancing on a bar or whatever; but we don't use those actions on a bike. Our hands are normally planted on the handlebars and we cannot use the inertia of

our limbs to jerk this way and that to maintain equilibrium. It would make for a very risky and dysfunctional style of riding.

Another response

Sure. The explanation is technical. The reason that the bike stays up is the gyroscopic inertia of the wheels. When a wheel spins, it has a form of circular momentum that keeps the axle in a single plane and hard to dislodge from its orientation. That's why a bike going quickly is more stable than a bike going more slowly. The gyroscopic inertia is proportional to the speed of rotation, and this explains the instability of a slow bicycle.

> No. Gyroscopic inertia may play a very minor part but it cannot explain how a slow-moving bicycle still manages to stay up. Further, as the gyroscopic inertia must be proportional to the mass of the wheel, it would mean that heavy old bikes are more stable than nimble light ones. But that suggestion is not consistent with experience. A heavy bike can tumble just as easily as a light one.

Another response again

Heavens, you're right. And I can tell if I run a bike along the ground with a mighty thrust and let it roll, the unsupported bike will still capsize, even though it has a certain amount of speed. So it must be something else. I think that it's got something to do with the rider. It's a quality of the inner ear that keeps us balanced. It's a biological system, not just an engineering system.

> Maybe. We certainly need our balance, but that's for everything, like getting out of bed. It's right in one sense but still doesn't explain how the bicycle stays up. Like what mechanical action takes place to maintain the balance? But hey, isn't there a clue with what you just suggested with the motif of pushing the bike by hand very forcefully and then letting go? What is the bike not doing with all that speed,

such that it falls over?

A final response

Okay, I get it. It's the steering. The bike hurtling down a hill without a rider has no steering and that's why it very quickly falls over. As riders, when we begin to fall to the left, we automatically or instinctively steer to the left. That means that our inertia moves the weight to the right of the line between the two points of contact with the ground. All of those small shifts reposition the wheels to remain beneath our centre o gravity.

> Yes. In riding a bike, our centre of gravity always has to be on the same vertical plane as the footprint of the two wheels. When we steer, our momentum can be thrown either to the left or the right of that plane, so we can giggle our base, a bit like holding up a stick in the palm of your hand: you're constantly repositioning the base so that it remains more or less underneath your centre of gravity, with constant corrections.

9. References

Boud, D., & Associates. (2010). *Seven propositions for assessment reform in higher education*. Sydney: Australian Learning and Teaching Council.

Kubinger, K. D., Holocher-Ertl, S., Reif, M., Hohensinn, C., & Frebort, M. (2010). On Minimizing Guessing Effects on Multiple-Choice Items: Superiority of a two solutions and three distractors item format to a one solution and five distractors item format. *International Journal of Selection and Assessment, 18*(1), 111-115. doi:10.1111/j.1468-2389.2010.00493.x

Lin, J., Chu, K.-l., & Meng, Y. (2010). *Distractor Rationale Taxonomy: Diagnostic Assessment of Reading with Ordered Multiple-Choice Items.* Paper presented at the American Educational Research Association, Denver, Colorado.

Marsh, E., Roediger, H., Bjork, R., & Bjork, E. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review, 14*(2), 194-199. doi:10.3758/bf03194051

Nelson, R., & Dawson, P. (2013). A contribution to the history of assessment: how a conversation simulator redeems Socratic method. *Assessment & Evaluation in Higher Education, 39*(2), 195-204. doi:10.1080/02602938.2013.798394