

International Studies Association Annual Conference 26-29 March 2014, Toronto, Canada	Innovative Panel FA05: Advancing the Learning Environment in the Digital Age <i>Friday, 28 Mar, 0815-1000, Dominion South, Sheraton Centre Toronto</i>
--	--

“Assessment Strategies in Simulation Games”

Simon Usherwood

School of Politics
University of Surrey
Guildford GU2 7XH
UK

s.usherwood@surrey.ac.uk

ABSTRACT

One of the big challenges in bringing simulations into the classroom is the question of how (or even whether) to assess them. In this presentation, I will consider the underlying logics of simulations, which in turn suggest a number of assessment strategies. These include assessing knowledge acquisition, skills development and critical reflection. Beyond that immediate challenge, the presentation will also throw some light on related questions of feedback and simulation design.

Keywords: simulation games, assessment, alignment, learning objectives

Draft paper – do not cite without author’s permission

Assessment Strategies in Simulation Games¹

The use of simulation games in Higher Education has increased markedly in recent years, as part of a wider process of engaging with active learning techniques. This has been driven by advances in pedagogical research, student demand and increasing institutional imperatives to deliver innovative learning. However, at the same time, it has been evident that many actual and potential users of simulations encounter a number of barriers to their effective use (see Usherwood 2014 for an overview): this suggests that the full potential of simulation games is not being realised. In this paper, we will consider one aspect of this challenge, namely the relationship of assessment to simulation games.

Assessment is an important aspect of simulation use, since it invites fundamental questions about the general function of those simulations and because the nature of the pedagogical environment potentially renders traditional assessment modes invalid. At the very least, it is a subject that requires careful thought by any simulation user, since the range of options involved is substantial (see Chin *et al* 2009 and Bellotti *et al* 2013 for overviews). Even if the conclusion is that assessment is not needed, assessment is closely bound up with feedback, which is essential in simulations, so it is important to unpack that relationship as well.

The paper starts by considering the functions of simulations, before turning to how we can best gain from their use. This provides a framework within which it is possible to explore the variety of assessment strategies available – considering both their merits and difficulties – and the relationship between assessment and feedback.

¹ This paper builds on earlier work, including Raymond and Usherwood (2013) and Usherwood (2014). As such it remains a work in progress.

THE PURPOSE OF SIMULATIONS

The proliferation of simulations reflects their adaptability of purpose: as a pedagogy, it possesses great flexibility in a number of basis dimensions, from size and level to complexity and duration. However, we can identify three main purposes that are usually understood to be in operation when running a simulation.

Firstly, simulations allow us to consider decision outcomes. Here, the focus is primarily on the substantive issues under discussion and the materials that are drawn into (and pushed out from) that discussion. The simulation here is seen as a way to allow participants to integrate a wide range of source materials into a more coherent whole and then to reflect on the dimensions and interactions which that whole contains. Thus we might run a simulation on a foreign policy decision, to allow participants to see how hard and soft elements of policy can work together in strengthening a particular policy position. Zeff (2003) suggests that using a simulation of the European Council offers improved understanding by students, as compared to more traditional teaching means, while Galatas (2006) points to a strong positive impact of his Council exercise.

Secondly, simulations allow us to consider the dynamics of negotiation and institutional dynamics more generally (Lantis 1998). This is a very common feature of political science and IR simulations, where the specific institutional practices contained within rules of procedure and the varied structural power of different actors can be more meaningfully communicated to participants than in a passive learning environment: instructors might get participants to compare the institutional logics of the UN Security Council and its General

Assembly, for example. Here, the focus rests on skills development, as participants come to appreciate the role of research and preparation, presentation, rhetoric and consensus-building as fundamental parts of the institutional life that they are recreating.

Finally, and more rarely, simulations allow for the development of a group identity. This is rarely the primary motivation, but simulations provide good opportunities to function as ice-breakers or as introductions to problem-solving techniques. Here the boundary to games in the broader sense becomes less clear, but certainly within HE and other educational environments, we can observe the development of group affiliation through the shared experience (Schick 2008).

These purposes of substance, process and group-building are often not explicitly articulated to this degree by simulation designers, who typically are looking to a more nebulous objective of improving participant understanding. However, as we will discuss below, having a clear and focused purpose is a common issue for simulations and designers and instructors would do well to reflect on how these potential purposes relate to their simulation. This is particularly true when considering the overlap and potential reinforcement between them.

MAXIMISING UTILITY IN SIMULATIONS

If the pitfalls of designing and running a simulation appear numerous, then it is also important to observe that typically no one of them proves fatal to the success of a particular simulation, in large part because of the multiple objectives that they can serve, as discussed

above. In addition, much of a simulation is contingent upon factors outside the designer's or user's hands, most obviously the individual participants playing any given iteration.

Since simulations can be organised in so many different ways, it is counter-productive to suggest a single model for so doing. However, it is still useful and pertinent to observe that whatever approach one takes to building and running a simulation, three core requirements have to be satisfied as a precondition to a successful outcome.

The first core requirement is that the learning objectives must be clear to all participants.

This is as true for the simulation designer and leader as it is for students participating in the simulation (Gredler 1992). As has been previously discussed, without clarity of purpose simulations become little more than diversions. The designer needs to have a well-defined set of learning objectives in order to create a simulation that speaks to them, as well as an awareness of what else it might be conveying to participants: the multi-faceted nature of simulations means that it is almost impossible (and probably counter-productive) to create a mono-dimensional scenario. Likewise, for participants, there is the need to set out in explicit terms the objective of the task, the nature and degree of support and/or preparation that is available and allowed, as well as the nature of any assessment attached to the exercise.

The learning objectives can most obviously be related back to the three categories of purpose identified above: decision outcomes/substantive knowledge, negotiation dynamics/skills development or group socialisation. Their over-lapping nature makes clarity of purpose all the more important. The articulation of these objectives helps to guide everyone involved, especially in larger simulations, which by their nature tend to be less defined, particularly

when participants might be asked to create or modify procedures or practices from real-world examples, or where the outputs might permit a degree of flexibility in their construction.

This leads into the second key requirement that the learning objectives have to be aligned with the game play and with any assessment. Alignment has been a key theme in education research since Bigg's work in the 1990s (e.g. 1996, 2003), not least because of its general application to all areas of teaching practice. This should follow logically from the designer describing the objectives in clear terms, since it becomes much simpler to see whether the game play allows the participants to focus on the objectives and whether the assessment tests the achievement of them. From the perspective of the participant, clear alignment of the elements reduces the potential for dislocation, improves immersion into the simulated environment and ultimately creates the opportunity for a much fuller learning experience.

Thus, a simulation that wants to build understanding of the internal institutional dynamics of the United Nations Security Council would be advised to re-create the elements around the Council itself (e.g. bilaterals, contact groups, etc.) and make use of the full rules of procedure, while a simulation that was more interested in the tensions between member states in producing policy might run on simplified rules, but allow for iterated decision-making in a crisis scenario. Likewise, European Parliament games might treat political groups as undifferentiated or with internal tensions, depending upon the purpose the designer has in mind. At the level of assessment, if the focus is on negotiation dynamics, then it is possible to assess on the basis of a reflective piece by each participant that stresses such elements in their experience, while a simulation that wants to develop abilities in substantive policy might tie the game play to real-world decisions in the same field to explore similarities and contrasts.

The final key element that has to be put in place is a meaningful system of feedback to participants. Regardless of the quality of the simulation, or of its design, without feedback the exercise cannot be properly brought back into the rest of the students' learning experience. Of all the three points set out here, this is the one that is most overlooked and the most consequential (Newmann & Twigg (2000) provide a rare example of how this can be done). This point will be returned to in the final section

ASSESSMENT STRATEGIES

The preceding section has outlined some overarching concerns, but these must also be connected to the more specific issue of assessment. In essence, this requires answers to the two questions: do we need to assess, and how do we assess? We will consider each in turn.

The first question in any consideration of assessment strategies is that of why we might assess at all (see Raymond & Usherwood 2013 for a broader discussion). In essence, the answers in favour of assessing boil down to one of three options. At a pedagogical level, assessment might be desirable if it allows students to access a particular form of learning. This is most evident when thinking about developing student reflection, and the production of a reflective report that gains feedback from a marker is a efficient and effective way of achieving this. At the practical level, assessment serves as a system of valorisation, focusing students' attention onto a particular aspect of an activity. Thus if we tell students they will have a paper after a simulation, assessing their knowledge of the procedural rules involved, then we would expect students to pay more attention to those rules within the simulation itself. And finally, at an institutional level, we might simply be required to assess. This is

rare, given the principle of academic discretion, but in some systems, internal and external quality assurance systems would expect any substantial activity within a degree programme to be evaluated and assessed. In a softer form, the alignment of learning objectives and game play mentioned in the previous section might logically lead to a requirement to assess.

In contrast, assessment might be avoided if it offers marginal benefits to its associated costs, or if the simulation element is only a relatively small one within a course/module. Where such boundaries lies is a matter beyond this paper (see Knight 2002 for an interesting discussion), but it is something that needs to be given a suitable amount of thought in either direction, since the consequences can potentially be quite significant.

If a decision to assess is made, then it is then necessary to consider what that assessment should look like. Considered in broad terms, the key dimension is that of proximity to the simulation *qua* simulation. The further one moves from that, the more the options that present themselves falls within conventional assessment approaches, which are more recognisable to new users, but with the cost that they do not access all the pedagogic value that simulations have to offer.

Furthest from the simulation itself, assessment can focus on students' wider learning from the course/module. The assumption here would be that any simulation was only one element of the teaching package and that assessment was structured to make connections across elements within that package. Thus, a course/module might run for a semester, with one week devoted to a simulation that allows students a different perspective on the given topic: a UN Security Council (UNSC) simulation to let students see how the theoretical discussion about the dynamics of that institution work in practice, for example.

The form of this assessment would look like a conventional piece of coursework or a final exam (“what are the key factors in the operation of the UNSC?” in this example). By integrating the simulation with the rest of the course/module, such assessment promotes more holistic reflection, coupled to a more rounded set of experiences on the part of the student. However, this does come at a price. Because the assessment does not link directly to the simulation, it does not valorise it for students, so they might choose not to engage so fully with it: in the example given, it is possible to answer the question whether or not you attend the simulation. This disconnect from the simulation (and particularly from any of the personal skills development aspect within it) means that this is a low level of alignment to the simulation game play and potentially to the learning objectives. More particularly, it raises the question of whether a simulation is really needed at all.

A second strategy is to focus assessment on the simulation topic itself. Necessarily, this requires that there is enough within the simulation to be meaningfully assessed. That might imply an extended simulation, either in time or in relative importance within the course/module. To use our example, the UNSC simulation might be run over several weeks and act as a means for students to discover dynamics and join it to wider reading. As in the previous strategy, either coursework or a final exam could be used to ask the same kinds of questions, the difference being that the simulation is the primary delivery mode for substantive knowledge.

Because the simulation becomes the key vehicle for learning, the assessment more clearly links to the activity and so valorises the simulation in the learning process. At the same time, it is exactly that link that poses the key challenge – which is also true of the other simulation-

focused strategies – namely does the simulation offer sufficient depth and scope to allow the students to answer the assessment questions. This matters because simulations are intrinsically uncertain in their operation: we should expect there to be variation between iterations (see Usherwood 2013). In this case, because the questions relate to the substantive knowledge aspect, much care must be given to designed a simulation that allows and encourages students playing it to find, use and reflect upon that knowledge. Thus, if the UNSC game focuses on states' positions on a given dossier, that might not help with answering a question about negotiation dynamics. In practice, this type of strategy requires a close dialogue between game design and assessment design, to ensure that the two align properly.

The third strategy moves much closer to the simulation itself. Here, students are evaluated by an external assessor on their performance within the simulation. Again, this requires a simulation of sufficient scope to allow all students to have a reasonable opportunity to perform: as such, it is most commonly seen in simulations that run over a full day (think here of non-Higher Education events such as Model United Nations that use judging). The appeal here is clear: students know that they are being watched and evaluated and so have clear and direct incentive to perform to the best of their abilities. Moreover, by keeping the assessment synchronised with the activity, there is scope for very rapid turnaround of assessment.

Despite such attractions, teacher evaluation is highly problematic. While all assessment has a degree of subjectivity, it is much more marked in this instance. This starts with the difficulty of establishing clear criteria: what is to be considered? How do we measure it? How do we weight different elements? Consider two students, one of whom works assiduously throughout the simulation, making repeated and constructive interventions, the other of

whom does nothing until the very end when she uses a simple procedural point to secure her objectives: who is the better student?

This problem extends into gathering evidence to support the assessment decision. In practical terms, it is impossible for an individual to observe more than five or six people for any length of time. This in turn implies that either other assessors need to be present (which will heighten the difficulty of evenly applying the assessment criteria) or some form of recording of the simulation (audio or visual) is needed. The difficulty with the latter option is that one risks missing the pertinent aspects of the simulation, such as the conversation in the hallway, or the online traffic between participants. In any large-scale simulation, such a proliferation of communication and negotiation points is a given and must be borne in mind.

To some extent this is a more philosophical question than anything else. Can we assure ourselves that we have sufficient evidence to make an informed decision? To some extent, one could sidestep the issue by assessing on the basis of ‘success’ in the simulation: did the student achieve their aims? The danger there is that it might not be possible all everyone to win and – more importantly – it might be prototypical for there to be winners: the author recalls a European Parliament simulation with such a mechanism, which encouraged students, but which didn’t give them a very useful insight into how that institution works as a consensual body.

Logically, discussion of teacher evaluation leads to the final assessment strategy: student evaluation. This form of assessment is closest to the simulation itself, since it is generated by a participant and set within a framework of that participant’s own understanding. Crucially, and possibly problematically, it requires that students are able to reflect on their own learning

processes and are able to integrate substantive knowledge with performative skills: while this should be a given with Higher Education students, it becomes more problematic when using simulations with those not yet at that stage.

Student evaluation also differs from the other strategies in that its focus is not so much on the substantive knowledge, but rather on the skills of critical reflection and integration of understanding. In practice, this simply enlarges the difficulty noted above, namely that the scope of possible answers to a question on the lines of “what have you learnt from this simulation?” is necessarily very much larger than it is for any of the substance-based question outlined above. Even if it is framed more narrowly (“show how your experiences in the simulation illustrate the difficulties of finding agreement within the UNSC”), there is still the possibility – indeed, likelihood – that individual students will produce very different accounts.

This intrinsic flexibility of answer must therefore be accommodated within both the framing of the assessment questions and in the range of what is considered acceptable as a response. This can be done more easily in some contexts than in others. The author runs a module on negotiating in politics, where the assessment is a reflective review of the students’ experience of what they have learnt through a series of negotiations, which they are then asked to link back to the academic literature. Because the module is focused on skills development, informed by acquisition of substantive knowledge, rather than the other way around, this assessment strategy works well in reinforcing the central objective of promoting self-criticality.

It is this last point that is perhaps the most important one. No one of these assessment strategies outlined above is the ‘correct’ one: each is potentially valid, but only within the terms of the learning objectives. Ultimately, how (or whether) one assesses must be a function of what one aims to achieve: without an understanding of the latter, the former cannot be properly determined. Seen in a more practical light, that requires a repeated interrogation of objectives, game play and assessment throughout the design and development process to ensure that they continue to match up and reinforce one another.

FEEDBACK AND ASSESSMENT

One particular aspect of assessment must be considered, namely the relationship with feedback. Whatever assessment strategy is used, it must be integrated with feedback: this latter is an essential element of any simulation, as has already been noted. As such, feedback should always be present, and so assessment (if it is also present) must work with the chosen feedback mechanisms.

The importance of feedback has already been discussed, but primarily relates to the need to connect the simulated experience with that of the real-world: we use simulations not to tell us about the simulated world, but its actual counterpart, so we have to draw that back in. With this in mind, we might usefully consider the practice of feedback, with its concomitant points of connection to assessment.

The substance of feedback can be focused on processes, actors or outputs within the simulation: again, the emphasis would logically follow from the learning objectives, but

where these are multiple in nature then it would be advisable to look at all three, since they form the basic units of any simulated interaction.

The timing of feedback is also flexible. Usually, it would come immediately after the main game play, since this is the point where participants are best able to recall detail and connect it to their wider learning: it also helps with letting participants 'exit' from the simulated environment, a particular issue with larger and more immersive simulations. For assessment that looks at student-centred evaluations, the connection between feedback and assessment is clear. But feedback can also take place at other points. If there is any substantial preparation required prior the main game play, then feedback can be provided on negotiating briefs or positions, in order to ensure participants enter with a more fully prepared approach.

Likewise, it is possible to design interim feedback for longer simulations, although these needs to be done with care, in order not to disrupt proceedings too much from their nature flow. One way of achieving this is to create a two-level game, with the simulation leader (maybe with colleagues) acting as national governments or parties, requiring participants to report back periodically on their progress.

Whenever feedback occurs and regardless of what the focus is meant to be, the process is most usefully driven by the participants themselves, be that through verbal or written contributions. As the participants in the simulation, they have insights into their actions and outputs that might have not been noticed by the game leader or other observers: by giving primacy to their thoughts and reflections, we can strengthen their confidence in self-evaluation and self-criticism. This participant-led feedback can then be supplemented by inputs from observers, documentary evidence (e.g. video, logs from online resources that have been used, observer blogs, etc.), as well as reflection on the simulation *qua* simulation,

this last being instructive in promoting discussion on how the scenario differs from the real-world situation. In all of this, the possible connection to assessment strategy can be usefully explored, with feedback materials matching up to assessment objectives.

CONCLUSIONS

This paper has considered the ways in which assessment can be used with simulations. As noted in the introduction, simulations do not necessarily require assessment, but they do require that the question of need is considered: assessment might well be needed for particular simulations, dependent upon the learning objectives that have to be identified at the outset of the design process. By providing a system of valorisation, assessment offers a range of opportunities to draw students further into the immersive experience that simulations offer (Usherwood 2009) and can reinforce the dynamics underpinning the learning objectives.

At a more practical level, and in keeping with most aspects of simulation design, the approach to assessment should be one of keeping it simple at the beginning. This means recognising limitations – be they pedagogic, practical or institutional – and pursuing an assessment strategy with which one can be comfortable. From that starting point, it is possible to develop more ambitious options, as that sense of comfort expands with each iteration of a simulation.

BIBLIOGRAPHY

- Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of Serious Games: an overview. *Advances in Human-Computer Interaction*, 2013, 1.
- Biggs J. (2003) Teaching for Quality Learning at University, Buckingham, Open University Press.
- Biggs J. (1996) 'Enhancing teaching through constructive alignment', *Higher Education* 32(3): 347-364.
- Chin, J., Dukes, R., & Gamson, W. (2009). Assessment in Simulation and Gaming A Review of the Last 40 Years. *Simulation & Gaming*, 40(4), 553-568.
- Galatas S. (2006) 'A Simulation of the Council of the European Union: Assessment of the Impact on Student Learning', *PS: Political Science & Politics* 39(1): 147-151.
- Gredler M. (1992) Designing and Evaluating Games and Simulations: A Process Approach, London, Kogan Page.
- Knight, P. (2002), "Summative Assessment in Higher Education: practices in disarray", *Studies in Higher Education*, 27(3):275-286.
- Lantis, J. (1998) 'Simulations and Experiential Learning in the International Relations Classroom', *International Negotiation* 3(1): 39-57.
- Newmann W. & Twigg J. (2000) 'Active Engagement of the Intro IR Student: A Simulation Approach', *PS: Political Science and Politics* 33(4): 835-842.
- Raymond, C., & Usherwood, S. (2013). Assessment in simulations. *Journal of Political Science Education*, 9(2), 157-167.
- Schick L. (2008) 'Breaking frame in a role-play simulation: A language socialization perspective', *Simulation Gaming* 39(2): 184.197.
- Usherwood, S. (2009). Enhancing Student Immersion in Negotiation-based Learning Environments. *International Journal of Learning*, 16(7).
- Usherwood, S. (2013). Building resources for simulations in political science. Paper presented to American Political Studies Association Learning & Teaching Conference, Long Beach CA, February 2013.

Usherwood, S. (2014). Constructing Effective Simulations of the European Union for Teaching: Realising the Potential. *European Political Science*. 13(1),

Zeff E. (2003) 'Negotiating in the European Council: A Model European Union Format for Individual Classes', *International Studies Perspectives* 4(3): 265-274.