Assessment in Simulations

Chad Raymond
Salve Regina University, chad.raymond@salve.edu

Simon Usherwood
University of Surrey, s.usherwood@surrey.ac.uk

Follow this and additional works at: https://digitalcommons.salve.edu/fac_staff_pub

Part of the Educational Assessment, Evaluation, and Research Commons, and the Educational Methods Commons

Raymond, Chad and Usherwood, Simon, "Assessment in Simulations" (2013). Faculty and Staff - Articles & Papers. 45.
https://digitalcommons.salve.edu/fac_staff_pub/45

Rights Statement
In Copyright - Educational Use Permitted. URI: http://rightsstatements.org/vocab/InC-EDU/1.0/
This Item is protected by copyright and/or related rights. You are free to use this Item in any way that is permitted by the copyright and related rights legislation that applies to your use. In addition, no permission is required from the rights-holder(s) for educational uses. For other uses, you need to obtain permission from the rights-holder(s).
Assessment in Simulations

Chad Raymond, PhD
Assistant Professor of Political Science and International Relations
Salve Regina University
100 Ochre Point Ave
Newport, Rhode Island, USA
chad.raymond@salve.edu
401.341.3294

Simon Usherwood, PhD
Senior Lecturer and Director of Learning & Teaching
School of Politics
University of Surrey
Guildford, United Kingdom
s.usherwood@surrey.ac.uk
+44.1483.689962

Simon Usherwood is Senior Lecturer at the School of Politics at the University of Surrey in Guildford, UK. He read for his Ph.D. in European Politics at the London School of Economics, following periods of study in the UK, Belgium and Germany. His teaching covers a range of comparative and European politics topics, as well as courses on negotiation in politics. His learning and teaching publications have focused on the optimization of simulations and negotiation-based learning environments in political science.

Chad Raymond is an Assistant Professor in the Department of Political Science at Salve Regina University in Newport, Rhode Island. He received his Ph.D. in Political Science from the University of Hawai‘i after language training and dissertation research in Vietnam. He teaches courses in international relations, comparative politics, and political economy, with a particular interest in political and economic development in Asia and the Middle East. He has published a number of journal articles on the use of simulations in the teaching of international relations.
ABSTRACT

Simulations are employed widely as teaching tools in political science, yet evidence of their pedagogical effectiveness, in comparison to other methods of instruction, is mixed. The assessment of learning outcomes is often a secondary concern in simulation design, and the qualitative and quantitative methods used to evaluate outcomes are frequently based on faulty paradigms of the learning process and inappropriate indicators. Correctly incorporating assessment into simulation design requires that an instructor identify whether a simulation should produce positive changes in students’ substantive knowledge, skills, and/or affective characteristics. The simulation must then be assessed in ways that accurately measure whether these goals have been achieved. Proper assessment can help demonstrate that simulations are productive tools for learning and that their popularity in the classroom is justified.
SIMULATIONS IN THE CLASSROOM

Simulations have been and are being used in a wide variety of academic disciplines (Keller et al. 2010; Koh et al. 2010; Finkelstein et al. 2005; Tonks 2002; Vaidyanathan and Rochford 1998; Rodgers 1996). Simulations were first employed in political science as a means of understanding complex social processes that did not lend themselves to experimental testing or analytical evaluation, and their use soon expanded in scope to include the teaching of political science itself (Guetzkow and Jensen 1966:264; Dorn 1989:1; Lantis 1998:41-42). Such “educational” simulations typically involve the use of role-play, gaming, computer models, and/or negotiation and bargaining. During a simulation, students’ applications of these skills are constrained by rules that are intended to allow them to experience a high degree of complexity while prohibiting them from engaging in actions that are impossible in the real world (Lane 1995:607; Hensley 1993:64). In some cases, participants in a simulation interact only with a computer-based, mathematically-driven representation of reality; more frequently, the simulation functions as a platform for student to student interaction (see Williams and Williams 2010; Mayer 2009; Baranowski 2006; Chasek 2005; Hobbs and Moreno 2004).

Though they are attractive to political science instructors for a number of reasons, simulations can negatively affect student learning outcomes if they lack clarity in their learning objectives, replace other, more effective teaching methods, require an
inordinately complex application of skills or technology; do not incorporate adequate levels of pre-simulation training; or lack meaningful post-simulation assessment mechanisms (Lane 1995: 615-6).

This paper focuses on the last of these issues. Simulations are popular instructional tools in the political science classroom, but it remains unclear whether or why simulations are pedagogically productive, despite the collection of large amounts of quantitative and qualitative data over a period of decades (cf Heitzmann 1973; Wentworth and Lewis 1975; Smith and Boyer 1996; Lantis 1998; Lean et al 2006:230) – suggesting that the practice of assessing simulations needs to be approached in a more systematic fashion.

Institutions of higher learning are increasingly asked to defend curricular and pedagogical outcomes, and faculty are now under pressure to close the assessment loop (Walvoord et al. 1998). Demonstrating that simulations are effective and productive tools for learning, rather than simply enjoyable exercises for students, is critical. University faculty must ask themselves what a simulation adds to a student’s knowledge base that cannot be learned more efficiently in a traditional classroom setting, and how can this be measured. This has been particularly apparent in countries where there are national frameworks for discipline content and quality assurance. In the UK, for example, the requirement for all assessment to be validated ex-ante and...
moderated *ex-post* by external assessors requires both an objective justification for the inclusion of simulations and a transparent mechanism for their assessment.

**LEARNING**

Simulations are broadly believed to foster learning in three ways. First, simulations are thought to improve students’ motivation to learn by delivering information to students in a way that heightens their interest in understanding it. It has long been recognized that motivation is an important factor in learning (Lei 2010; Tuckman 1996; Ehrman and Oxford 1995; Weiner 1990). Educators often assert simulations enhance students’ extrinsic (reward-driven) and intrinsic (satisfaction-driven) motivation because simulations require active engagement, provide prompt feedback, present challenging but achievable goals, have uncertain outcomes, and generate feelings of autonomy, accomplishment, relevance, relatedness to others, and pleasure among participants (Carnes 2011; Koh et al. 2010; Mitchell and Savill-Smith 2004). According to de Freitas (2006:350),

> [o]ne of the main perceived advantages of using games and simulations [is] increased motivational levels for learners . . . The majority of those experts interviewed thought that simulations and games significantly improved learner motivation . . . In particular, games-based learning content was found to be particularly positive.
Parsing the relationship between simulations, motivation, and student learning can be problematic. Much of the evidence reported in the literature is anecdotal (Stroessner 2009:605; de Freitas 2006:35) and cause and effect are not clearly identified. Motivation “often is inferred from learning [but] learning usually is the indicator of motivation” (Weiner 1990:618). Learning is influenced by many factors in addition to motivation, and emotion is both a product of and an influence on motivation (Katt and Condly 2009:219; Hannula 2006:175). Intrinsic motivation will be high only if students’ expectancy of success and the value placed on success are high (Ehrman and Oxford 1995:68).

Students’ extrinsic motivation is affected by the type and size of rewards that are used. Rewards mean different things to different individuals, and “each connotation can have different motivational implications” (Weiner 1990:618), making it exceedingly difficult to demonstrate that rewards reinforce desired behaviors.

Some students are reluctant to engage in the competitive, group-oriented, or role-play activities that are part of some simulations, which is likely to decrease their motivation (de Freitas 2006:351). It has been found that highly structured teaching methods can maximize the motivation of some students, while students’ perceptions of a lack of instructor control of discussions or activities and unclear expectations can act as powerful demotivators for other students (Hancock 2002; Katt and Condly 2009). Given that simulations often cede classroom control to students and favor open-ended
outcomes to predetermined ones, it is likely that simulations have important effects on student motivation that have yet to be adequately explored.

Second, simulations are assumed to favorably alter the environment in which learning occurs – e.g., by replacing a dyadic relationship in which a teacher delivers content to a student who must passively receive it with activities that promote peer-to-peer instruction and self-discovery. Claims that simulations alter the learning environment in such a fashion typically rest on an epistemological distinction between thought and experience. This distinction is commonly associated with the American educational philosopher John Dewey (1916) and was more fully elaborated by Kolb (1984).

Kolb (1984) theorized that learning is facilitated if students are immersed in a concrete experience that requires their active participation. Students initially observe and reflect upon the effects of their behavior then they form abstract generalizations as they seek to understand overarching principles that may apply to what they have experienced. Lastly, students test these generalizations against new observations. According to Kolb (1984:30) the processes of experience, reflection, abstraction, and experimentation are all necessary components of learning. In contrast, traditional forms of instruction require that students first passively receive information from texts or lectures and then understand it. Students next “infer particular applications of what is learned to general principles . . . [and finally] they learn to use the general principles to act in some way”
Due to the delay between students’ first encounters with new knowledge and the opportunity to apply it, the relevance of and incentives for learning may not be apparent to students during much of the learning process.

While Kolb’s theory has been much acclaimed (Brock and Cameron 1999), validation of it is rare in the social science literature (Mandel 1987:339). Authors who cite its validity often rely on anecdotal information (see Brock and Cameron 1999:254; Stice 1987:293 and 296) or misquote the work of other scholars (see Starkey and Blake 2001:537-538; Winham, 1991:417; Brademeier and Greenblat 1981:316 and 320; Brewer 1984:805 and 810).

Assumptions about learning environments are also frequently tied to a belief in typologies of learners, such as visual, auditory, and kinesthetic “learning styles,” personality inventories like the Myers-Briggs Type Indicator, or Howard Gardner’s (1983) “multiple intelligences.” Educators often interpret these typologies as indicating that learning is maximized if forms of instruction match distinct cognitive characteristics of students (cf Moizer et al 2009:213). Simulations, because of their use of role-play, digital visual representations, and other features, are thought to be better at addressing these characteristics than traditional teaching methods.

The problem with basing teaching techniques on learning typologies is that students' expressed preferences for particular forms of instruction do not correlate to observable cognitive or skill aptitudes, and only a handful of studies citing the existence
of “learning styles” have conducted valid experimental tests (Paschler et al. 2008). In a review of thirteen commonly used learning style inventories, Coffield et al. (2004:139) found that twelve did not meet one or more basic criteria for internal consistency, test-retest reliability, construct validity, and predictive validity. The field of learning styles “is bedevilled by vested interests because some of the leading developers of learning style instruments have themselves conducted the research into the psychometric properties of their own tests, which they are simultaneously offering for sale in the marketplace” (Coffield et al 2004:137).

The hypotheses that simulations improve student motivation and the learning environment are usually implicitly linked to a third hypothesis -- that improvements in student motivation and the learning environment created by simulations lead to gains in substantive learning about external processes or phenomena (content knowledge) or the self (affective knowledge). This third hypothesis is rarely investigated (Brademeier and Greenblat 1981:319).

CURRENT ASSESSMENT METHODS

The assessment of simulations for instructional efficacy should include testing of one or more of the above hypotheses with appropriate methods; however, such testing is rarely explicitly performed. In many cases, the qualitative assessment of simulations’ pedagogical value has relied upon the anecdotal and subjective impressions of the instructor in combination with student self-reflection, commonly in the form of oral or

\[ \text{Comment [SMU1]: Lose all of this?} \]

Deleted: Kolb’s Learning Styles Inventory (LSI) in particular has been found to be “an unreliable instrument designed in such a way that its results spuriously support” Kolb’s theory (Freedman and Stumpf 1980:446-7); according to Basuray (1982:62), whatever the LSI measures is “measured by chance.” Other studies, such as Vaidyanathan and Rochford (1998), have found that student academic performance is driven more by a preference for traditional methods of learning (e.g., reading) rather than “kinesthetic” instructional exercises. Even Howard Gardner (1995), the originator of “multiple intelligences,” cautions that intelligence is a capacity rather than a learning or cognitive style and that types of intelligence cannot be accurately identified by standardized instruments.

While individuals may prefer particular methods of acquiring new information, cognitive researchers broadly agree that learning involves a wide variety of mental processes (Willingham 2009) and that preferences for one kind of instruction do not indicate an in ability to learn from other methods; for example, students “who prefer examples with concrete numbers to abstract mathematical expressions may be responding to a lack of familiarity with algebra rather than a lack of innate ability” (Redish 2002: 37). In general, personality inventories show little correlation with proficiency and “personality preferences give no information about the cognitive level of the individual” (Ehrman and Oxford 1995:82).

Evidence also suggests that students employ methods of learning that are context-dependent -- students are cognizant of their environment and their approach to learning is determined by their interpretation of that environment. It would therefore be hazardous for an investigation of learning to proceed on the assumption that learning is a process that is independent of other external factors, or that students possess inherent, invariant sty...

However, as noted by Stroessner et al. (2009:608), a “sizable literature suggests that people tend to be quite inaccurate in judging both the nature and the extent of the impact of experience.” Although students can provide valuable information about how they perceived what happened during a simulation, this information should not be automatically construed as a reliable indicator of the influence of the simulation on their learning, or even on their responses to the questions (Wilson and Nesbitt 1978:130). After an event, individuals “try to detect or retrieve causal antecedents that potentially caused the outcome,” often leading to hindsight bias (Nestler and von Collani 2008:482). Conversely, students’ perceptions of a simulation can also be affected more by the simulation’s confirmation or disconfirmation of prior beliefs than by their experience of the simulation itself – a situation of confirmation bias (Maznick and Zimmerman 2009:34; Eiser et al. 2008:1023). As pointed out by Brademeier and Greenblat (1981:318) “students’ liking the experience may not necessarily mean they learned anything from it.” Student introspection about the causes of their behaviors during a simulation also may not reveal the true causes of those behaviors (Wilson and Nesbitt 1978:118); for example, Steck, Buonnano, and Eagles (1996:24) reported that
students from the same school tended to have the same opinions. Reactions were less individualized that one might expect and apparently more shaped by the student’s immediate peer group. To the extent that one might hypothesize that a ‘good’ overall experience produced better ‘learning,’ a good collective experience . . . would enhance both group and individual learning.

Quantitative findings on simulations and learning are mixed, despite several decades of investigation. Frederking (2005) and Baranowski (2006) reported a statistically significant improvement in exam scores among students who participated in simulations in American government courses, while Krain and Lantis (2006:404) found that both a diplomacy simulation and traditional lecture and discussion had “statistically significant positive effects on student learning, regardless of instructor or issue area,” though possibly in different ways. Powner and Allendoerfer (2008) concluded that students who participated in a brief role-play activity scored better on multiple choice questions after the activity than students who participated in classroom discussion, but that there was no statistically significant differences in the overall performance of the two groups. Stroessner et al. (2009:614) found that in comparison to traditional instruction, Reacting to the Past role play exercises produced no statistically significant benefit in writing skills among first-year college students, and that the benefit to students’ rhetorical skills was “marginal.” Raymond (2010) found no statistically significant improvements in exam scores among students who participated
in a role-playing simulation compared to students who received traditional lectures and assignments in an international relations course.

In general, the ability of simulations to effectively produce desired learning outcomes in students has not been rigorously or comprehensively assessed. Smith and Boyer (1996:693-694) state that though “large amounts of anecdotal evidence [support] the idea that simulation promotes greater depth of understanding [and] stronger critical thinking and analytical skills . . . none of this information has been collected, standardized, or quantified,” and Lantis (1998:51) writes that “there are very few studies that confirm our experiences (and convictions) that [role-playing simulation] exercises are truly effective.” Krain and Lantis (2006:399, 400) find that “very few studies confirm our experiences (and convictions) that [simulation] exercises are truly effective methods for teaching political science and international relations” because such exercises “have remained generally untested in any rigorous fashion.”

EFFECTIVE ASSESSMENT

In order for assessment to be effective, both in terms of guiding students towards particular learning objectives and of providing meaningful measure of student attainment, then a number of points must be addressed. Moreover, there is a need for an overarching framework and conceptualization, since assessment must perforce be an integral part of the learning experience. Even in academic systems where assessment is placed in an extensive regulatory and external framework (e.g. Quality Assurance
Agency (2011)), Knight (2002) has argued that assessment practice has long been in
disarray and that central to any reform is the need to restructure curricula in order to
reconceptualize assessment as a communicative practice, by which is understood the
interaction between educator and student. Seen as such, much of what follows builds

We might observe here that where authors have presented their simulations,
their consideration of how assessment fits into their design is typically scant. Thus,
McCarthy & Anderson (1999) provide an evaluation of simulations against
conventional learning, but without exploring whether their assessment pattern allowed
students to properly access all of their learning, whatever form it took. Likewise,
Dougherty (2002) makes no mention of assessment as part of the simulation and
authors including Kaunert (2009) and Crossley-Frolick (2010) make only passing
reference. As such, Chin, Dukes and Gamson (2009) are a notable exception, although
this must be seen in the context of that work’s primary aim of reviewing assessment.

Firstly, there has to be a clear understanding on the part of the simulation
designer of what the simulation is intended to achieve. This starts from an initial
evaluation of whether simulations preclude the delivery of information critical to
student performance, so necessitating the provision of other learning environments. The
demands of the simulation may have also consumed students’ time and energy outside
of class in ways that might have been unproductive for learning (O’Toole and Absalom
Providing students with outlines or other instructional supplements, decreasing the amount of time in and out of class devoted to the simulation, or re-orienting the simulation’s online component away from collaborative writing assignments and toward individual knowledge assessment might improve student performance.

Assuming that a simulation can provide that appropriate learning environment, then the designer must identify the specific learning outcomes that are to be achieved, be it substantive knowledge acquisition, skills development or group socialization. This is essential for creating appropriate assessment regimes. A methodologically well-designed simulation integrates student participation, learning objectives, and learning outcomes (see Facer et al., 2004; Egenfeldt-Nielsen, 2005). If this integration is absent “simulations are not effective in achieving their broader learning outcomes regardless of how motivated the learners are during game-play. This issue has not been of such concern with using simulations, where the simulation and the learning context are clearly connected together, allowing learners upon reflection to take full advantage of understanding how simulations relate to a real-life task, occupation or set of experience.” (de Freitas 2006 p. 350). Thus, one of the authors runs a course on ‘negotiating politics’, where the development of specific aspects of negotiation skills is the core objective, so the entire assessment is based upon a dossier of reflective pieces, where students analyze and evaluate their practice in the classroom against the relevant
literature. Such an alignment of learning objectives and assessment makes good pedagogic sense (see Biggs 1999), in facilitating deep learning. It also is essential in those systems where there is a requirement for individual courses to contribute to an overall program’s learning objectives or where external moderation of assessment takes place.

In a second step, the simulation or game (in the narrow sense) must be clearly aligned to the assessment regime. This is obviously an extension of the previous point, in that the gameplay must allow for the learning objectives and the assessment to be achieved satisfactorily. This is perhaps more of an issue than is immediately apparent with simulations, for the simple reason that simulations are so adaptable in so many different dimensions. Thus a legislative simulation might cover just one sitting or an entire legislative process; it might be entirely closed, with predefined positions, or entirely open, with students taking positions of their own; it might be an illustration of a lecture or the entire course; it might be about formal procedures or informal bargaining, or both. All of these choices can be valid ones, but they need to be carefully aligned to any assessment. Thus Kaunert (2009) uses a mix of course paper and reflective discussion to connect with his students’ European Union simulation with their wider reading on that organisation’s institutional architecture (see also Usherwood 2009).
This process of alignment does require recognition of the simple fact that simulations are unusual learning environments, in the sense that students rarely encounter them. Students can find simulations, especially in their initial phases, confusing if the purpose of the simulation is difficult to grasp. Although some students might be, as described by Lantis (1998:49), “more comfortable with the rules of procedure from the outset and [are] ready to use them to their own advantage,” others can be hesitant to act. Consequently, both the simulation and the assessment should be designed to take account of such disruptive effects. Typically, this is addressed by focused assessment on the students’ appreciation of the activity and their agency within it, rather than on ‘winning’, although it must be recognized that many political science and IR simulations do revolve around competitive behavior. Such considerations also address the possibility of free riders. Rivera and Simons (2008:302) recommend that simulations be designed to include several checkpoints that allow students to be assessed on both individual efforts and group results. Individual writing assessments can be graded before being used to create a collaborative version in the form of a negotiating brief. Likewise, grades for both individual and group tasks could also serve to prevent an unequal division of responsibilities.

Thirdly, any assessment has to be designed with a strong element of debriefing and feedback within it. As Petranek, Corey, and Black (1992) note it is through oral debriefing and written reflection that the full benefits of simulations can be realized.
This builds on the understanding that simulations and games are not self-teaching and so require an externalization of students’ reflections (see Gillespie 1973): while students do internalize practice during the simulation, this is typically an implicit process, rather than an explicit one. Feedback depends upon the ability of the student to recall, discuss and thus analyze their actions and motivations. Seen as such, assessment can be a powerful way by which to encourage students to reconnect the experience of the simulation to other areas of their learning; an obvious example here would be the use of reflective discussion, wherein the student can benchmark their experience against academic literature (both substantive and procedural). In this way, assessment can contribute to - and by informed by - a wider set of feedback processes, such as post-game group discussions, reviewing of game materials (such as video or online caches) and observer- or peer-led debate (see Petranek 2000 for a discussion of the added value of written debriefing).

CONCLUSION

It would not be stretching the point too far to say that assessment is both students’ and educators’ least enjoyable part of the learning experience. Too often, it is disconnected from the rest of the process, a chore to be endured, rather than a valuable learning moment. In this paper, we have argued that this has often been even more of an issue in relation to the use of simulations, since the focus of attention has very largely been on the design of scenarios and game-play. However, if the full potential of
simulations is to be realized, then assessment has to be integrated from the start, informed by learning objectives chosen by the instructor and incorporated into simulation design. Educators must obtain the feedback necessary to understand the potential benefits and costs of using simulations and the impact that they can have on learning outcomes: “If educators assume that it is important to use data to assess student learning, then they should be using data to make decisions about how they teach” (Ebert-May et al. 2003:1221).

Simulations offer a potentially deeply rewarding learning environment, in which students can practise and actively develop a wide range of skills and knowledge, leading to creative outputs and perspectives on complex topics of study. This is not an automatic outcome, however, and depends upon the educator identifying what learning objectives students are to achieve, what is the most appropriate pedagogy for facilitating learning, and how well outcomes have been achieved (Gilchrist and Zald, 2008). Thus assessment becomes a central concern, rather than an add-on.
REFERENCES


Quality Assurance Agency (2011) *Code of practice for the assurance of academic quality and standards in higher education*


Kolb’s Learning Styles Inventory (LSI) in particular has been found to be “an unreliable instrument designed in such a way that its results spuriously support” Kolb’s theory (Freedman and Stumpf 1980:446-7); according to Basuray (1982:62), whatever the LSI measures is “measured by chance.” Other studies, such as Vaidyanathan and Rochford (1998), have found that student academic performance is driven more by a preference for traditional methods of learning (e.g., reading) rather than “kinesthetic” instructional exercises. Even Howard Gardner (1995), the originator of "multiple intelligences," cautions that intelligence is a capacity rather than a learning or cognitive style and that types of intelligence cannot be accurately identified by standardized instruments.

While individuals may prefer particular methods of acquiring new information, cognitive researchers broadly agree that learning involves a wide variety of mental processes (Willingham 2009) and that preferences for one kind of instruction do not indicate an inability to learn from other methods; for example, students “who prefer examples with concrete numbers to abstract mathematical expressions may be responding to a lack of familiarity with algebra rather than a lack of innate ability” (Redish 2002: 37). In general, personality inventories show little correlation with proficiency and “personality preferences give no information about the cognitive level of the individual” (Ehrman and Oxford 1995:82).
Evidence also suggests that students employ methods of learning that are context-dependent -- students are cognizant of their environment and their approach to learning is determined by their interpretation of that environment. It would therefore be hazardous for an investigation of learning to proceed on the assumption that learning is a process that is independent of other external factors, or that students possess inherent, invariant styles of learning (Laurillard 1979:408).