negotiation and conflict management research

INTERNATIONAL ASSOCIATION IACM FOR CONFLICT MANAGEMENT

Negotiation and Conflict Management Research

Taking Action to Reduce Waste: Quantifying Impacts of Model Use in a Multiorganizational Sustainability Negotiation

Ellen Czaika[1] and Noelle E. Selin[2]

1 Institute for Data, Systems and Society, Massachusetts Institute of Technology, Cambridge, MA, U.S.A.

2 Institute for Data, Systems and Society and the Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, U.S.A.

Keywords
sustainability negotiation, Life Cycle Assessment Model (LCA), negotiator cocreated model, expert-made model, model use.

Correspondence
Ellen Czaika, Institute for Data, Systems and Society, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, U.S.A.; e-mail: eczaika@mit.edu.

Abstract

We use a role-play simulation to examine how using quantitative models influences the process and outcome of sustainability negotiations. Our experimental approach involved 74 teams of five parties negotiating the details of a pilot test to compost and/or recycle used paper coffee cups. Approximately half of these negotiation teams were given a quantitative model—a life cycle assessment (LCA). We measured both negotiation process and outcome variables, in particular identifying favorable agreements—the mutually exclusive set of agreements that either minimized carbon dioxide emissions or maximized the parties' collective earned value. We found that most teams used a quantitative model; nearly half of those cocreated their own while negotiating. In our sample, teams that used a model, even those cocreating the model while negotiating, reached agreement more quickly than teams not using a model. Teams that cocreated the LCA reached a higher number of favorable agreements. We observed two dominant manners of model use: using the model to test alternatives while developing an agreement and verifying that a tentative agreement would sufficiently reduce carbon dioxide emissions. We conclude that using a quantitative model during a sustainability negotiation can help to increase the chances of obtaining a favorable agreement without lengthening the negotiation duration.

Introduction and Theory

Addressing complex environmental and sustainability challenges often requires collaboration among stakeholders to design and implement solutions. Sustainability negotiations frequently involve some issues that are based on physical and environmental constraints and other issues that involve stakeholder preferences. Collaborative modeling has been applied in multistakeholder environmental challenges as a means for stakeholders to make sense of the physical world components of sustainability. However, much remains unknown about the degree to which model use affects multistakeholder negotiations, and

We would like to thank Starbucks Coffee Company for partial funding, Paper Recovery Alliance for allowing us to attend meetings, Lawrence Susskind for consultation on the game matrix, and all those who participated in the Cup Game. We especially thank Ofer Sharone, who provided feedback on the game development and who allowed us to collect data in his classes.

Volume 9, Number 3, Pages 237–255

what this implies for best practices in using quantitative models. By using a serious gaming simulation to provide repeated and comparable instances of the same contextual situation, varying only model use, this study quantifies the influence of model use on negotiation outcomes, and identifies negotiation process insights about how modeling can inform complex, multistakeholder negotiations.

A substantial body of work has examined the use of modeling in environmental policy applications. Modeling enables the study of full physical systems including interrelated components (Dowlatabadi, 1995; van Delden, Seppelt, White, & Jakeman, 2011) and allows for the consideration of multiple different futures and possibilities (Langsdale et al., 2009; Morgan, 2011). In particular, researchers who study collaborative modeling (also called mediated modeling and cooperative modeling, among other names (Tidwell & van den Brink, 2008) and used interchangeably herein) typically examine situations involving complex decisions with many stakeholders. These studies are frequently set in the public sector. The models used tend to be complex, large, and predictive in nature and are frequently dynamic. Furthermore, there is often a focus on expert modelers working alongside decision makers and stakeholders. For example, system dynamics modelers who practice mediated or participatory modeling (Langsdale et al., 2009; van den Belt, Schiele, & Forgie, 2013) may employ expert modelers to collaborate with stakeholders to build a system dynamics model to address the stakeholders' problem. Likewise, integrated assessment researchers often work with stakeholders directly (see Mustajoki, Marttunen, Karjalainen, Hokkanen, & Vehmas, 2013). Collaborative modeling research has focused on evaluating the participatory process. Collaborative modeling involves the coproduction of knowledge (as defined by Funtowicz & Ravetz, 1993, 1994); some scholars identify social or participant learning as an important component of collaborative modeling (Ducrot, van Paassen, Barban, Daré, & Gramaglia, 2015; McIntosh et al., 2011).

Much of the research assessing collaborative modeling utilizes a case study approach, as for example, managing the Ria Formosa coastal zone in Portugal (Videira, Antunes, Santos, & Gamito, 2003; Videira, van den Belt, Antunes, Santos, & Gamito, 2004) and the decisions about sage-grouse protection and land use in Washington state (Beall & Zeoli, 2008). Videira, Antunes, Santos, & Gamito (2003) and Videira, van den Belt, Antunes, Santos, & Gamito (2004) found that the participants credit the collaborative modeling process with organizing their ideas and conversations. Beall and Zeoli (2008) concluded that when stakeholders and experts learn how to work together, scientific findings can be coalesced with long-term planning. Although the case study method has many benefits (e.g., using the actual decision makers (see van den Belt, et al. (2013) for another example)), a drawback is this method's inability to test repeatedly with the same situation. Testing repeated instances in the same situation allows for statistical analysis and enables comparison across controlled differences. Because of the complexity of the case studies in collaborative modeling research and their protracted timeline, it is often hard to isolate concepts of interest such as the manner in which a group uses a model and the impacts model use has on the resulting outcome.

In contrast to a case study approach, serious gaming role-play simulations offer an opportunity to identify, in a more controlled setting, the pathways of influence by which model use might impact collective sustainability decision making. Many serious gaming studies combine unstructured social interactions with a dynamic computer simulation (for examples, see Kuit, Mayer, & de Jong (2005), Mayer, Carton, de Jong, Leijten, & Dammers (2004), Mayer, Meijer, Nefs, Gerretsen, & Dooghe (2010), Meijer, Mayer, van Luipen, & Weitenberg (2012), and Nefs, Gerretsen, Dooghe, Mayer, & Meijer (2010)). These computer simulations are central to the game and the participants' experience with it. Other types of serious games exist (e.g., role-play, board games, behavioral (Klabbers, 2005)). Serious games that do not use a computer simulation can also involve scientific information (see, for example, Schenk and Susskind (2015)); many of these are role-play simulations. Often, role-play simulations of environmental situations are used to educate the participants (see, for example, MaKinster (2010) and Stokes and Selin (2014)). Other role-play simulations are used to represent real-world negotiations and study negotiators and negotiations (for examples, see Curhan, Neale, & Ross (2004), Curhan, Elfenbein, & Kilduff (2009), Curhan & Pentland (2007), De Dreu, Beersma, Stroebe, & Euwema (2006), and

Elfenbein, Curhan, Eisenkraft, Shirako, & Baccaro (2008)), although these are often not set in environmental or sustainability situations. Serious gaming leaves the social interactions unfettered (Corrigan, Zon, Maij, McDonald, & Mårtensson, 2015), which enables the study of, as Mayer (2009) terms it, the "chaotic and messy" collective decision process.

In the following sections, we present the methods used, the results from the 74 teams of five participants—throughout this article, we refer to the participants negotiating with each other around a table as a "negotiating team," not in the sense that they are organizationally a team, but rather in the more common usage that they are "two or more people working together" ("Team," 2016)—that have played the Cup Game, and discussion. In the section *Research Design*, we describe the serious game role-play simulation, the experimental setup and the data we collected, and the research questions. In *Research Results*, we discuss our analytical methodology and the findings, classified by negotiation process and the negotiated outcome. In *Discussion and Conclusions*, we discuss the real-world relevance of the findings and suggest future research.

Research Design

This study complements existing studies of collaborative modeling by using methods that can compare repeated instances of the same situation. Using a serious gaming role-play simulation research design, we use the game to identify whether using a life cycle assessment (LCA) of the carbon dioxide (CO_2) emissions reduced in the new postconsumer paper coffee cup system influenced the process or outcome of negotiations, or both. We also investigate how the authorship of the model (whether it is cocreated by the negotiating parties or made by experts) and two different means of using the model impact these elements. In the game, teams of five negotiators are asked to make decisions about the used paper coffee cup system, and a subset are given an LCA model.

The Simulation Design

The Cup Game is a negotiation role-play simulation loosely based on the initiative Starbucks Coffee Company launched to recycle and/or compost used paper coffee cups wherein both recycling and composting these cups entail technological uncertainty. The Cup Game's roles and issues are informed by workshops and interviews held in 2009 and 2010 with individuals involved in this initiative (Czaika, 2010). In these role-play simulations, each participant was randomly assigned a negotiating table and one of the five roles at the table: a coffee retail company (the convening party), a cup manufacturing company, a recycling company, a compost company, and a waste collection (hauling) company. Each such group of five negotiators—called a negotiating team—sought to make decisions about an upcoming pilot test to determine the feasibility of composting and recycling used paper coffee cups. These decisions took the form of five issues each with two to five alternatives. Each role (party) had individualized utilities—called preference points—associated with each alternative; participants were told to keep these points confidential. For an agreement to be valid, all five parties had to meet or exceed their individualized point thresholds. It was possible for teams not to reach agreement.

In approximately half of the teams, one party (the cup manufacturer) was given a life cycle assessment (LCA) model, and the encouragement to "share it [the LCA model] with the other parties to determine the amount of carbon dioxide saved with different alternatives for [some issues]." When, how, and whether to share the model were left up to the discretion of the participant. In the other half of the teams, none of the parties received an LCA model. Each team received all the data necessary for the LCA, regardless of whether or not the team received the set of equations forming the LCA (the model). However, these data were distributed among the five parties. Participants had to decide whether, when, and how accurately to share the LCA data they were given. The participants themselves—on some teams the cup manufacturer alone and on others the parties collectively—determined whether or not to use a model; that is, a team of negotiators that was not given the model could cocreate it and a team given the model could ignore it. Furthermore, negotiators were free to use the model however they wanted during the negotiation.

The LCA model ruled out some alternatives on two of the issues as not fulfilling the environmental goal that the parties were working toward—reducing carbon dioxide emissions by 110,000 lbs. (49.9 metric tons) in the pilot test. If a team's agreement met or exceeded 110,000 pounds of CO_2 savings, each of the five parties received bonus points added to their negotiated points. Alternatively, teams could reach this environmental goal strictly through negotiation based on their preference points. The parties were not told outright that they share this same specific environmental target and the same reward for reaching it, although they could establish this shared goal during their negotiation talks.

The research design forced the participants to choose between maximizing their point value and maximizing the environmental benefits of their agreement—measured in the amount of reduction in carbon dioxide emissions. Furthermore, we designed the game such that the difference between the amount of carbon dioxide emissions reduced in the value maximizing agreement (the second largest amount of emissions reduction) and the amount reduced by the set of best environmental agreements (largest reduction) could not be easily estimated without the use of a model. That is, the difference among the environmental consequences of the alternatives is small enough that the participants could not estimate the impact each combination of alternatives would have on the reduction in carbon dioxide emissions. Therefore, in choosing among the alternatives on the environmental issues, the parties had either to calculate the emissions reduction using a LCA, or to use their respective personal utilities.

Experiment Setup and Data

The participants were diverse professionals and students in negotiation and sustainability-related fields from different world regions. Butler (1991) demonstrated that the outcomes of student subjects matched those of experienced professionals and Herbst and Schwarz (2011) confirmed that students trained in negotiation perform comparably well to professional negotiators. We collected data from 370 participants grouped into 74 teams of five participants at each negotiating table. Approximately half of the 74 teams were given an LCA. Fifty-nine teams used a model (80%), with 32 teams using the given model and 27 creating their own. Seven teams did not use the LCA they were given, and eight were neither given an LCA nor created one (Table 1). Seventy-three teams negotiated an agreement; one team failed to reach agreement. We include this latter team in our analyses because "no agreement" is a possible outcome.

We collected presurvey and postsurvey responses from participants (n = 291 presurvey and n = 215 postsurvey). The survey data served to help rule out alternate explanations and to gather information about the participants' experience playing the Cup Game. To rule out the explanation that model usefulness depends on the individuals' comfort with quantitative data and models, participants indicated their comfort with numerical data. The presurvey also helped to establish their initial opinion about the use of numerical data in sustainability negotiations, and their starting opinion of each role's commitment to sustainability. The postsurvey repeated the above measures, in addition to asking more questions about

Table 1
Distribution of Data Across Initial Condition and What the Teams Did With That Initial Condition

	Used an LCA	Did not use an LCA	Totals
Given LCA	32	7	39
Not given LCA	27	8	35
Total	59	15	74

their use of the LCA model, how much they trusted the data other parties shared, and their opinions about the outcome.

Some participants were interviewed (n = 37 interviews) 1–2 weeks after the role-play simulation to gain more insight into how they used the model or why they did not use it. The semistructured interviews were between 20 and 30 min, were recorded if participant permission was given, and transcribed. The interview data helped to triangulate whether and how teams used the LCA model, to provide insight into the nature of the conversation during the negotiation, and to elucidate participants' opinions about model use in sustainability negotiations and their opinions about the roles' commitment to sustainability.

Research Questions

We use our role-play simulation to investigate questions related to negotiation process and outcome. We expect that model use will impact the negotiation by providing a structured means for the negotiators to make sense of the alternatives they are selecting among, as researchers studying modeling, such as Dowlatabadi (1995), Morgan (2011), and van Delden et al. (2011), have found happens in settings other than negotiations. We expect models to provide this structured means of assessing alternatives either while the parties are creating an agreement package or while they are verifying a tentative agreement. Given the similarities of the participants (e.g., in self-reported comfort with quantitative information, in self-assessed trust for other parties, in perception of the roles' commitment to sustainability), we can reasonably attribute the observed negotiation outcome and process differences to the difference between model use and the lack of model use and differences in model authorship (i.e., cocreating a model versus using an expert-given model).

We analyzed the data in two sets of categories: (a) the two-way comparison of model used (combining both types of model authorship) versus no model used and (b) a three-way comparison considering model authorship (teams that used the LCA they were given, teams that cocreated an LCA during their negotiation, and teams that did not use a model). The category "no model used" comprises teams that did not use the LCA they were given and those that neither received an LCA nor created one. Treating these teams as one category is consistent with the focus of our research questions. We are interested in comparing teams that use a model with those that do not use a model; therefore, exploring the differences between negotiators that ignored a model and those that did not cocreate one is left for future research. To be counted as having cocreated their own model, the negotiators in a team had to report using a model and the team must not have received the expert (given) model. For most teams, we also observed whether they used a model during the negotiation. We collected and analyzed many of the models cocreated by these teams; at least one contained a mistake.

Regarding the negotiation process, we hypothesize (H1a) that teams who use a model will reach agreement more *quickly* than those who do not use a model; (H1b.1 and H1b.2) that teams who cocreate the model will reach agreement more *slowly* than teams who use the given model (H1b.1) and than teams who do not use a model (H1b.2); and (H1c) that teams who use the given model will reach agreement more *quickly* than those who do not use a model. Using a given model may help the negotiators make sense of the alternatives they are deciding among, speeding up their selection of a viable agreement. When the negotiating parties cocreate their own model during the course of the negotiation, they are taking time away from negotiating to create a tool to help them make sense of the alternatives. Because of this competing use of their time, we expect that teams of negotiating parties that cocreate a model will have longer negotiation durations than teams of negotiators that do not use a model and than teams using the given model.

We expect that participants will use models in multiple ways. First, the negotiators can consult the model as each new alternative is discussed and build an agreement based on the model's output. Alternatively, the negotiators can negotiate as normal and then verify that an agreement they have tentatively reached reduces the carbon dioxide emissions satisfactorily. A third possibility, as elaborated below, is that some parties could use the model secretly without letting the other parties know that the model exists, and use it for individual gain. Therefore, our second process hypothesis (H2) is that model-using teams will use the model in one of three distinct manners: collectively to test alternatives while building an agreement, collectively to verify a tentatively accepted agreement, and individually to benefit personally (by the party introducing the model). Because model use provides more insight into the alternatives being considered (Dowlatabadi, 1995; Morgan, 2011; van Delden et al., 2011), it is possible that a negotiator could use the given model to capture more value for himself or herself. To investigate whether this is happening, we assess the value distribution among the negotiating parties. In teams that cocreate a model, the model logic is more transparent and discussed. Therefore, these teams are less vulnerable to individual parties not sharing the model insights than are teams using a model given to them.

With respect to negotiation outcome, we hypothesize (H3a) that the percentage of teams who reach a favorable agreement will be higher for teams who use a model (whether given or cocreated) than for teams who do not use a model; favorable agreements are the set of twelve agreements that minimize carbon dioxide emissions and the one agreement that maximizes the negotiated value, which are mutually exclusive. Model use will help the negotiators more clearly understand the environmental consequences of alternatives (Dowlatabadi, 1995; Morgan, 2011; van Delden et al., 2011). With that understanding, the parties can decide whether to maximize the environmental savings, at some detriment to the negotiated value, or to maximize their negotiated value, at some detriment to the environmental outcome. Furthermore, when a team of negotiating parties cocreates a model for use during the negotiation, their level of understanding of the context is higher than if they had blindly used a model given to them. Therefore, we hypothesize (H3b.1 and H3b.2) that the percentage of teams who reach a favorable agreement will be higher for teams who cocreate a model than for teams who use the given model (H3b.1) and than for teams who do not use a model (H3b.2). We hypothesize (H3c) that the percentage of teams who reach a favorable agreement will be higher for teams that use the given model than for teams that do not use a model. Furthermore, when considering the value of the negotiated agreements, we hypothesize (H4a) that teams who use a model (whether given or cocreated) will reach agreements with higher team scores than those who do not use a model. Additionally, we expect that model authorship will also make a difference: we hypothesize (H4b.1 and H4b.2) that teams who cocreate a model will reach agreements with higher team scores than those who use the given model (H4b.1) and than teams not using a model (H4b.2), and (H4c) that teams who use the given model will reach agreements with higher team scores than those who do not use a model. For each hypothesis, we tested the significance of our findings using statistical analyses as described further below.

Research Results

Negotiation Duration

We hypothesized that model use and model authorship will impact the duration of the negotiations. As stated above, we hypothesized:

Hypothesis 1a: Teams who use a model (whether given or cocreated) will reach agreements more *quickly* than those who do not use a model.

Hypothesis 1b.1: Teams who cocreate a model will reach agreement more *slowly* than those who use the given model.

Hypothesis 1b.2: Teams who cocreate a model will reach agreement more *slowly* than those who do not use a model.

***Hypothesis 1c*:** Teams who use the given model will reach agreement more *quickly* than those who do not use a model.

Use of the LCA impacted the duration of the negotiation in our sample. Teams using the LCA model, whether using the given model or cocreating their own, had shorter mean negotiation durations than teams not using an LCA model (see Figure 1). Using independent-samples t tests (Welch's unequal variances), we examined the differences in duration in response to model use or model authorship, as hypothesized above. Furthermore, we used F tests to compare the variances of these categories.

In our sample, the teams that used a model had a shorter mean negotiation duration (46.3 min) than teams that did not use a model (51 min), offering only limited support for H1a. The means were not statistically significantly different ($t = -1.079$; $df = 16.26$; $p = 0.2965$), suggesting little support for H1a (see Table 2).

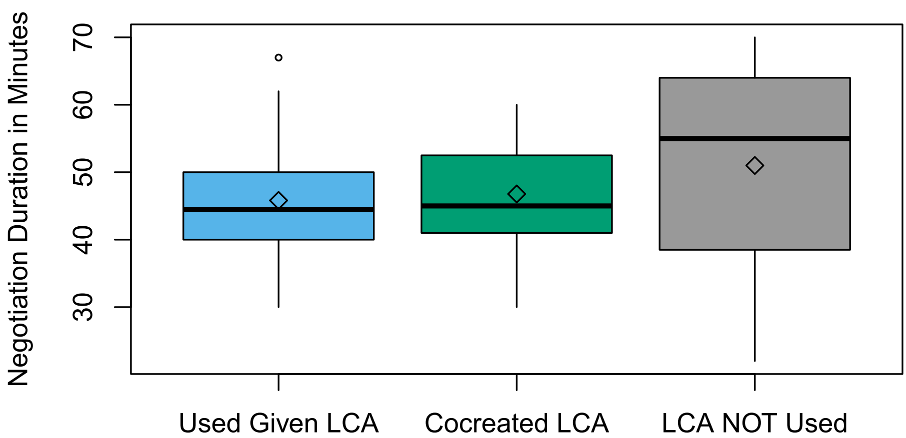

Figure 1. Negotiation Duration. Mean (diamond), median, and 25th and 75th percentiles for the model-use and authorship categories: Used given LCA, cocreated LCA, and LCA not used. The tables not using an LCA had a longer mean and median negotiation duration than the other model-use categories.

Table 2

Results of Testing H1a, H1b.1, H1b.2, and H1c and Test of Variances: Negotiation Durations for Comparisons of Model Use and Authorship Categories

		Comparing means			Comparing variances			
Comparison	Means (in min)	*t* statistic	*df*	*p*-value	*F* test	Numerator *df*	Denominator *df*	*p*-value
H1a: Model used vs. no model used	Model used 46.3 No model 51	−1.079	16.26	0.2965	0.308	58	14	0.0016**
H1b.1: Cocreated vs. used given model	Cocreated 46.8 Used given 45.8	−0.406	56.593	0.6865	1.193	31	26	0.6512
H1b.2: Cocreated vs. no model used	Cocreated 46.8 No model 51	−0.9264	18.5	0.3662	0.283	26	14	0.0053**
H1c: Used given model vs. no model used	Used given 45.8 No model 51	−1.138	18.569	0.2697	0.337	31	14	0.0115**

Note. Significance codes: **** = 0; *** = 0.001; ** = 0.05; * = 0.1.

The teams that cocreated the model had to do so while negotiating (the prenegotiation preparation period was individual). Even with this extra burden of creating a model, these teams in our sample still had shorter mean negotiation times compared to teams not using the model. These teams that cocreated their own model had just slightly longer mean negotiation durations (mean: 46.8 min) compared to teams that used the given model (mean: 45.8 min), offering limited support for H1b.1. The differences are not statistically significant, however. Indeed, we find that whereas we expected the former to take longer negotiating—because they had to recognize the need for a model and develop it, while the latter had only to become familiar with a given model—they actually took about the same amount of time statistically. For H1b.2, in our sample, the teams that cocreated a model had a shorter mean negotiation duration (46.8 min) than the teams that did not use a model (51 min); this is opposite what we expected. The difference in means is not statistically significantly different. Therefore, we find no support for H1b.2. Table 2 delineates these results.

Of the teams that received the model, none of the participants had seen the model before the 20-min negotiation preparation period (the time during which participants read the confidential instructions for their roles and the general instructions for all roles). Therefore, there was a learning curve associated with the LCA model itself. Despite this, as we expected, teams using the given model had a shorter negotiation time (mean: 45.8 min) than teams not using a model (mean: 51 min) in our sample, offering limited support for H1c. The difference in means is not significant in Welch's unequal variances t test. Therefore, there is little evidence supporting H1c, suggesting that the time efficiencies gained from using the model—better understanding of the available alternatives, the reduction in the number of preference-based issues, and the social learning from model use—were offset by the addition of the learning curve, resulting in a statistically equivalent negotiation time for negotiation teams using the given LCA model and those not using the LCA model. See Table 2 for statistical results.

Despite the mean durations being appreciably equal regardless of model use and model authorship, we did observe significant differences in their variances. The teams using an LCA had a smaller variance in their negotiation times. When teams that used the given LCA and those that cocreated an LCA are considered as a composite category (teams that used a model), the variances for the teams using a model and those not using a model significantly differ in a two-sided F test ($F = 0.30804$, numerator $df = 58$, denominator $df = 14$, $p = 0.0016$). Considering model authorship, the variances of the teams that cocreated a model and those that used the given model are not significantly different from each other. However, the variance of each is statistically different from the variance of teams that did not use a model, in a two-sided F test: Teams that cocreated their own model and those that did not use a model have a significant difference in variance in a two-sided F test ($F = 0.28289$, numerator $df = 26$, denominator $df = 14$, $p = 0.0053$), and the teams that used the given model and those that did not use a model have a significantly different variance ($F = 0.33744$, numerator $df = 31$, denominator $df = 14$, $p = 0.0115$). Therefore, although mean durations of negotiation do not vary in statistically significant patterns, the variances are significantly different in H1a, H1b.2, and H1c.

Using the LCA Model for Agreement Building vs. Agreement Verification vs. Personal Gain

As stated above, we hypothesized:

> ***Hypothesis 2*:** Model-using teams will use the model in one of three distinct manners: collectively to test alternatives while building an agreement, collectively to verify a tentatively accepted agreement, and individually to benefit personally (by the party introducing the model).

Survey, interview, and observational data revealed how the teams used the model. We observed all three outcomes, but in most teams, the parties used the model collectively. While the cup maker (the party who received the model) could use the model to maximize his or her own utility by collecting data from the other parties, plugging these data into the LCA model, and not sharing the results, this occurred

Table 3
Ways to Use the Model. The Predominant Two Ways Were to Test Alternatives or to Verify Agreements, Although a Third Manner (Self-Benefit of the Party Introducing the Model) Was Also Observed

Verify agreement	"We looked at it to make sure we had the right decisions to trigger the incentive [bonus for reaching the CO_2 savings goal]."
Test alternatives	"His [cup maker's] spreadsheet kinda became our fact checking or like a litmus test tool. . . . [W]hen we started looking at different proposals, we started looking at the Cup Maker's spreadsheet to see [which were] above the threshold."
Self-benefit	"He [cup maker] just went around and asked us for numbers to estimate the CO_2 per ton that we would reduce. . .. So we thought maybe there was some sort of spreadsheet he was working with already. We don't know if he was calculating it separately or not. . .."

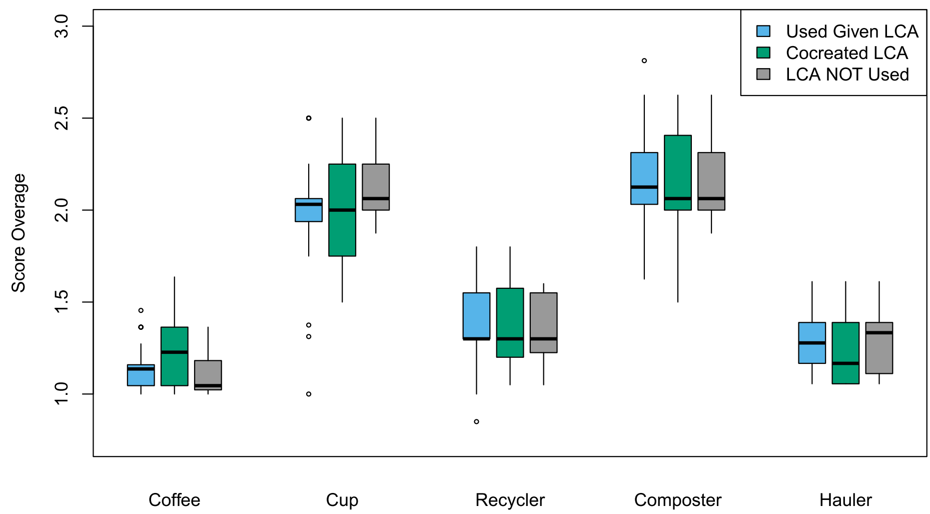

Figure 2. Role score overages by used given LCA, cocreated LCA, and LCA not used. A score overage value of 1.0 is the point threshold to join the agreement. Model use did not substantively change the value distribution among the roles.

in only a few teams. Negotiators that collectively used the model used it either to test the emissions of alternatives during the process of developing an agreement or to verify the emissions of an agreement tentatively agreed upon. Table 3 provides interview quotes representative of each way of using the LCA.

More teams who used the LCA to test alternatives as they built agreements achieved high team scores than did teams who used the LCA to verify an agreement; that is, the 75th percentiles were higher in the test alternatives group than in the verify agreement group despite their similar 50th percentiles. However, these differences are not statistically significant. Of the teams for whom we could document how they used the model (26), the teams who used the model to test alternatives were the only teams to reach the value maximizing agreement.

To investigate whether some parties are gaining advantage over other parties through model use, we investigate the value distribution among parties. We measure this by looking at the score overage, which is the normalized number of points (utility)—for each participant, the points he/she earned divided by the minimum number of points his/her role must attain to enter into an agreement. The thresholds for joining an agreement differ by role. Small differences in the score overages exist within the roles depending on whether the team used the given model, cocreated a model, or did not use a model (for medians and interquartile ranges, see Figure 2); note that the team that did not reach an agreement is included in the data generating Figure 2, though the scale of the axis does not display this team's 0 points. Note that the variances are not expressly visible in Figure 2. The coffee retailer had the largest median score overage for teams that cocreated a model and the largest variance for teams that did not use a model. The cup maker's (the role that was given the model and could therefore have used it secretly) median score overage was comparable for teams using the given model and teams not using a model, and slightly lower when the team cocreated a model. The variance of the cup maker's score overage was the largest for teams that did not use a model and smallest for teams that cocreated a model. The recycler had equal medians among the three cases and the largest variance for teams that did not use a model. Note that the recycler outlier with a normalized score under 1.0 is the individual who accepted an agreement he or she should not have, as discussed below. The composter had the largest median for teams that used the given model and the largest variance for teams that did not use a model. The hauler had the highest median overage and largest variance on teams that did not use a model. When considering median score overages, the coffee retailer and the hauler were most impacted by model use and authorship; when considering variance of the score overages, the cup maker and the composter were the most impacted.

Favorable Agreements: Most Environmental Protection or Value Maximum

Here, we examine favorable agreements—the subset of feasible agreements that either provide the most environmental protection or, mutually exclusively, maximize the value earned. The former is a set of twelve agreements that minimize carbon dioxide emissions, and the latter is one agreement that maximizes agreement value (also called team score). The twelve environmental maximizing agreements each have a lower team score than the one score maximizing agreement, and the score maximizing agreement offers less environmental protection than the environmental maximizing agreements.

As stated above, we hypothesized:

Hypothesis 3a**:** The percentage of teams who reach a favorable agreement will be higher for teams who use a model (whether given or cocreated) than for teams who do not use a model.

Hypothesis 3b.1**:** The percentage of teams who reach a favorable agreement will be higher for teams who cocreate a model than for teams who use the given model.

Hypothesis 3b.2**:** The percentage of teams who reach a favorable agreement will be higher for teams who cocreate a model than for teams who do not use a model.

Hypothesis 3c**:** The percentage of teams who reach a favorable agreement will be higher for teams who use the given model than for teams who do not use a model.

The LCA model enabled the negotiators to calculate the amount of carbon dioxide their agreement would save and thus relates to two of the five issues being negotiated. That is, the LCA model provides information to address the physical world issues that teams not using the LCA model had to address with preference-based information. In our sample, the percentage of teams reaching a favorable agreement—the number of teams reaching a favorable agreement divided by the number of teams in that category—for teams using a model (34%) was higher than for teams not using a model (27%), offering limited support for H3a. However, Fisher's exact test, a one-sided comparison of these, shows that they are not statistically significant (odds ratio = 1.40, p-value = 0.4196).

Teams that cocreated their own LCA outperformed the other two categories in reaching a favorable agreement in our sample (limited support for H3b.1 and H3b.2). The percentage of teams reaching a favorable agreement for teams that cocreated a model (44%) and for those that used the given model (25%) are significantly different (odds ratio = 2.36, p = 0.0975) from each other in a one-sided, 90% Fisher's exact test (chi-squared test of equal proportions with Yates continuity correction gives similar results). This further and more strongly supports H3b.1. While the percentage of teams reaching a

Table 4
Results of Testing H3a, H3b.1, and H3b.2, and H3c: Ratio of Favorable Agreements for Comparisons of Model Use and Model Authorship Categories

Comparison	Ratio of favorable agreements	Odds ratio	*p*-value
H3a: Model used vs. no model used	Model used 0.34 No model 0.27	1.40	0.4196
H3b.1: Cocreated vs. used given model	Cocreated 0.44 Used given 0.25	2.36	0.0975*
H3b.2: Cocreated vs. no model used	Cocreated 0.44 No model 0.27	2.16	0.2116
H3c: Used Given vs. no model used	Used given 0.25 No model 0.27	0.918	0.6900

Note. Significance codes: **** = 0; *** = 0.001; ** = 0.05; * = 0.1

favorable agreement for teams that cocreated a model (44%) was larger than for those that did not use a model (27%), the difference is not statistically significant in a one-sided Fisher's exact test (odds ratio = 2.16, $p = 0.2116$). Therefore, we find little support for H3b.2. We do not find support for H3c: the percentages of teams who reached a favorable agreement are statistically similar (in a one-sided Fisher's Exact test) for teams that used the given model and teams that did not use a model (25% for the former; 27% for the latter). See Table 4.

In addition, two other findings should be highlighted. First, half (50%) of the cocreating teams in our sample that reached a favorable agreement chose an environmentally favorable agreement (over the point maximizing), compared to only 37.5% of the teams using the given model and 25% of the teams not using a model. Second, our data show that half (50%) of the teams that did not use a model reached a unique final agreement (a final agreement not reached by any other team in that category), compared to 38% of teams that used the given model and 41% of teams that cocreated their own model. That is, most model-using teams reached agreements within the same small subset of agreements. This concentration of their negotiated agreements lends support to the finding that cocreating teams were better able to target their final agreement to a favorable agreement.

Agreement Value

The team score of the agreement is the sum of the parties' individual scores; it is also called the agreement value. As discussed above, in the Cup Game design, the agreements that maximize the emissions reduction do not have the highest team values (these team scores range from 590 to 695 points), and the value maximizing agreement (765 points) offers less emissions reduction. When choosing among the set of agreements that maximize emissions reduction, the teams would benefit from choosing an agreement with a high score. Furthermore, the value maximizing agreement forms a ceiling for the team score data (765) and not reaching an agreement forms the floor (0). In this section, we consider the team score regardless of emissions savings.

As stated above, we hypothesized:

***Hypothesis 4a*:** Teams who use a model (whether given or cocreated) will reach agreements with *higher* team scores than those who do not use a model.

Hypothesis 4b.1: Teams who cocreate a model will reach agreements with higher team scores than those who use the given model.

Hypothesis 4b.2: Teams who cocreate models will reach agreements with higher team scores than those who do not use a model.

Hypothesis 4c: Teams who use the given model will reach agreements with higher team scores than those who do not use a model.

Figure 3 shows the scores of teams who used the given LCA model, who cocreated their own LCA, and who did not use an LCA model. In this figure, the lowest visible outlier of the teams who used the given LCA is an infeasible agreement. In this team, one party either miscalculated his or her score or knowingly, possibly reluctantly, agreed to a score lower than his or her threshold. The only team not to reach agreement did not use a model; this team is included in the data generating Figure 3, though the scale of the axis does not display this team's 0 points. Considering model use (using a model versus not using a model), the mean score of teams using a model is 717 points compared to 682 points for teams not using a model, offering only limited support for H4a. The difference in means is not significant ($W = 424.5$, $p = 0.402$) in a one-sided asymptotic Wilcoxon rank sum test. Therefore, we find little support for H4a. The difference in variance between these two categories is significant ($F = 0.049091$, numerator $df = 58$, denominator $df = 14$, $p = 2.2 \times 10^{-16}$) in an F test to compare the variances. Note that the variances are not directly visible in Figure 3.

The teams cocreating a model had the highest mean score (719 points), followed by teams using the given model (716 points), offering limited support for H4b.1, but the difference in these means is not significant ($W = 399$, $p = 0.3057$) in a one-sided exact Wilcoxon rank sum test. Furthermore, for H4b.2, teams cocreating a model earned more points (719 points) than teams not using a model (682 points), but again the difference in means is not significant in a one-sided exact Wilcoxon rank sum test. The variance for these latter two categories is significantly different ($F = 0.0514$, numerator $df = 26$, denominator $df = 14$, $p = 5.5 \times 10^{-10}$) in an F test.

For H4c, the difference in mean team scores between teams using the given model (mean: 716 points) and teams not using a model (mean: 682 points) is not significant in a one-sided exact Wilcoxon rank sum test ($W = 223.5$, $p = 0.6572$), but the difference in variance is significant (F test: $F = 0.0486$, numerator $df = 31$, denominator $df = 14$, $p = 1.5 \times 10^{-11}$). Therefore, we find limited but statistically insignificant support for H4a, H4b.1, H4b.2, and H4c; we find strong evidence that the variances are significantly different in H4a, H4b.2, and H4c (see Table 5).

Table 5

Results of Testing H4a, H4b.1, H4b.2, and H4c and Test of Variances: Team Scores for Comparisons of Model Use and Model Authorship Categories

		Comparing means		Comparing variances			
Comparison	Means (in points)	*W* statistic	*p*-value	*F* statistic	Numerator *df*	Denominator *df*	*p*-value
H4a: Model used vs. no model used	Model used 717 No model 682	424.5	0.4015	0.049	58	14	2.2×10^{-16}****
H4b.1: Cocreated vs. used given model	Cocreated 719 Used given 716	399	0.3057	0.9459	31	26	0.8746
H4b.2: Cocreated vs. no model used	Cocreated 719 No model 682	201	0.5183	0.0514	26	14	5.5×10^{-10}****
H4c: Used given vs. no model used	Used given 716 No model 682	223.5	0.6572	0.0486	31	14	1.5×10^{-11}****

Note. Significance codes: **** = 0; *** = 0.001; ** = 0.05; * = 0.1

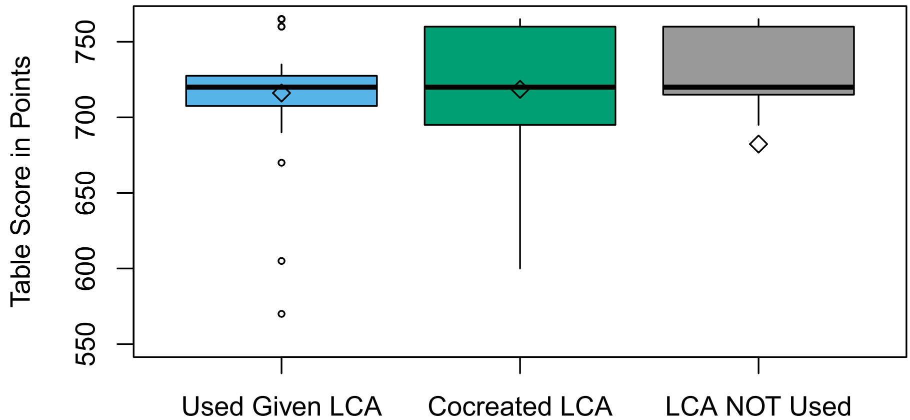

Figure 3. Agreement value. Mean (diamond), median, and 25th and 75th percentiles for the model-use and authorship categories: used given LCA, cocreated LCA, and LCA not used. The model-use and authorship categories all had the same median value (720). Teams cocreating their own LCA had the highest mean (719). The highest possible team score is 765 points.

Alternative Explanations

To verify these results, we tested three alternate explanations: comfort with quantitative information; assessment of trust among the parties; and assessment of role commitment to sustainability. First, we tested whether participants self-selected into model use by nature of their comfort with quantitative information. The participants' self-reported quantitative comfort is not significantly different between teams given a model and those not given a model (Welch two-sample t test: $t = -1.1736$, $df = 276.94$, $p = 0.2416$); it is also not significantly different for teams that used a model versus those that did not use a model (Welch two-sample t test: $t = -0.091766$, $df = 83.185$, $p = 0.9271$). Thus, quantitative comfort is not an alternative explanation for the results found.

Additionally, we assessed whether the participants' trust for the parties differed between teams given a model versus those not given a model, or teams that used a model versus those that did not. Similarly, we assessed whether the participants' perception of the parties' commitment to sustainability differed by these categories. The differences in participants' assessments are not significant—with the exception of other parties' trust for the composter—for the teams given versus not given a model nor for teams that

Table 6
Participants' Ratings of Trust: t Statistics, df, and p-Values of Participant Ratings of Their Trust of the Information Shared by Each Role

	Given vs. not given			Used vs. not used		
	t	*df*	*p*	*t*	*df*	*p*
Trust of info coffee shared	1.3799	193.69	0.1692	−0.1757	24.99	0.8619
Trust of info cup shared	1.3884	189.52	0.1667	−0.7150	23.54	0.4817
Trust of info recycler shared	0.8993	185.4	0.3696	−0.2658	24.73	0.7926
Trust of info composter shared	1.6915	190.61	0.0924*	1.1922	25.26	0.2443
Trust of info hauler shared	1.1006	177.43	0.2726	−0.2469	25.05	0.807

Notes. None of the values are significant at better than 90% level, with the exception of trust for the composter in given versus not given, which is significant at a 90% level.
Significance codes: **** = 0; *** = 0.001; ** = 0.05; * = 0.1

Table 7
Participants' Ratings of Roles' Commitment to Sustainability: t Statistics, df, and p-Values of Participant Ratings of Role Commitment to Sustainability

	Given vs. not given			Used vs. not used		
	t	*df*	*p*	*t*	*df*	*p*
Coffee commitment to sustainability	1.2207	284.63	0.2232	−0.0975	70.33	0.9226
Cup commitment to sustainability	0.51783	281.31	0.605	−0.4013	79.25	0.6893
Recycler commitment to sustainability	1.5735	284.71	0.1167	−1.518	65.00	0.134
Composter commitment to sustainability	0.1806	278.03	0.8568	−0.6923	73.62	0.4909
Hauler commitment to sustainability	−0.3426	282.26	0.7322	0.7663	75.92	0.4459

Notes. None of the values are significant at better than 90% level.
Significance codes: **** = 0; *** = 0.001; ** = 0.05; * = 0.1

used versus those that did not use a model (see Table 6 for *t* statistics, *df*, and *p*-values by role for trust ratings and Table 7 for a listing of *t* statistics, *df*, and *p*-values by role for ratings of commitment to sustainability). Tables 6 and 7 use Welch two-sample *t* tests. Neither trust nor perceived commitment to sustainability explains the results observed in this study.

Discussion and Conclusions

This study complements existing studies of collaborative model use (such as those by Videira, Antunes, Santos, & Gamito, 2003; Videira, van den Belt, Antunes, Santos, & Gamito, 2004; and Beall & Zeoli, 2008) by utilizing a method that allows for repeated instances within the same context. Here, we used a serious game role-play simulation wherein participants enact one of five roles negotiating together. Approximately half of our 74 teams of five negotiators received a model. These repeated instances enable us to provide quantitative information about how model use impacts the negotiation outcome and process, and the ways models are used in sustainability negotiations.

This study was contextualized within a sustainability negotiation that contains a mixture of issues involving negotiator preference and scientific information. We chose the particular case of recycling and/or composting used paper coffee cups because it was ongoing in the real world, involved many parties all wanting a sustainable outcome, and required that the parties address both preference-based and scientific issues. In addition, we had access to real-world stakeholders while developing the role-play simulation. We anticipate that models could be helpful in other types of negotiations, beyond sustainability, that also require negotiators to grapple with both preferences and scientific information.

By allowing participants the freedom to use or not use an expert-provided model, to cocreate a model, or to negotiate without a model, we were able to study the participants' inclination toward using a model in the negotiation and to compare the use of an expert-given model with that of a cocreated model. Through surveys issued before and after the role-play simulation and interviews a few weeks later, we gained more insight into the respective impacts of model use, model authorship, and the ways negotiators use models while negotiating.

In the role-play simulation, more teams (80%) used a model than did not, even if they had to cocreate the model themselves. This suggests that the negotiators found a model helpful in their negotiation process and confirms the collaborative modeling research findings that model use can help sustainability decision makers (see, for example, Tidwell & van den Brink, 2008). We quantify the potential magnitude of this effect.

We found quantitative evidence that model use did not prolong the negotiation process. In fact, we found some evidence that model use might speed up the negotiation: in our sample, on average, teams that used a model reached agreement faster than those that did not use a model. Even those teams that had to cocreate a model while negotiating negotiated faster than teams that did not use a model. However, our analysis on negotiation duration shows that, statistically, the time it takes to negotiate using a model compared to the time it takes to negotiate without a model is approximately the same. This is important because a frequently stated reason for not using a negotiated decision making approach is concern for how long it takes. Parties already concerned with a prolonged process might be unwilling to add a component—such as a model—that might lengthen the process. Our results indicate that using a model does not lengthen the negotiation, and in our sample of 74 teams of negotiators, model use sped up the negotiation. Furthermore, we found that model use reduced the variance in negotiation time significantly compared to not using a model; this could potentially allow negotiators to better plan their negotiation process, with a more accurate estimate of how long it is likely to take. Quicker negotiations also increase the likelihood that the negotiation will reach a conclusion rather than being aborted due to a lengthy process. Situations where some parties are better off if no agreement is reached notwithstanding, it is generally the case that an agreement must be reached for the group to take action toward improving sustainability. Therefore, by not lengthening the negotiation duration—and potentially shortening it and thereby increasing the likelihood of a negotiated agreement—using a model in sustainability negotiation can improve the chances of collective sustainability action.

This study demonstrates that using models to test alternatives while agreement building, rather than to verify a tentative agreement, can lead to higher team scores more frequently. By exploring as they negotiate, the negotiators can see more clearly the impact of each choice. Seeing the environmental impact of alternatives considered can enable negotiators to create agreements that benefit the environment. Additionally, seeing the impact while negotiating allows the negotiators to identify which parties lose value on which alternatives and to find potential areas to give more value to those parties. Finding ways to convey more value is particularly pertinent if some parties lose value due to choices that protect the environment. Rather than stopping the entire negotiation process, parties that lose on environmental issues can gain enough on other issues to join the agreement, thus enabling collective environmental action. Future studies could investigate these mechanisms directly.

Furthermore, the party with ownership of the expert-provided model—the cup maker—could have used the model secretly for personal gain (rather than sharing it with the other negotiators in his or her team). If he or she were using it for individual gain, we would expect the cup maker's score overage to be highest in teams that used the given model. However, there was no appreciable difference in the cup maker's score overage for teams using the given model and teams not using a model. In teams that used the given model, the coffee retailer (convener of the negotiation) and, to a lesser extent, the cup maker (owner of the model) were able to target agreements more precisely than the other roles in their team and than those two roles in teams that cocreated a model or did not use a model.

Our study indicates that not only did model use help the negotiation process, but it also improved the negotiated outcome. The model-using teams in our sample had a higher average team score and targeted their agreements. Additionally, we found evidence that involving negotiators in model creation can improve the negotiated outcome: Negotiators who cocreated a model reached the highest ratio of favorable agreements (agreements that either maximize the environmental protection or maximize the negotiated value). These teams also had the highest mean team score in our sample, although statistically no different than other teams. Therefore, cocreating a model offers the possibility of targeting more favorable agreements and could result in a team score as good as, if not better than, using a given model. The strong performance of the teams cocreating their model supports the suggestion that researchers such as van den Belt et al. (2013) have made to involve the parties in model development.

In addition to having the highest mean team score, negotiators cocreating a model had the second largest variance. The large variance could potentially be related to the differences (including errors) in the models they created. Contrasting with the variance of the teams that used the given model (the smallest variance of the three categories) supports this reasoning because these teams used the same model as each other. Future research could test the relationship between model correctness and score variance by giving some teams a correct model and other teams a model with an error in it. Teams that used the given model had the smallest variance, indicating that they were able to consistently earn value in the negotiation.

It is possible that there are other explanations besides model use and model authorship for the results observed here. We measured several variables to help rule out some of these possible explanations. Our evidence suggests that comfort with quantitative information, trust among the parties, and perceived commitment to sustainability did not determine whether negotiators used a model or not and were not aligned with whether they received a model or not.

While this serious gaming approach allows us to test repeated instances of the same sustainability context, the role-play simulation constraints required that the physical world scenario be simplified. Further research could explore the influence a more complex model has on negotiation duration, manner of model usage, deliberate agreement choices, and team score. Ideally, the model complexity should be on the level of those in case study and action research-based collaborative modeling studies (Beall & Zeoli, 2008; van den Belt et al., 2013; Videira, Antunes, Santos, & Gamito, 2003; Videira, van den Belt, Antunes, Santos, & Gamito, 2004). Moreover, future research could investigate further the different manners of model use. Such a study could also test whether model authorship reinforces the manner of model use in impacting the negotiation process and negotiated outcomes. Additionally, future studies could investigate why some teams of negotiators do not use a model, including looking at why some ignore a given model and others do not cocreate a model. Future research can also investigate why parties that cocreate a model while negotiating took no longer to negotiate, even though they split their time between negotiating and cocreating the model.

Model use can help to explore the implications of various alternatives (Dowlatabadi, 1995; Morgan, 2011; and van Delden et al., 2011) within the negotiation. Furthermore, it provides a structure to the dialogue (Tidwell & van den Brink, 2008; van den Belt, 2006). Exploring the various alternatives in a structured fashion can help teams explore the solution space of the negotiation—the zone of potential agreement—and identify the most favorable agreements. Future research can determine whether model use impacts the agreements considered en route to the final agreement. In this study, we did not collect data about the agreements considered by each team (only data about the agreement finally reached); therefore, we cannot address the teams' trajectory through the solution space. However, our findings do show that model-using teams reached a smaller set of final agreements and that teams cocreating their model reached more favorable agreements, which together indicate that model-cocreating teams were better able to target their agreements.

Based on the results of this study, we find qualified support that using a model in sustainability negotiations might help negotiators achieve more overall value more consistently and reach agreements that are more protective of the environment, without lengthening the negotiation. Along with other researchers (Langsdale et al., 2013; van den Belt, 2004; van den Belt et al., 2013; Videira, Antunes, Santos, & Gamito, 2003; Videira, van den Belt, Antunes, Santos, & Gamito, 2004; among others), we further suggest that the way models are used and who creates them might be more important than just having a model. Teams cocreating their own model reached the highest ratio of favorable agreements (value maximizing or maximizing the environmental protection). Teams that used the model to test alternatives as they were negotiating reached higher scores than those using the model to verify a tentative agreement. In aggregate, the results of this study suggest the importance of model use—especially models collaboratively built by the negotiators—in sustainability negotiations.

References

Beall, A., & Zeoli, L. (2008). Participatory modeling of endangered wildlife systems: Simulating the sage-grouse and land use in central Washington. *Ecological Economics*, *68*, 24–33. doi:10.1016/j.ecolecon.2008.08.019

Butler, J.K. (1991). Toward understanding and measuring conditions of trust: Evolution of a conditions of trust inventory. *Journal of Management*, *17*, 643–663. doi:10.1177/014920639101700307

Corrigan, S., Zon, G.D.R., Maij, A., McDonald, N., & Mårtensson, L. (2015). An approach to collaborative learning and the serious game development. *Cognition, Technology & Work*, *17*, 269–278. doi:10.1007/s10111-014-0289-8

Curhan, J.R., Elfenbein, H.A., & Kilduff, G.J. (2009). Getting off on the right foot: Subjective value versus economic value in predicting longitudinal job outcomes from job offer negotiations. *Journal of Applied Psychology*, *94*, 524–534. doi:10.1037/a0013746

Curhan, J.R., Neale, M.A., & Ross, L. (2004). Dynamic valuation: Preference changes in the context of face-to-face negotiation. *Journal of Experimental Social Psychology*, *40*, 142–151. doi:10.1016/j.jesp.2003.12.002

Curhan, J.R., & Pentland, A. (2007). Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes. *Journal of Applied Psychology*, *92*, 802–811. doi:10.1037/0021-9010.92.3.802

Czaika, E. (2010). *Starbucks cups: Trash or treasure? An example of facilitated systems thinking assisting stakeholders in designing their own system to recycle take-away cups*. Cambridge, MA: Massachusetts Institute of Technology.

De Dreu, C.K.W., Beersma, B., Stroebe, K., & Euwema, M.C. (2006). Motivated information processing, strategic choice, and the quality of negotiated agreement. *Journal of Personality and Social Psychology*, *90*, 927–943. doi:10.1037/0022-3514.90.6.927.

Dowlatabadi, H. (1995). Integrated assessment models of climate change: An incomplete overview. *Energy Policy*, *23*, 289–296. doi:10.1016/0301-4215(95)90155-Z

Ducrot, R., van Paassen, A., Barban, V., Daré, W., & Gramaglia, C. (2015). Learning integrative negotiation to manage complex environmental issues: Example of a gaming approach in the peri-urban catchment of São Paulo, Brazil. *Regional Environmental Change*, *15*, 67–78. doi:10.1007/s10113-014-0612-1

Elfenbein, H.A., Curhan, J.R., Eisenkraft, N., Shirako, A., & Baccaro, L. (2008). Are some negotiators better than others? Individual differences in bargaining outcomes. *Journal of Research in Personality*, *42*, 1463–1475. doi:10.1016/j.jrp.2008.06.010.

Funtowicz, S.O., & Ravetz, J.R. (1993). Science for the post-normal age. *Futures*, *25*, 739–755. doi:10.1016/0016-3287(93)90022-L

Funtowicz, S.O., & Ravetz, J.R. (1994). The worth of a songbird: Ecological economics as a post-normal science. *Ecological Economics*, *10*, 197–207. doi:10.1016/0921-8009(94)90108-2

Herbst, U., & Schwarz, S. (2011). How valid is negotiation research based on student sample groups? New insights into a long-standing controversy. *Negotiation Journal*, *27*, 147–170. doi:10.1111/j.1571-9979.2011.00300.x

Klabbers, J.H.G. (2005). Enhancing policy development through actor-based simulation. In R. Shiratori, K. Arai & F. Kato (Eds.), *Gaming, simulations, and society: Research scope and perspective* (pp. 1–10). Tokyo: Springer.

Kuit, M., Mayer, I.S., & de Jong, M. (2005). The INFRASTRATEGO game: An evaluation of strategic behavior and regulatory regimes in a liberalizing electricity market. *Simulation & Gaming*, *36*, 58–74. doi:10.1177/1046878104272666

Langsdale, S., Beall, A., Bourget, E., Hagen, E., Kudlas, S., Palmer, R., et al. (2013). Collaborative modeling for decision support in water resources: Principles and best practices. *Journal of the American Water Resources Association*, *49*, 629–638. doi:10.1111/jawr.12065

Langsdale, S.M., Beall, A., Carmichael, J., Cohen, S.J., Forster, C.B., & Neale, T. (2009). Exploring the implications of climate change on water resources through participatory modeling: Case study of the Okanagan Basin, British Columbia. *Journal of Water Resources Planning and Management*, *135*, 373–381. doi:10.1061/(ASCE)0733-9496(2009)135:5(373)

MaKinster, J. (2010). Unraveling the scientific, social, political, and economic dimensions of environmental issues through role-playing simulations. In A. M. Bodzin, B. Shiner Klein, & S. Weaver (Eds.), *The inclusion of environmental education in science teacher education* (pp. 237–253). Netherlands: Springer.

Mayer, I.S. (2009). The gaming of policy and the politics of gaming: A review. *Simulation & Gaming*, *40*, 825–862. doi:10.1177/1046878109346456

Mayer, I.S., Carton, L., de Jong, M., Leijten, M., & Dammers, E. (2004). Gaming the future of an urban network. *Futures*, *36*, 311–333. doi:10.1016/S0016-3287(03)00159-9

Mayer, I., Meijer, S., Nefs, M., Gerretsen, P., & Dooghe, D. (2010). Gaming the interrelation between rail infrastructure and station area development: Part 2 – Insights from the serious game "SprintCity". Presented at the

2010 *Third International Conference on Infrastructure Systems and Services: Next Generation Infrastructure Systems for Eco-Cities (INFRA), IEEE*, pp. 1–6. doi:10.1109/INFRA.2010.5679218

McIntosh, B.S., Ascough, J.C., Twery, M., Chew, J., Elmahdi, A., Haase, D., et al. (2011). Environmental decision support systems (EDSS) development – Challenges and best practices. *Environmental Modelling and Software*, *26*, 1389–1402. doi:10.1016/j.envsoft.2011.09.009

Meijer, S.A., Mayer, I.S., van Luipen, J., & Weitenberg, N. (2012). Gaming rail cargo management: Exploring and validating alternative modes of organization. *Simulation & Gaming*, *43*, 85–101. doi:10.1177/1046878110382161.

Morgan, M. G. (2011). Technically focused policy analysis. In K. Husbands Fealing, J. I. Lane, J. H. Marburger III, & S.S. Shipp (Eds.), *The science of science policy, a handbook* (pp. 120–130). Stanford, CA: Stanford University Press.

Mustajoki, J., Marttunen, M., Karjalainen, T.P., Hokkanen, J., Vehmas, A. (2013). "Developing practices for supporting EIA with multi-criteria decision analysis. Proceedings of the 33rd Annual Meeting of the International Association for Impact Assessment (IAIA13), Calgary, Canada, May 13–16, 2013, pp. 5, (http://conferences.iaia.org/2013/final-papers.html)

Nefs, M., Gerretsen, P., Dooghe, D., Mayer, I. S., & Meijer, S. (2010). Gaming the interrelation between rail infrastructure and station area development: Part 1 – Modeling the serious game "SprintCity". Presented at the 2010 *Third International Conference on Infrastructure Systems and Services: Next Generation Infrastructure Systems for Eco-Cities (INFRA), IEEE*, pp. 1–6. doi:10.1109/INFRA.2010.5679221

Schenk, T., & Susskind, L. (2015). Using role-play simulations to encourage adaptation: Serious games as tools for action research. In A. van Buuren, J. Eshuis, & M. van Vliet (Eds.), *Action research for climate change adaptation: Developing and applying knowledge for governance* (pp. 148–165). New York City: Routledge.

Stokes, L.C., & Selin, N.E. (2014). The mercury game: Evaluating a negotiation simulation that teaches students about science-policy interactions. *Journal of Environmental Studies and Sciences*, pp. 1–9. doi:10.1007/s13412-014-0183-y

Team. (2016). In the Oxford Dictionaries. Retrieved from http://www.oxforddictionaries.com

Tidwell, V.C., & van den Brink, C. (2008). Cooperative modeling: Linking science, communication, and ground water planning. *Ground Water*, *46*, 174–182. doi:10.1111/j.1745-6584.2007.00394.x.

van Delden, H., Seppelt, R., White, R., & Jakeman, A.J. (2011). A methodology for the design and development of integrated models for policy support. *Environmental Modelling and Software*, *26*, 266–279. doi:10.1016/j.envsoft.2010.03.021

van den Belt, M. (2004). *Mediated modeling: A system dynamics approach to environmental consensus building*. Washington, DC: Island Press.

van den Belt, M. (2006). Mediated modeling of the impacts of enhanced uv-b radiation on ecosystem services. *Photochemistry and Photobiology*, *82*, 865–877. doi:10.1562/2005-10-19-1R-722

van den Belt, M., Schiele, H., & Forgie, V. (2013). Integrated freshwater solutions – A New Zealand application of mediated modeling. *Journal of the American Water Resources Association*, *49*, 669–680. doi:10.1111/jawr.12064

Videira, N., Antunes, P., Santos, R., & Gamito, S. (2003). Participatory modeling in environmental decision-making: The Ria Formosa Natural Park case study. *Journal of Environmental Assessment Policy and Management*, *5*, 421–447. doi:10.1142/S1464333203001371.

Videira, N., van den Belt, M., Antunes, P., Santos, R., & Gamito, S. (2004). Decision support in coastal zone management in the Ria Formosa, Portugal. In Marjan van den Belt (Ed.), *Mediated modeling: A system dynamics approach to environmental consensus building* (pp. 205–229). Washington, DC: Island Press.

Ellen Czaika received her Ph.D. from the Massachusetts Institute of Technology (MIT) in Engineering Systems: Technology, Management, and Policy as part of the Engineering Systems Division (now the Institute for Data Systems and Society). She researches data-driven decision making in sustainability contexts and is particularly interested in how mathematical models can be used in sustainability negotiations and policy making. She has over 9 years of experience in the private sector as a management

consultant, systems engineer, and engineering manager. She also has a bachelor of science in mathematics from Millsaps College and masters of science degrees in Applied Statistics from the University of Oxford and in Engineering and Management (with the System Design and Management Program) from MIT.

Noelle E. Selin is Esther and Harold E. Edgerton Career Development Associate Professor at the Institute for Data, Systems, and Society and the Department of Earth, Atmospheric and Planetary Sciences at the Massachusetts Institute of Technology. She also serves as Associate Director of MIT's Technology and Policy Program. Her research focuses on using quantitative modeling and analysis to inform decision-making strategies on air quality and sustainability issues. She has also published articles and book chapters on the interactions between science and policy in international environmental negotiations, in particular focusing on global efforts to regulate hazardous chemicals and persistent organic pollutants. She has a Ph.D. (Earth and Planetary Sciences), M.A. (Earth and Planetary Sciences), and B.A. (Environmental Science and Public Policy) from Harvard University.