Lustick and Tetlock (2021) argue for the use of theory-guided simulation for aiding geopolitical planning and decision-making. This is a very welcome contribution, as in our experience, geopolitical analyses are often qualitative in nature and prone to group think. The complexities of geopolitical issues, however, really call for the use of theory-guided simulations. These issues are too complex for mental simulation (Atkins et al., 2002; Sterman, 1989, 1994). Dynamic simulations, if properly grounded in appropriate theories and well-motivated assumptions, can derive the possible dynamics from interacting nonlinear processes, and thus aid human reasoning about system behavior (Sterman, 2002). Moreover, since geopolitical issues are subject to uncertainty, scarce data, and conflicting information, using an ensemble modeling approach is appropriate. An ensemble of simulations enables reasoning across alternative assumptions consistent with the available data and information. Such an ensemble can capture much more of the available theories, information, and educated guesses than any single model in isolation (Bankes, 2002). With the rising computational power, ensemble modeling is increasingly a feasible research strategy.

Despite our broad agreement with Lustick and Tetlock (2021), we have three major comments on their work. First, from the broader perspective of modeling and simulation, they offer little that is truly novel or surprising. The envisioned approach of ensemble simulations is already well established under the label of exploratory modeling. By not engaging with this literature, the authors have deprived themselves from a rich set of analytical techniques that could have substantially strengthened their case study, as well as relevant theories and concepts which would have strengthened the appeal of the manifesto. Second, we content that validating simulation models of complex systems with partially open system boundaries should focus on perceived usefulness, not supposedly predictive accuracy captured through brier scores. Third, in interpreting results from simulation models it is possible and useful to try and increase understanding between the system’s structural characteristics and theories, instead of using simulations as point predictions.

1 | EXPLORATORY MODELING IS ALREADY WELL ESTABLISHED

There is ample literature that has emerged over the last 30 years on the use of computational experimentation with simulation models to aid planning and decision-making. Hodges (1991) identified six things that could be done with simulation models in the absence of good data. Hodges and Dewar (1992) identified a seventh use case. Bankes (1993) moved away from enumerating the number of use cases, and simply spoke of exploratory modeling. Since these formative ideas from the early 1990s, a large body of literature on exploratory modeling has emerged (see, e.g., Kwakkel, 2017; Moallemi et al., 2020 for detailed overviews). Perhaps most importantly, exploratory modeling has become one of the cornerstones of the field of decision-making under deep uncertainty (Marchau et al., 2019).

It is quite unfortunately that the authors seem to be utterly unaware, with the exception of Davis et al. (2007), of this rich existing body of literature on exploratory modeling. The reported case study could have benefitted from various techniques for model analysis like scenario discovery (Bryant & Lempert, 2010; Kwakkel & Jaxa-Rozen, 2016) and global sensitivity analysis (Razavi et al., 2021). In particular, dynamic scenario discovery (Kwakkel et al., 2013; Steinmann et al., 2020) could have been useful in identifying characteristics dynamics over time, as well as their conditions for occurring. Dynamics scenario discovery and global sensitivity analysis could substantially enhance the mapping of the state space of possibilities by identifying the conditions under which the different possibilities would occur. So, in their Bangladeshi case study, these techniques could have helped in answering questions about the representativeness of the three scenarios for the entire ensemble, and the regions...
of the model input space from which these three scenarios originate. This, in turn, would enable grounding the resulting narratives much more strongly in the underlying model (Greeven et al., 2016).

Lustick and Tetlock (2021) strongly emphasize the importance of theoretically grounding the simulation models. Unfortunately, for many social phenomena, we have a variety of possible theoretical explanations. Rather than narrowing down on a preferred theory, exploratory modeling is increasingly being used for opening up the conversation by exploring over a range of alternative theories. For example, Pryut and Kwakkel (2014) explore the rise of home-grown terrorism following two rival theories and use this to identify policy interventions that work in either case. Both Mitchell (2003) and Page (2018) argue more generally for using a plurality of structurally different simulation models because such an ensemble sheds a rich light on the phenomenon under study, which is particularly appropriate in case of complex systems.

Next to the potential of various techniques that have emerged over the last two decades for exploratory modeling, there have also been ample relevant theoretical developments that are directly relevant to the manifesto. Most importantly, what exactly is the status of individual computational experiments, and what is the status of the ensemble? A key question in this context is whether one can take the frequencies of occurrence of different classes of model results within an ensemble, and directly translate these frequencies into probabilities that hold in the real world. For obvious reasons, you cannot. The frequencies found in model land are critically dependent on the way in which one samples the model input space. Are the various dimensions that characterize the model input space taken as independent, or are correlations assumed? What prior is assumed for each of the dimensions? Modeling choices related to these questions strongly alter the observed frequencies (Quinn et al., 2020). So, instead of treating frequencies of possibilities in model land as probabilities that also hold in the real world, Shortridge and Zaitchik (2018) and Taner et al. (2019) explore ways of using scenario discovery results as input to an a posteriori Bayesian probabilistic assessment. This might offer an alternative way of synthesizing the inside and outside view as discussed by Lustick and Tetlock (2021) because of equifinality. Building on this, Oreskes (1998) argues that one cannot proof the predictive reliability of models of complex systems prior to their actual use. Like with natural systems, in case of geopolitical systems one is confronted with limits to measurability, accessibility and a lack of spatiotemporal stationarity. The yardstick for model quality is thus not to be found in its purported predictive quality (e.g., brier scores), but rather in its perceived usefulness for guiding planning and decision-making.

3 | USING ENSEMBLES OF SIMULATIONS SHOULD MOVE BEYOND SUMMARY STATISTICS OVER THE ENSEMBLE

Ensembles of simulation experiments can meaningfully inform decision-making beyond merely the summary frequencies of types of outcomes. For example, the absence of a particular type of outcome in an ensemble can be highly decision relevant. An example of this is offered by Auping et al. (2014) who analyzed the geopolitical implications of the US’ shale revolution. In short, they first generated an ensemble of energy price dynamics at a global scale. A set of exemplars from this ensemble was subsequently used as input to a state stability model for stress testing a wide variety of rentier states to identify which of those were most vulnerable to price swings. The global energy model, across the ensemble showed low oil prices for the decade of 2010–2020. Under no condition explored by the model would there be high prices. This ran counter to conventional wisdom at, for example, NATO, where the results were presented. However, because of the underlying theory-informed causal model, there was a very good explanation for the absence of such high price simulation runs. This example highlights the value of analyzing the ensemble in more detail, and carefully explaining how the different types of dynamics arise out of the structural assumptions within the simulation models.

4 | CLOSING THOUGHTS

We are in broad agreement with the plea of Lustick and Tetlock (2021) for the use of theory-guided simulations in aiding decision-making on complex issues, and content that this appeal has a reach well beyond geopolitical questions. Saltelli et al. (2019) lament the fact that simulation and modeling is not its own field (c.f., Padilla et al., 2017), yet is being practiced across many fields, because it hampers the development of shared best practices. The simulation manifesto of Lustick and Tetlock (2021) unfortunately exemplifies this regrettable state of affairs because it seems that they are trying to reinvent the wheel. We have highlighted here primarily the rich body of literature on exploratory modeling which could have meaningfully informed and strengthened the manifesto.

Other examples where Lustick and Tetlock (2021) bypass much relevant literature could easily have been given as well. For example, both the shale gas study and the homegrown terrorism study
relied on system dynamics models, rather than agent-based models, as advocated by Lustick and Tetlock (2021). Rather than fruitlessly debating which is the better approach, we content that both can be useful for representing key causal mechanisms, and both produce emergent aggregate dynamics. The primary difference is the level at which causal mechanisms are described. System dynamics models use a lumped, mesoscopic representation where the aggregate dynamics are an emergent property of interacting feedback loops involving accumulation and delay. If there is no obvious way of lumping, or compartmentalizing, your agents, a microscopic representation where the aggregate dynamics arise out of local interactions among heterogeneous agents as used in agent-based models is more suitable (c.f., Rahmandad & Sterman, 2008). Similar remarks could be made with respect to federating models, which under labels such as multimodeling, multimodel ecologies, and co-simulation is being practices in various scientific fields (Nikolic et al., 2019).

**ORCID**
Jan Kwakkel https://orcid.org/0000-0001-9447-2954

**REFERENCES**

Atkins, P. W., Wood, R. E., & Rutgers, P. J. (2002). The effects of feedback format on dynamic decision making. *Organizational Behavior and Human Decision Processes, 88*(2), 587–604. https://doi.org/10.1016/S0749-5979(02)00002-X

Auping, W., de Jong, S., Pruyt, E., & Kwakkel, J. H. (2014). The Geopolitical Impact of Shale Gas: The Modelling Approach. Paper presented at the 32nd International Conference of the System Dynamics Society, Delft, The Netherlands, 20-24 July 2014; Delft, the Netherlands.

Bankes, S. C. (1993). Exploratory modeling for policy analysis. *Operations Research, 4*(3), 435–449. https://doi.org/10.1287/opre.41.3.435

Bankes, S. C. (2002). Tools and techniques for developing policies for complex and uncertain systems. *Proceedings of the National Academy of Sciences of the United States of America, 99*(3), 7263–7266. https://doi.org/10.1073/pnas.92081399

Bryant, B. P., & Lempert, R. J. (2010). Thinking inside the box: A participatory computer-assisted approach to scenario discovery. *Technological Forecasting and Social Change, 77*(1), 34–49. https://doi.org/10.1016/j.techfore.2009.08.002

Davis, P. K., Bankes, S. C., & Egner, M. (2007). Enhancing strategic planning with massive scenario generation. https://www.rand.org/pubs/technical_reports/TR392.html

Greeven, S., Kraan, O., Chappin, E. J. L., & Kwakkel, J. H. (2016). The emergence of climate change mitigation action by society – An agent-based scenario discovery study. *Journal of Artificial Societies and Social Simulation, 19*(3), 9. https://doi.org/10.18354/jasss.3134

Hodges, J. S. (1991). Six (or so) things you can do with a bad model. *Operations Research, 39*(3), 355–365. https://doi.org/10.1287/opre.39.3.355

Hodges, J. S., & Dewar, J. A. (1992). Is it you or your model talking? A framework for model validation (R-4114-FA/OSD). http://www.rand.org/pubs/reports/2006/R4114.pdf

Kwakkel, J. H. (2017). The Exploratory Modeling Workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. *Environmental Modelling and Software, 96*, 239–250. https://doi.org/10.1016/j.envsoft.2017.04.054

Kwakkel, J. H., Auping, W. L., & Pruyt, E. (2013). Dynamic scenario discovery under deep uncertainty: the future of copper. *Technological Forecasting and Social Change, 80*(4), 789–800. https://doi.org/10.1016/j.techfore.2012.09.012

Kwakkel, J. H., & Jaxa-Rozen, M. (2016). Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes. *Environmental Modelling and Software, 79*, 311–321. https://doi.org/10.1016/j.envsoft.2015.11.020

Lustick, I. S., & Tetlock, P. E. (2021). The simulation manifesto: The limits of brute-force empiricism in geopolitical forecasting. *Futures and Foresight Science, e64*. https://doi.org/10.1002/fff2.64

Marchau, V. A. W., Walker, W. E., Bloemen, P., & Popper, S. (Eds.). (2019). *Decision making under deep uncertainty – From theory to practice*. Springer.

Mitchell, S. (2003). *Biological complexity and integrative pluralism*. Cambridge.

Moallemi, E., Kwakkel, J. H., de Haan, F., & Bryan, B. A. (2020). Exploratory modeling for analyzing coupled human-natural systems under uncertainty. *Global Environmental Change, 65*, 102186. https://doi.org/10.1016/j.gloenvcha.2020.102186

Nikolic, I., Warnier, M., Kwakkel, J. H., Chappin, E. J. L., Lukszo, Z., Braizer, F., & Palensky, P. (2019). *Principles, challenges and guidelines for a multi-model ecology*. https://doi.org/10.4233/uuid:1aa3d16c-2acd-40ce-b6b8-0712fd947840

Oreskes, N. (1998). Evaluation (not validation) of quantitative models. *Environmental Health Perspectives, 106*(6), 1453–1460. https://doi.org/10.1289/ehp.9810661453

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science, 263*(5147), 641–647.

Padilla, J. J., Diallo, S. Y., Lynch, C. J., & Gore, R. (2017). Observations on the practice and profession of modeling and simulation: A survey approach. *Simulation, 94*(6), 493–506. https://doi.org/10.1177/0037597817737159

Page, S. E. (2018). The model thinker. *Basic Books.*

Pruyt, E., & Kwakkel, J. H. (2014). Radicalization under deep uncertainty: a multi-model exploration of activism, extremism and terrorism. *System Dynamics Review, 30*(1–2), 1–28. https://doi.org/10.1002/sdr.1510

Quinn, J. D., Hadjimichael, A., Reed, P. M., & Steinschneider, S. (2020). Can exploratory modeling of water scarcity vulnerabilities and robustness be scenario neutral? *Earth’s Future, 8*(11), e2020EF001650. https://doi.org/10.1029/2020EF001650

Rahmandad, H., & Sterman, J. D. (2008). Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science, 54*(5), 998–1014. https://doi.org/10.1287/mnsc.1070.0787

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo, Pian, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J. H. A., Jakeman, J., Gupta, H., Mellilo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., ... Maier, H. R. (2021). The future of sensitivity analysis: An essential discipline for systems modeling and policy support. *Environmental Modelling and Software, 137*, 104954. https://doi.org/10.1016/j.envsoft.2020.104954

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., & Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling and Software, 114*, 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012

Shortridge, J. E., & Zaitchik, B. F. (2018). Characterizing climate change risks by linking robust decision frameworks and uncertain productivity projections. *Climatic Change, 151*, 525–539. https://doi.org/10.1007/s10584-018-2324-x

Steinmann, P., Auping, W. C., & Kwakkel, J. H. (2020). Behavior-based scenario discovery using time series clustering. *Technological Forecasting and Social Change, 156*, 120052. https://doi.org/10.1016/j.techfore.2020.120052
Sterman, J. D. (1989). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43*(3), 301–335. https://doi.org/10.1016/0749-5978(89)90041-1

Sterman, J. D. (1994). Learning in and about complex systems. *System Dynamics Review, 10*(2–3), 291–330. https://doi.org/10.1002/sdr.4260100214

Sterman, J. D. (2002). All models are wrong: reflections on becoming a systems scientist. *System Dynamics Review, 18*(4), 501–531. https://doi.org/10.1002/sdr.261

Taner, M. Ü., Ray, P. A., & Brown, C. (2019). Incorporating multidimensional probabilistic information into robustness-based water systems planning. *Water Resources Research, 55*, 3659–3679. https://doi.org/10.1029/2018WR022909

**How to cite this article:** Kwakkel J, Auping W. A commentary on Lustick and Tetlock (2021). *Futures & Foresight Science* 2021;3:e84. https://doi.org/10.1002/ffo2.84