Modeling Decisions in Collective Risk Social Dilemma Games for Climate Change Using Reinforcement Learning

Medha Kumar*, Kapil Agrawal, and Varun Dutt
Applied Cognitive Science Laboratory
School of Computing and Electrical Engineering
Indian Institute of Technology Mandi, Kamand, India - 175005
* medha751@gmail.com, varun@iitmandi.ac.in

Abstract*— Prior research has used reinforcement-learning (RL) models like Expectancy-Valence-Learning (EVL) and Prospect-Valence-Learning (PVL) to investigate human decisions in choice games. However, currently little is known on how RL models would account for human decisions in games where people face a collective risk social dilemma (CRSD) against societal problems like climate change. In CRSD game, a group of players invested some part of their private incomes to a public fund over several rounds with the goal of collectively reaching a climate target, failing which climate change would occur with a certain probability and players would lose their remaining incomes. Next EVL and PVL models were calibrated to human decisions across two between-subject conditions in CRSD (Info-all: N=120; No-Info: N=120), where half of the players in each condition possessed lesser wealth (poor) compared to the other half (rich). A symmetric Nash model was also run in both conditions as a benchmark. In Info-all condition, players possessed complete information on investments of other players after every round; whereas, in the No-info condition, players did not possess this information. Our results showed that for both rich and poor players, the EVL model performed better than the PVL model in No-info condition; however, the PVL model performed better than the EVL model in the Info condition. Both the EVL and PVL models outperformed the symmetric Nash model. Model parameters showed reliance on recency, reward-seeking, and exploitative behaviours. We highlight the implications of our model results for situations involving a collective risk social dilemma.

Keywords*— *Collective risk social dilemma, decision-making, reinforcement-learning, Expectancy-Valence-Learning model, Prospect-Valence-Learning model.

I. Introduction

Climate change is a global phenomenon and, since the late 19th century, Earth's average surface temperature has already risen by about 1.8 degrees Fahrenheit (1.0 degree Celsius) [1]. Although the average surface temperature has been increasing and this increase poses a threat to mankind, people continue to show a waiting approach towards climate change [2-4]. This waiting approach has been prevalent in climate negotiations and a likely reason for this approach could be the dilemma that people face when they need to decide whether to keep their private wealth to themselves or to contribute some part of it for mitigating climate change [5-9].

Prior research has used a collective risk social dilemma (CRSD) game to study climate negotiations in the laboratory [6-7, 9]. In CRSD, a group of six-players are provided with initial private endowments that they can contribute for mitigating climate change across several rounds. In each round, all players contribute either 0, 2, or 4 units against climate change, and after all players have decided their investments, all players get to know how much other players contributed in the last round and since the beginning of the game. The first three rounds are operated by a computer, where three randomly chosen players are made to contribute 4 units (poor players) and the other three players are made to contribute 0 units (rich players). All players need to collectively reach a climate goal after 10-active rounds of investments, where players decide the investments themselves. If players fail to collectively reach the target, they lose their leftover endowments completely with a probability because of climate change.

Prior literature has used the CRSD game to test the effects of varying endowments and pledges on investments against climate change [9]. However, this literature assumed that players got full-information about other players' investments [9]. Reference [6] extended this information limitation and tested how information availability influences investments in the CRSD game. Specifically, [6] presented two conditions to their participants, where the conditions differed in terms of information available about investments to different players (rich and poor). Reference [6] found that possessing information about investments of other players produced an overall higher investment to the climate fund and higher success rates. Although there are differences in people's investment behaviour in the presence or absence of investment information, little is known on how different underlying cognitive mechanisms contribute to these investment differences. One way of studying the influence of underlying cognitive mechanisms is to develop computational cognitive models [10].

The primary objective of this research is to understand how certain reinforcement learning (RL) mechanisms in cognitive models account for people's decision-making in the CRSD game in the presence or absence of investment information. Specifically, we use two computational cognitive RL models [11-14] to account for people's decision-making in the CRSD game: Expectancy-Valence-Learning (EVL) [15] and Prospect-Valence-Learning (PVL) [16]. EVL model has three parameters, where one parameter each account for people's loss-aversion, recency, and explorative-exploitative behaviours. The PVL model improves the EVL model with an additional fourth parameter from Prospect Theory [17], where this additional parameter captures the shape of people's utility function for losses and gains. In this paper, we use the EVL and PVL models to account for people's investment decisions in the presence and absence of investment information in the CRSD game.

In what follows, first, we detail background literature. Next, we propose the EVL and PVL models in the CRSD game and detail how we calibrated different model parameters in each of the two information conditions in CRSD. Finally, we discuss our results and highlight their implications for decisions in collective risk social dilemma situations.

II. Background

A. EVL and PVL Models

Prior research has used the EVL and PVL models in choice tasks [14]. For example, [14] used both the EVL and PVL models to understand the decision-making of different population (brain-damaged subjects, drug-abusers, Asperger subjects, and older-aged subjects) on the Iowa Gambling Task (IGT) [18]. Reference [19] have tested variations of the EVL and PVL models by calibrating these models to each participant's choices in choice tasks. Furthermore, [20] have examined the impact of losses on expensive exploratory search in a binary choice tasks using EVL and PVL models. Reference [21] investigated the exploration behaviour before making final decisions in binary-choice tasks using the EVL and PVL models. These authors found that losses caused more exploration compared to gains and the PVL model outperformed the EVL model in fitting human exploration behaviour.

Prior research involving IGT has revealed that among both the EVL and PVL models, the PVL model and its variants fit to human data better compared to the EVL model [11, 12, 16, 22, 23]. For example, [11] found that an alternate utility function in the PVL model helped this model to provide a better fit to human decisions in IGT. Similarly, [16] calibrated different variants of the PVL model to different classes of drug users. These authors found that the PVL model and its variants explained the data better compared to heuristic rules. Reference [22] employed the PVL model to decompose IGT performance into component processes in healthy and marginally housed persons with substance use disorders (MHP-SUD). Application of the PVL model revealed a better fit to human data in IGT among MHP-SUD subjects. Furthermore, [24] have used the EVL and PVL models to assess motivational, memory, and response processes among chronic cannabis abusers and control participants. These authors found a variant of the PVL model to perform better in fitting human data in IGT compared to the EVL model.

Although a number of prior attempts have investigated EVL and PVL models in the IGT and binary-choice tasks, to the best of authors' knowledge, there is still lack of research that investigates these models in applied judgement tasks involving multiple players. In this paper, we address this research question by proposing EVL and PVL models for the CRSD game and investigating the performance of these models to fit human data collected in the CRSD game.

B. Collective Risk Social Dilemma (CRSD)

A collective-risk social dilemma implies that personal endowments will be lost if the collective group contributions to a common pool are too small [7, 25]. Reference [7] proposed a CRSD game in connection to climate change to investigate a group's contribution to reach a target knowing that climate change could occur with a probability if the group failed to reach the target. Results revealed that under high risk of simulated dangerous climate change, half of the groups succeeded in reaching the target sum, whereas the others only marginally failed. Recently, [26] investigated how residual risk of failure of climate change policies affects willingness to contribute to such policies in CRSD. These authors found that investments were higher at least in the final part of treatments including a residual risk. In this paper, we calibrate the EVL and PVL models possessing different cognitive parameters to players' investments in the CRSD game.

III. An Experiment involving Information Asymmetry in CRSD

A. Experimental Design

Participants were randomly assigned to different groups of six-participants each across two between-subject conditions in a laboratory experiment: Info (N = 20 groups) and No-info (N = 20 groups). In Info condition, after each round, all players were provided information about other players' individual investments in the last round as well as information about the cumulative investments of the entire group since the start of the game. In the No-info condition, players did not possess investment information about other players' individual investments in the last round. However, players were provided information about the total cumulative investment of the group since the start of the game.

B. The CRSD Game

In CRSD, players have to collectively reach a target failing which they lose their personal endowments with a probability. A group of six-players are provided with initial private endowments (= 52 units) that they can invest to a public fund for mitigating climate change across 13-rounds. In each round, all players invest either 0, 2, or 4 units from their initial endowment against climate change. After all players have decided their investments in a round, all players are provided feedback. As part of the feedback, players may or may not get to know about other players' investments in the last round. However, all players are provided feedback about their group's cumulative investments since the beginning of the game. There is a total of 13-rounds in CRSD, where the first three rounds are operated by a computer. In the first three rounds, three randomly chosen players are made to contribute 4 units (poor players) and the other three players are made to contribute 0 units (rich players). All six players need to collectively reach a climate target (= 156 units) after 10-active rounds of investments (from round 4 to round 13), where players decided the investments themselves. If players were successful in reaching the climate target, then they could keep their leftover endowments in real money as per a conversion rule. However, if players fail to collectively reach the climate target, then they would lose their leftover endowments completely with a 50% probability of climate change.

C. Participants

There were 240 students recruited through an email advertisement for a climate change study at the Indian Institute of Technology Mandi, India. Students were randomly divided into 40 groups of 6 participants each (i.e., 20 groups per condition). Participants were undergraduate and graduate students in computer engineering, mechanical engineering, electrical engineering, basic sciences, and humanities and social sciences. Age ranged from 18 years to 31 years (Mean = 20 years; Min = 18 years; Max = 31 years). Two hundred and sixteen participants were males and rest females. Participants were paid as per the following rule: INR 30 as base payment and a performance bonus. The performance bonus was computed using participants' left-over endowments as per the following formula: 1-unit endowment left = INR 0.50 left in real money.

D. Procedure

Participation was voluntary, and participants signed a consent form before starting their experiment. First, participants provided their demographic information and read the instructions related to the study. This step was followed by game play where participants were asked to play CRSD game in their group for 13-repeated rounds. After completing the study, participants were thanked and paid their compensation. Results of this experiment will be presented with model results ahead in this paper.

IV. EVL AND PVL MODELS

Table 1 shows the equations for the utility function, learning rule, choice rule, and sensitivity in the EVL and PVL models [12]. In the CRSD game, people can contribute an outcome *k*, where *k* ε {0 units, 2 units, or 4 units} on a round t. First, in any round t, in both models, we calculate the utility functions for different possible contributions *k* (0, 2, or 4) using the appropriate equations and parameters shown in Table 1. The EVL and PVL models start-off with a utility equation wherein it is important to specify the win and loss that any decision-maker receives along the game. At any round t, the win function W(t) is defined as the player's remaining endowment after making t-1 investment decisions in CRSD. Mathematically, W(t) = Initial endowment – Sum of investments made till t-1 rounds. Similarly, the loss function at any round t is calculated by taking the minima of the investments all other players made subtracted by their own decision at round t. Mathematically, L(t) = Min (Investment at round t by all other players) – Investment by the player at round t.

After defining the win and loss functions for any round t, we can calculate the utility, learning, and choice rules for both EVL and PVL models from Table 1. In the EVL model, the w parameter (the loss-aversion parameter), is the weight that participants assign to losses relative to gains [11, 12]. A small value of w, i.e., w ≤ 0.5, characterizes decision-makers who put more weight on the rewards and can thus be described as reward-seeking. Whereas, a large value of w, that is, w > 0.5, characterizes decision makers who put more weight on losses and can thus be described as loss-averse [16].

Table 1. The EVL and PVL models, model parameters, and parameter ranges

Concept	Models	Model Equation	Free parameter	Range
Utility function	EVL	$V_k(t) = (1-w) * W(t) + w * L(t)$	w: Loss aversion parameter	[0,1]
	PVL	$u(t) = \{ \begin{matrix} x(t)^{\alpha} & if\ (x(t) \geq 0) \\ -w\lvert x(t)\rvert^{\alpha} & if\ (x(t) < 0) \end{matrix}$	α: Shape parameter w: Loss aversion parameter	[0, 1] [0, 5]
Learning rule	EVL	$EV_K(t+1)) = (1-a) * E(V_K(t)) + a * V_K(t)$	a: Recency parameter	[0, 1]
	PVL	$\mathrm{EV_k}(t+1) = a * \mathrm{EV_k}(t) + \delta_k(t) + u_k(t)$	a: Recency parameter	[0, 1]
Choice Rule	All	$P[(S_k(t+1)] = \frac{e^{\{(\theta(t))(E(V_k)\}}}{\sum_{j=1}^{4} \theta(t) EV_j}$		
Sensitivity	EVL	$\theta(t) = (t/10)^c$	c: Consistency parameter	[-5, 5]
	PVL	$\theta(t) = 3^c - 1$	c: Consistency parameter	[0, 5]

Note: W(t) and L(t) are the rewards and losses, respectively on trial t. x(t) is the net outcome on trial t, x(t) = W(t) - |L(t)|. $\delta_k(t)$is a dummy variable that takes the value 1 if outcome k is chosen on trial t and 0 otherwise.

The PVL utility function contains the two parameters—the shape parameter, A, and the loss aversion parameter, w [11, 12]. As A approaches zero, the shape of the utility function approaches a step function. In contrast, as A approaches one, the subjective utility, $u_k(t)$, increases in direct proportion to the net outcome, $x(t)$. In the PVL model, a value of w larger than one indicates a larger impact of losses than gains on the subjective utility; whereas, a value of w of one indicates equal impact of losses and gains. As w approaches zero, the PVL model predicts that losses will be neglected. Second, in both models, we calculate the expected utility ($EV_K(t)$) as per the appropriate learning rule. This updating process in EVL model is influenced by the recency parameter, a [11, 12]. A value of a close to zero indicates slow forgetting and weak recency effects; whereas, a value of a close to one indicates rapid forgetting and strong recency effects. In contrast, in the PVL model, a small value of a indicates rapid forgetting and strong recency effects. Whereas, a large value of a parameter indicates slow forgetting and weak recency effects. For both models under consideration, we initialized the expectancies of all outcomes (0, 2, and 4) to zero, $EV_K(0) = 0$. In the next round, the models assume that the expected utilities of each outcome are used to guide the choices of participants [11, 12]. This assumption is formalized by the softmax choice rule, which in both models computes the probability of choosing a particular outcome on a round [27].

The choice rule contains the sensitivity parameter, θ, which indexes the extent to which round-by-round choices match the expected utilities of different outcomes. In the EVL model, the sensitivity parameter θ changes based upon the response consistency parameter c. If c is positive, the sensitivity of round-by-round choices to the expected utilities of different outcomes increases over rounds; otherwise, the sensitivity decreases. In PVL model, small values of c cause a random choice pattern; whereas, large values of c cause a deterministic choice pattern.

V. Symettrical Nash Model

Nash equilibria provide optimal solutions in games [28]. In the CRSD, there are several Nash equilibria possible as there are multiple players and different combinations of investments over 13-rounds may achieve a target of 156 units for the group. However, one Nash equilibria, which ensures symmetry among all players, assumes that all players contribute 2 units per round (i.e., 12 units per group per round or 156 units across 13-rounds). We use this symmetrical Nash model as a baseline to compare the performance of the EVL and PVL models. As EVL and PVL possess cognitive assumptions, they are likely to outperform the symmetric Nash model [28].

VI. Model Parameter Calibration

In both models, we ran the same number of simulated participants as the number of human participants that participated in the experiment. Next, we computed the Sum of Squared Deviation (SSD) using the following formula:

$$SSD = \sum_{i=1}^{13}[(model\ cum\ invest\ rich_i - human\ cum\ invest\ rich_i)^2 + (model\ cum\ invest\ poor_i - human\ cum\ invest\ poor_i)^2)] \quad (1)$$

Where $model\ cum\ invest\ rich_i$, $human\ cum\ invest\ rich_i$, $model\ cum\ invest\ poor_i$ and $human\ cum\ invest\ poor_i$ refer to the average cumulative investments from model and human players, respectively, in round i. To compare models with different parameters, we used the Akaike information criterion (AIC) that takes into account both a model's ability to predict human data and its complexity in terms of number of parameters contained [29]. The AIC was defined in the following manner:

$$AIC = 13 * \ln\left(\frac{SSD}{13}\right) + 2 * k \quad (2)$$

Where, k refers to the number of free parameters calibrated in a model. For EVL, PVL, and Nash models, the value of k were 3, 4, and 0. We used genetic algorithm to minimize the AIC values in both EVL and PVL models. The optimization ran for a minimum of 250 generations for each model. The genetic algorithm has population size = 20, a crossover rate of 80%, and a mutation rate of 1%. The algorithm stopped when any of the following constraints were met: stall generations = 100, function tolerance = $1\text{x}10^{-8}$, and when the average relative change in the fitness function value over 100 stall generations was less than function tolerance ($1\text{x}10^{-8}$).

VII. Results

We analysed the average cumulative investments from human data and both EVL and PVL models. First, in agreement with [6], the Info condition showed much larger investments over rounds compared to those shown in the No-Info condition. Second, in the Info condition, based upon AIC minimization, the PVL model fitted the human data better compared to the EVL model for rich and poor players (AIC_{PVL} (24.412) < AIC_{EVL} (78.166)). Third, as expected, the symmetric Nash model was outperformed by both the EVL and PVL models across both Info and No-Info conditions (Info: AIC_{Nash} (84.471); No-Info: AIC_{Nash} (83.682)). In order to check the stability of AIC, we also varied individually parameters a, w, c and A around their optimum values. Results revealed that the AIC did not become flat at the local minima.

Table 2 shows the calibrated values of EVL and PVL model parameters for rich and poor players in the Info and No-Info conditions, respectively. The best set of parameters (corresponding to the lowest AIC values) has been italicized.

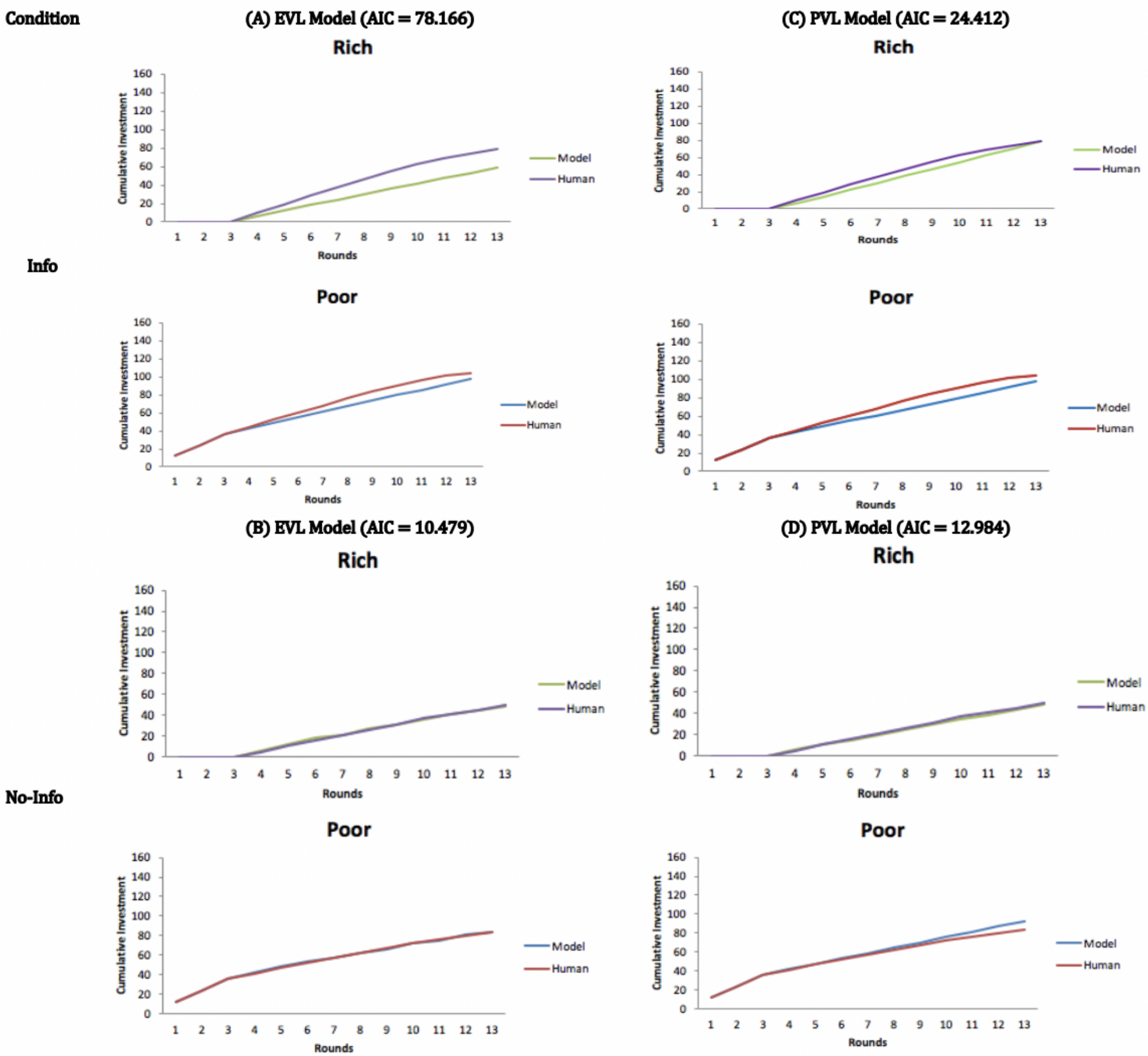

Figure 1. Human data and model data from EVL and PVL models for rich and poor players over 13-rounds in Info and No-Info conditions. (A) Human data and EVL model in Info condition. (B) Human data and EVL model in No-Info condition. (C) Human data and PVL model in Info condition. (D) Human data and PVL model in No-Info condition.

Table 2. Calibrated parameter values from the EVL and PVL models in different experimental conditions.

Model	Info				No-Info			
	Rich		Poor		Rich		Poor	
EVL	a	0.001	a	0.740	*a*	*0.999*	*a*	*1.000*
	w	0.043	w	0.409	*w*	*0.001*	*w*	*0.144*
	c	-1.480	c	3.542	*c*	*0.516*	*c*	*3.974*
PVL	*A*	*0.153*	*A*	*0.730*	A	1.000	A	0.325
	a	*0.001*	*a*	*0.709*	a	1.000	a	1.000
	w	*4.517*	*w*	*0.000*	w	4.111	w	0.350
	c	*2.992*	*c*	*4.143*	c	1.908	c	1.911

Note. The italicized parameters correspond to the model that fitted the human data better in the respective condition.

In the Info condition, the PVL model revealed that rich players' utilities were not much influenced by the net outcomes x(t) (the A parameter was close to 0). Second, the model showed strong influence of recency among rich players' decisions the a parameter possessed a very small value). Third there was presence of loss-aversion in rich players' decision--making (the w parameter was far exceeding 1.0). Fourth, rich players' decisions seemed to be deterministic (the c parameter possessed a large value). Furthermore, the PVL model revealed that poor players' utilities increased in direct proportion to the net outcome x(t)'s increase (the A parameter was close to 1). Second, the model showed weak influence of recency among poor players' decisions (the a parameter possessed a large value). Third, there was a strong neglect of losses among poor players' decision-making (the w parameter was 0). Fourth, poor players' decisions seemed to be largely deterministic (the c parameter possessed a very large value).

In the No-Info condition, first, the EVL model revealed strong influence of recency among rich players' decisions (the a parameter possessed a very large value close to 1). Second, there was a strong drive towards rewards and a neglect of losses in rich players' decision-making (the w parameter was close to 0). Third, rich players' decisions tended to be deterministic (the c parameter possessed a small value close to 0). Furthermore, the EVL model revealed strong influence of recency among poor players' decisions (the a parameter possessed a very large value close to 1). Second, there was a strong drive towards rewards and a neglect of losses in poor players' decision-making (the w parameter was close to 0). Third, poor players' decisions tended to be explorative (the c parameter possessed a large positive value).

VIII. Discussion and Conclusions

Although a number of prior attempt have investigated EVL and PVL models in the IGT and binary-choice tasks [12, 21], research has yet to explore the potential of these models in applied judgement tasks involving multiple players. The primary objective of this paper was to overcome this literature gap. Specifically, in this paper, we investigated the ability of EVL and PVL models to fit human investment decisions in the presence or absence of investment information in CRSD. Results revealed that in the presence of information about opponent's last investments, the PVL model performed better compared to the EVL model in fitting human decisions. However, in the absence of information about opponent's last decisions, the EVL model performed better compared to the PVL model in fitting human decisions. These results are in contrast to those in IGT, where the PVL model and its variants have been found to be consistently better compared to the EVL model [11, 12, 16, 22, 23]. Both the EVL and PVL models outperformed the symmetric Nash model. Furthermore, the model parameters best fitting human decisions across different conditions revealed differences and similarities among rich and poor players' decision-making when investment information about opponents was available and not available, respectively.

First, we found that the PVL model did not consistently outperform the EVL model across both information conditions. A likely reason for this finding could be the differences in the task used in our study compared to those used in prior research. In prior research, mostly the tasks used involve making choices between available options (e.g., IGT and binary-choice). Mostly, these tasks are played by a single decision-maker repeatedly. However, the CRSD task in this paper was a judgment task involving multiple players, where different players in a group had to collectively decide how much to invest against climate change repeatedly.

Second, we found that both the EVL and PVL models outperformed the symmetric Nash model in fitting to human decisions in CRSD. A likely reason for this finding could be the presence of cognitive assumptions and parameters in the EVL and PVL models and the absence of such assumptions in the Nash model. As explained above, both the EVL and PVL models possessed cognitive mechanisms like recency, weight to losses versus gains, net outcome's influence on decision-maker's utility, and the reliance on explorative versus exploitative behavior. Perhaps, these mechanisms enabled the

EVL and PVL models to fit human decisions accurately. However, the symmetric Nash model seem to rely solely upon mathematical (rational) assumptions without any reliance on cognitive (bounded-rational) assumptions [28]. Thus, deviations of the Nash model from human data basically showed people to be bounded rational beings who did not invest in a symmetrical manner.

Third, we found that recency played a dominant role in shaping our results among both rich and poor players when information about opponents' investments was known or when it was unknown. In fact, recency has been shown to influence choice behavior in different tasks, even in those tasks that belong to domains other than the environment [30-32]. According to [31-32], recency effects show-up in decisions from experience tasks via theories of cognition like the Instance-based Learning Theory. Thus, the presence of recency of information among the EVL and PVL models is in agreement with broader literature on judgement and decision-making involving both single decision-makers [31] as well as multiple decision-makers [33].

Fourth, we found that reward-seeking behavior and neglect of losses seem to be higher when information about opponents' investments was absent compared to when this information about opponents' investments was present. We speculate that when opponents' investments are not shown, players are less motivated towards reaching the climate goal. This lack of motivation, perhaps, makes them keep their endowments to themselves and seek rewards. In addition, the presence of information about opponent's investments may likely improve the players' drive towards the climate goal due to social influence [34]. Overall, across most conditions, players tend to show reward-seeking behavior to save their endowments.

Fifth, we found that rich (poor) players' utilities were influenced (not influenced) by net outcomes in the information condition. One likely reason for this finding could be that rich players possess greater endowment compared to poor players after the first three rounds in the game. It could be that the perception of greater endowments among rich players makes the PVL model disregard the net outcome in the game.

Overall, our results showed deterministic decision-making from both rich and poor players in conditions when the information about opponents' investments was present compared to when it was absent. However, the lack of investment information perhaps does not allow players to start trusting this information. Thus, poor players show less determinism and more explorative decision-making.

We did not consistently find a single model (EVL or PVL) to explain all experimental conditions. From our results, it seems that the PVL model is a better choice compared to the EVL model when information about opponent's decision is available. However, the opposite is true when information about opponent's decision is not available. Overall, the EVL and PVL models may help us to explore negotiations in CRSD over a longer time horizon lasting several rounds. Also, we may use EVL and PVL models to predict human performance in conditions involving different probabilities of climate change and different investment amounts.

As part of future work, we plan to extend our investigation of RL models to other models that either combine the EVL and PVL assumptions or those that make more optimal decisions. In addition, we would also like to explore the ability of RL models to explain experimental conditions where only a subset of players possess investment information, i.e., either the poor players possess rich players' investment information, or the rich players possess poor players' investment information. Some of these ideas form the immediate next steps in our research program involving CRSD.

ACKNOWLEDGMENTS

This research was partially supported by a grant (IITM/SG/VD/32) to Dr. Varun Dutt. Also, we are thankful to Indian Institute of Technology Mandi for providing computational resources for this project.

REFERENCES

[1] IPCC. Climate Change 2014: Mitigation of Climate Change. Vol. 3. Cambridge University Press, 2015.

[2] Dutt V, Gonzalez C. Decisions from experience reduces misconceptions about climate change. *Journal of Environment Psychology*, 32(1), pp.19-29, 2012a.

[3] Dutt V, Gonzalez C. Human control of climate change. *Climatic Change*, 111(3-4), pp.497-518, 2012b.

[4] Ricke KL, Caldeira K. Natural climate variability and future climate policy. *Nature Climate Change*, 4(5), pp.333-338, 2014.

[5] Kumar, M., Chouhan, R., & Dutt V. Role of Information Asymmetry in a Public Goods Game for Climate Change. 24th Conference on Behavior Representation in Modeling and Simulation (BRIMS), Washington DC, USA, 2015.

[6] Kumar, M., & Dutt, V. Collective Risk Social Dilemma: Role of information availability and income inequality in achieving cooperation against climate change (manuscript submitted for publication), 2018.

[7] Milinski M, Sommerfeld RD, Krambeck HJ, Reed FA, Marotzke J. The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, 105(7), pp.2291-2294, 2008.

[8] Milinski, M., Röhl, T., & Marotzke, J. Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Climatic Change*, *109*(3-4), 807-814, 2011.

[9] Tavoni A, Dannenberg A, Kallis G, Löschel A. Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences*, 108(29), pp.11825-11829, 2011.

[10] Busemeyer, J. R., & Diederich, A. *Cognitive modeling*. Sage, 2010.

[11] Dai, J., Kerestes, R., Upton, D. J., Busemeyer, J. R., & Stout, J. C. An improved cognitive model of the Iowa and Soochow Gambling Tasks with regard to model fitting performance and tests of parameter consistency. *Frontiers in psychology*, *6*, 229, 2015.

[12] Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., &Wagenmakers, E. J. Performance of healthy participants on the Iowa Gambling Task. *Psychological assessment*, *25*(1), 180, 2013.

[13] Sutton, R. S., & Barto, A.G. Reinforcement learning: An introduction. Cambridge, MA: The MIT Press, 1998.

[14] Yechiam, E., Busemeyer, J. R., Stout, J. C., &Bechara, A. Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science*, *16*(12), 973-978, 2005.

[15] Busemeyer, J. R., & Stout, J. C. A contribution of cognitive decision models to clinical assessment: decomposing performance on the Bechara gambling task. *Psychological assessment*, *14*(3), 253, 2002.

[16] Ahn, W.-Y., Busemeyer, J. R., Wagenmakers, E.-J., & Stout, J. C.Comparison of decision learning models using the generalization criterion method. *Cognitive Science*, *32*(8), 1376–1402, 2008.

[17] Tversky, Amos; Kahneman, Daniel. "Advances in prospect theory: Cumulative representation of uncertainty". Journal of Risk and Uncertainty. 5 (4): 297–323, 1992.

[18] Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1), 7-15, 1994.

[19] Kudryavtsev, A., & Pavlodsky, J. Description-based and experience-based decisions: individual analysis. *Judgment and Decision Making*, *7*(3), 316, 2012.

[20] Lejarraga, T., Pachur, T., Frey, R., & Hertwig, R. Decisions from experience: From monetary to medical gambles. *Journal of Behavioral Decision Making*, *29*(1), 67-77, 2016.

[21] Lejarraga, T., & Hertwig, R. How the threat of losses makes people explore more than the promise of gains. *Psychonomic bulletin & review*, *24*(3), 708-720, 2017.

[22] Baitz, H. A. *Component processes of decision making in persons with substance use disorders* (Doctoral dissertation, Arts & Social Sciences: Department of Psychology, 2016.

[23] Konstantinidis, E., Speekenbrink, M., Stout, J. C., Ahn, W. Y., & Shanks, D. R. To simulate or not? Comment on Steingroever, Wetzels, and Wagenmakers (2014).

[24] Fridberg, D. J., Queller, S., Ahn, W. Y., Kim, W., Bishara, A. J., Busemeyer, J. R., ... & Stout, J. C. Cognitive mechanisms underlying risky decision-making in chronic cannabis users. *Journal of mathematical psychology*, *54*(1), 28-38, 2010.

[25] Chen X, Szolnoki A, Perc M. Risk-driven migration and the collective-risk social dilemma. Physical Review E, 86(3), 036101, 2012.

[26] Farjam, M., Nikolaychuk, O., & Bravo, G. Does risk communication really decrease cooperation in climate change mitigation?. *Climatic change*, *149*(2), 147-158, 2018.

[27] Luce, R. Individual choice behavior. New York: Wiley, 1959.

[28] Camerer, C. F. Strategizing in the brain. *Science*, *300*(5626), 1673-1675, 2003.

[29] Pitt, M. A., Myung, I. J., & Zhang, S. Toward a method of selecting among computational models of cognition. *Psychological review*, *109*(3), 472, 2002.

[30] Dutt, V., Ahn, Y. S., & Gonzalez, C. Cyber situation awareness: modeling detection of cyber attacks with instance-based learning theory. *Human Factors*, *55*(3), 605-618, 2013.

[31] Gonzalez, C., & Dutt, V. Instance-based learning: Integrating sampling and repeated decisions from experience. Psych. Review, *118*(4), 523-551, 2011.

[32] Gonzalez, C., & Dutt, V. Refuting data aggregation arguments and how the instance-based learning model stands criticism: A reply to Hills and Hertwig (2012). Psych. Review 119(4), 893-898, 2012.

[33] Gonzalez, C., Ben-Asher, N., Martin, J. M., & Dutt, V. A cognitive model of dynamic cooperation with varied interdependency information. *Cognitive science*, *39*(3), 457-495, 2015.

[34] Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. The constructive, destructive, and reconstructive power of social norms. *Psychological science*, *18*(5), 429-434, 2007.