Impact of Latent State Cues on Behavior in Repeated Games

Ismail Guennouni (i.guennouni.17@ucl.ac.uk)
Department of Experimental Psychology, University College London

Kaitlyn Ng (kaitlyn.ng.19@ucl.ac.uk)
Department of Experimental Psychology, University College London

Maarten Speekenbrink (m.speekenbrink@ucl.ac.uk)
Department of Experimental Psychology, University College London

Abstract

In social interactions, inferring the interaction partner's hidden mental state is crucial for predicting their actions and optimizing our responses. Effective models for this inference must account for how these mental states evolve due to the interaction history and environmental changes. For example, recognizing someone's emotional state can help forecast their behavior. Our study investigates how making these latent states visible influences decision-making in social interactions. Using the repeated trust game paradigm, we show how to use hidden Markov models (HMM) to formally represent latent state dependent strategies of the players. HMMs fitted to human dyadic play in the trust game are then used to specify adaptive AI agents that simulate changes in mental dispositions of human players, such as the level of trust in the opponent, during a repeated interaction. Making these artificial HMM based agents take the role of the investor and interact with real human trustees, we then explore how displaying "emotion" cues to the opponent's latent state affects people's actions. We find that the presence of cues was associated with more cooperative behavior from the human trustees, and that patterns of behavior that promote the maintenance of cooperation emerged in the presence of latent state cues and were transferred to settings where the cues were subsequently hidden.

Keywords: Social interaction; Repeated economic games; Hidden Markov models; Cooperation; Hidden state cues.

Introduction

In social interactions, decision making is not solely determined by our understanding of the interaction's goals and rules. It is also shaped by our beliefs and expectations about the people we interact with, as well as by the influence of emotion. Trust, a fundamental element of social interaction, can be defined as the willingness to accept vulnerability based on positive expectations about another person's intentions or behavior (Rousseau, Sitkin, Burt, & Camerer, 1998). Emotions play a crucial role in shaping trust, influencing both our own decision-making and how we interpret the actions of others. Research, such as that reviewed by Angie, Connelly, Waples, & Kligyte (2011), has extensively explored how a decision maker's own emotions impact their actions. For example, Dunn & Schweitzer (2005) and Harlé & Sanfey (2007) have shown that inducing positive or negative emotions in the decider leads to an increase or decrease in decisions consistent with trust, respectively.

Emotions not only influence our decisions but also play a critical role in how we interpret and predict others' actions. The Trust Game (Berg, Dickhaut, & McCabe, 1995), particularly in its repeated form, serves as an effective paradigm for studying the interplay between the decision maker's beliefs, emotions, and behaviors. In this game, the "investor" decides how much of an endowment to send to the other player (the "trustee"). The amount that is sent is tripled and the trustee decides, in return, how much of the tripled amount to send back to the investor. The investor's total reward is the amount remaining of their endowment and the amount sent back by the trustee. The trustee's reward is the amount of the tripled investment they keep to themselves. In a one-shot game, a purely selfish trustee would keep all of the tripled investment. A purely selfish investor, realising this, would therefore not invest anything, and keep all of the endowment to themselves. In the repeated version, rewards for both players are maximised if they build trust and share the benefits of higher investments and returns. In this context, trusting behavior hinges on the belief in a partner's likelihood to reciprocate trust (Camerer, 2003). Emotional cues are important for forming those beliefs. When shown faces that invoke happiness, either through photographs of people (Averbeck & Duchaine, 2009; Scharlemann, Eckel, Kacelnik, & Wilson, 2001) or drawings (Eckel et al., 2003), participants exhibited a higher level of trust in one-shot games. Krumhuber et al. (2007) further found that the perceived genuineness of a smile affects trust levels, with genuine smiles garnering more trust than fake ones. Schug, Matsumoto, Horita, Yamagishi, & Bonnet (2010) highlighted that a lack of emotional expression can lead to perceptions of untrustworthiness.

In repeated economic games, most studies did not allow participants to perceive the emotional state of their opponent, either because the opponent was not shown or photographs of neutral faces were used. A notable exception is work by Tortosa, Strizhko, Capizzi, & Ruz (2013) that explored how faces displaying happiness and anger affect decisions in a repeated Trust Game. They found that participants were able to learn associations between the emotion displayed and the tendency of the opponent to cooperate. Displayed emotions can thus be used as cues to predict how an interaction partner might act, and as such, to inform best responses to the partner's actions. For instance, Stratou, Hoegen, Lucas, & Gratch (2015) showed that smiling players were more likely to be exploited in an repeated prisoner's dilemma, because a smile was a reliable cue of the opponent's intention to cooperate.

The current study builds on these foundations to explore how "emotion" cues for the opponent's latent state affects

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

people's actions in the repeated Trust Game. We use hidden Markov models of investor behaviour to program artificial agents that take the role of the investor. We make these artificial agents play against real human trustees, with either hidden or visible cues to the latent state of the HMM agent. We expect these "emotion" cues of the investor's latent state to lead to more cooperative behavior, as expressions of the investor's satisfaction may reinforce the reward from cooperative actions, and cues to displeased states may encourage the trustee to coax investors back into cooperation. We also expect the emotion cues to aid learning, as the hidden state of the investor can be accurately inferred from the emotion. As such, we expect participants to develop a good policy more quickly when emotions are displayed. Furthermore, this policy may be generalized to later games where emotions are no longer displayed.

Methods

Participants A total of 60 participants were recruited on the Prolific Academic platform (29 males, 30 females, and 1 preferred not to say). We invited participants with an age between 18 and 65 (M = 28.7 years, SD = 8.52 years) and an approval rating higher than 95% on the site. Participants were paid a fixed fee of £2.50 for completing the experiment plus a bonus of up £1.50 (average = £0.74) dependent on their performance on a random trial in which the opponent made an investment of 10.

Design The experiment had a 2 (Condition: Emotion Shown or Not Shown) by 2 (Order: Emotion First or No-emotion First) design, with repeated measures on the first factor. Participants were always assigned the role of the trustee in every condition. They were randomly assigned to one of the two levels of the second factor. For the first factor, participants either saw a cue to the investor's emotional state (as an emoji), or no cue was given. These conditions were separated into two blocks respectively, each containing a total of 50 trials to give participants sufficient opportunity to familiarise themselves with the task and learn from the outcomes of their decisions. The order in which participants played the two blocks was randomised. They would either play the block in which the investor's emotional states were shown first, followed by the block where the investor's emotional states were hidden, or vice versa.

Investor Strategy The strategy of the computerised investor was modelled on behavior of human investors in the Repeated Trust Game (RTG) over 10-rounds with the same co-player. The dataset consisted of a total of 381 games from two data sources: First, a total of 93 repeated trust games with healthy investors and a mix of healthy trustees and trustees diagnosed with Borderline Personality Disorder (BPD) (King-Casas et al., 2008). The second source was from data collected as part of a project investigating social exchanges at Virginia Tech and consists of 288 games including a mix of healthy participants and those suffering from either BPD or Anti-social Personality Disorder (ASPD). The deliberate inclusion of both healthy participants and those with personality disorders, particularly disorders known to impact social functioning, was strategic. This diversity enables our Hidden Markov Model (HMM) to capture a broad spectrum of behavioral responses. By incorporating participants with these specific disorders, our model not only reflects a wider range of human interactions but also enhances its predictive accuracy and generalizability in scenarios akin to real-world social exchanges

Using this data, we estimated a hidden Markov model (HMM) on investors' behavior with three latent states using maximum likelihood estimation via the Expectation-Maximisation algorithm as implemented in the depmixS4 (Visser & Speekenbrink, 2021) package for R. Each latent state was associated with a state-conditional discretized normal distribution over the possible investments from 0 to 20 (Figure 1). The fitted distributions can be seen to reflect "low-trust", "medium-trust", or "high-trust". Over rounds, the investor can move between latent states, and the probability of these transitions was modelled as a function of their net return (i.e return - investment) in the previous round (see Figure 2). On all rounds, the investor's actions were determined by randomly drawing an investment from the state-conditional distribution. The investor transition to a new state was determined by randomly drawing the next state from the state-transition distribution as determined by the net return on the previous round. The initial state for the HMM investor in each instance of the game was the "mid-trust" state.

The advantages of this approach is that it does not require strong a priori assumptions about the model features. The number of states, the policy conditional on the state, and the transition function between states can all determined in a purely data-driven way. These HMMs can in turn be used to simulate a human-like agent playing the trust game. This agent may transition to a new state depending on the other player's actions and adopt a policy reflecting its state, thus simulating changes in e.g. emotional dispositions of human players during a repeated game. When the investor gains from the interaction, they become more likely to transition to a state where their policy is more "trusting" with generally higher investments. However, faced with losses, the investor is more likely to transition to a more cautious policy with generally lower investments. The policies and the transitions between states are sufficient to build an agent that reflects this type of adaptive behavior and reacts to the trustee's action choices in a way that mimics a human player.

Materials The main component of the study was a gamified version of an investment trust game (Berg et al., 1995), which was produced using the Gorilla experiment builder (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020). Participants played the game with an artificial opponent, with participants taking on the role of the trustee and the artificial opponent taking on the role of the investor, or in this case, a

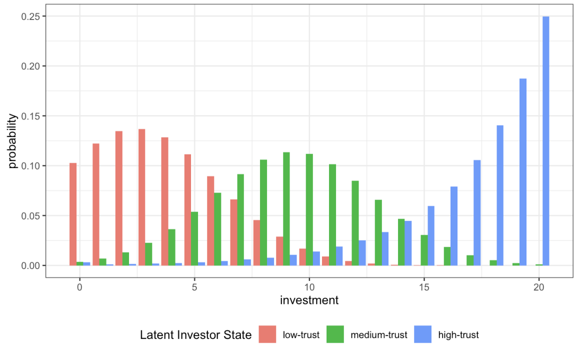

Figure 1: Probability mass function of the investor's policy conditional on its latent state as an output of the 3 state HMM using a truncated discretised Gaussian as a response function

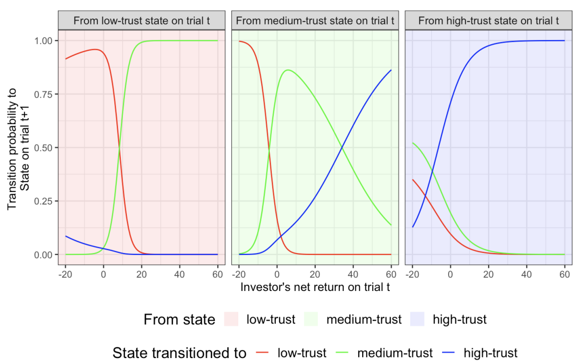

Figure 2: Transition probabilities of the investor state as a function of investor's net return in the current round. Each panel is the state transitioned from, and each line shows the probability of transitioning to the state identified by the line color.

farmer and a landowner respectively. Instead of money, participants would receive tomato seeds as a form of investment from the investor. Once fully grown, each seed produces three tomatoes, and the participant decided how much of the harvest to return to the investor. The investor was depicted using one of four images of a landowner: three of which were to be used in the condition where the investor's emotion was shown (images of a landowner with an unhappy, neutral and happy emoji respectively), with the remaining image being used in the condition where the investor's emotion was not shown (image of a landowner with an empty emoji). Each investment and return were represented by matching images of seeds and tomatoes.

Procedure Before beginning the study, participants were asked to report gender and age. Participants were then informed of their role as a farmer (trustee) and were told that on each trial they would receive a number of seeds from their landowner as an investment. After their crops had fully grown, they would yield a number of tomatoes three times the number of seeds, and would then need to decide how

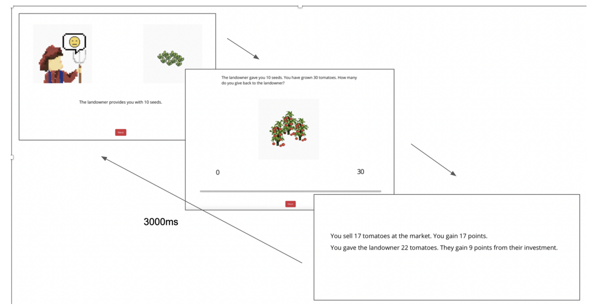

Figure 3: Overview of the timeline of the experiment detailing the various phases participants go through in each round. In this gamified version of the Trust game, participants first see the amounts of seeds given to them by the landowner (Investor), as well as an emoji representingn its latent state. This is followed by a screen where they indicate how many tomatoes to send back and finally a recap screen.

many tomatoes sell at the market and how many to return to the landowner (investor) as returns on their investment. The participants were encouraged to find a balance between making a profit whilst maintaining a good working relationship with the landowner. These instructions were then followed by three general attention check questions ("What is the maximum number of seeds the landowner can provide you with on a single trial?", "If you're given 10 seeds, how many tomatoes will you grow?", "What is the price of a seed compared to the price of a tomato?"), to ensure that the participants had understood the prior instructions correctly. There were an additional three attention check questions shown before the condition in which the investor's emotional states (unhappy, neutral, happy) were shown to ensure that participants could correctly identify the investor's emotions from the images.

For each trial, participants were shown an image of the investor alongside an image of the seeds given to them. Participants would always receive 10 seeds from the investor in the first trial, and the investor's latent state would always be neutral, irrespective of whether an emotion cue was shown or not. Participants were then shown an image of their total tomato yield and were asked to select how many tomatoes they wished to return to the investor using a slider. After participants made their decision, the results of that decision (i.e. the number of tomatoes sold at the market and returned to the investor) were shown to them for 3000ms before the next trial began. A timeline with screenshots of the various stages of a round is shown in Figure 3. After completing all 50 trials in a block, participants would then be redirected to the second block, consisting of 50 trials for the condition they had not undergone. After completing the experiment, participants were debriefed and compensated monetarily for their time, which included both their base payment rate of £2.50 and any bonus payment they may have earned from the experiment.

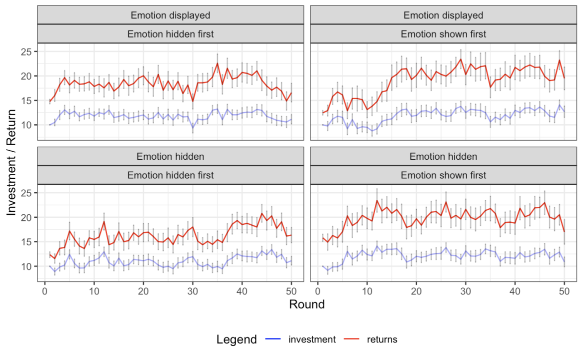

Figure 4: Average and standard errors of the investment and return per round for each condition and Order

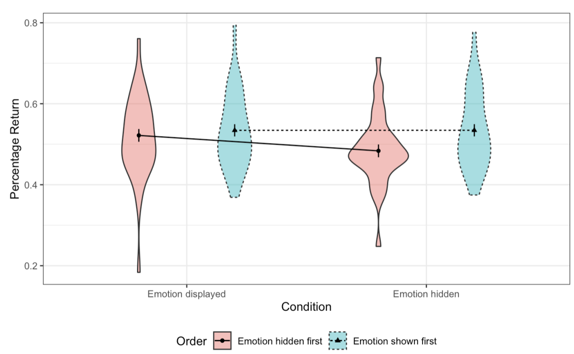

Figure 5: Marginal means and distributions of percentage trustee returns over all rounds, shown across participants by Order and Condition

Statistical analysis To determine the effects of the condition on returns (as a percentage of Investment) we estimated a linear mixed-effects model with fixed-effects of Order (emotion first, no-emotion first), Condition (emotion shown, emotion hidden), and Investment (normalised). We also include interaction terms between all the aforementioned covariates. We included participant-wise random intercepts and slopes for Condition and Investment and allowed correlation between these random effects to be estimated. For the F-tests, we used the Kenward-Roger approximation to the degrees of freedom, as implemented in the R package afex (Singmann et al., 2022). Unless otherwise stated, the default adjustment method for multiple comparison is Tukey's HSD. We Z-transform the Investment variable as centering is beneficial to interpreting the main effects more easily in the presence of interactions.

Results

Figure 4 shows the average investment (blue) and returns (red) and their standard error on each round by order and condition. We can see that in both conditions, the returns are higher than the investments throughout the rounds, indicating that on average people establish a cooperative pattern where the investors are rewarded for their trust.

Estimated marginal means of percentage returns and their standard errors are presented in Figure 5. We find a main effect of Condition ($F(1, 58.07) = 6.39, p = 0.01$) with higher percentage returns when emotion is displayed. We also find an interaction effect between Order and Condition ($F(1, 58.1) = 6.52, p = 0.01$). When participants played the emotion display block first, there was no significant difference between the percentage returns in the two conditions. However, when they played that block last, they sent back significantly higher returns to the investors when the emotion was displayed compared to when it was hidden ($\Delta M = 0.04$, 95% CI $[0.02, 0.06]$, $z = 3.79$, $p < .001$). Further, there was a three way interaction effect between the Condition, Order and Investment ($F(1, 58.1) = 13.6, p < 0.001$), indicating a different reaction to the investment depending on the condition and order of blocks.

In order to simplify the interpretation of this interaction, we replace the Investment variable by the investor's latent state in a new model. This has the added benefit of directly exploring whether making this latent state visible through the emotion display affects the return. We estimate a new linear mixed-effects model with fixed-effects Order (emotion first , no-emotion first), Condition (emotion shown, emotion hidden) and investor State (low-trust, medium-trust or high-trust). We also include interaction terms between all the aforementioned covariates. We included participant-wise random intercepts and slopes for Condition.

In this new model, we find a main effect for the investor's state ($F(2, 5835) = 18.72, p < 0.001$)). Percentage returns when the investor was in a low-trust state were significantly higher than when they were in the high-trust ($p < 0.001$) or medium-trust state ($p < 0.001$). There was no significant difference between the returns when the investor was in the medium-trust versus the low-trust State ($p = 0.16$). Furthermore, we find an interaction effect between Order and the investor's state ($F(2, 5835) = 5.18, p = 0.005$). This effect reflects higher returns in the low-trust state compared to the high-trust ($p < 0.001$) and medium-trust ($p < 0.001$) states when emotions are shown first, but not when the no-emotion games are played first ($p = 0.14$ for high vs low trust and $p = 0.16$ for the medium vs low trust). We also find an interaction effect between Condition and the investor State ($F(2, 5720.6) = 4.54, p = 0.01$). Post-hoc analysis shows percentage returns being higher when emotion is shown compared to when it is hidden in both the medium-trust ($p = 0.01$) and low-trust state ($p = 0.03$).

As in the previous model with investments as a covariate, we still find evidence for the interaction between Order and Condition ($F(1, 68.4) = 8.45, p = 0.004$). Finally, we find a three-way interaction between Order, Condition and investor state ($F(2, 5742) = 9.95, p < 0.001$). To make sense of this three-way interaction, Figure 6 shows the marginal means of the percentage returns for each latent investor state and Condition, with the panels representing the two levels of order. When the investor is in a low-trust state, we see a differen-

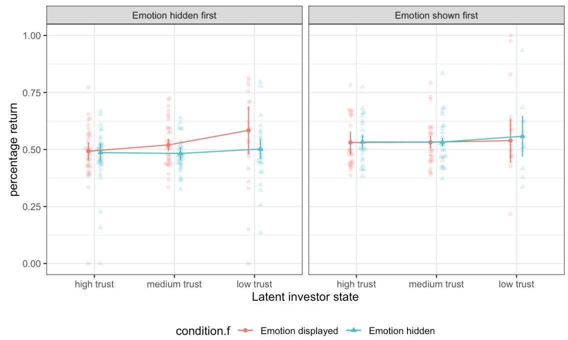

Figure 6: Marginal means, standard errors and beeswarm of percentage trustee returns over all rounds, shown across participants for each latent investor state by Order (panels) and Condition (color)

tiated behaviour depending on the order of play: When the emotion is hidden first, participant returns are significantly higher when the emotion is then displayed compared to when it was hidden ($p < 0.001$). The same applies in the medium-trust state ($p = 0.006$). However, when the emotion is displayed first, there is no significant difference between participant's returns between the two conditions irrespective of the investor's state.

In summary, when emotion is hidden initially, the subsequent display of emotions leads to a differentiated behavior, whereby less "trusting" investor states are associated with higher percentage returns by the trustee. This would be consistent with a coaxing behavior by the trustee to get the investor's back into cooperation. However, when emotion is shown first, the behavior of the trustee is similar across states between the emotion and no-emotion condition. This suggest that participants learn to associate the emotion to the true latent state of the investor and transfer their learning, through learned policy, to the second block where emotion is no longer displayed.

We can also test whether these differences in percentage returns were associated with a difference in the investor's behavior between conditions. To analyse the artificial agent investments, we fit a linear mixed effects model with fixed effects of Condition and Order and their interaction. We included participant-wise random intercepts and slopes for Condition. We find no main effects of Order and Condition and no interaction effect between the covariates, indicating that the HMM agent's investments were not affected by whether or not the emotions were shown or the order in which the participants played. In other words, any difference in returns cannot be explained by a differentiated behavior from the artificial agent playing the role of the investor.

Discussion

This experiment used adaptive artificial agents, modeled on human behavior, as investors in a repeated Trust Game (RTG) against human participants as trustees. A key finding is that regardless of whether or not the artificial investor displayed emotions as cues to its latent state, interactions resulted in sustained cooperation. Agents continually invested over half their endowment, and participants reciprocated with returns exceeding the investment (around 50% of the total yield).

The main research question of the current experiment was whether providing cues to the opponent's latent state in the RTG would lead to differentiated behavior by human participants. Whilst we found that displaying cues reflecting the latent state did lead to higher percentage returns, the more notable effect was the interaction between the order of play and the condition (emotions shown or not). When participants were initially presented with reliable cues (emotions) to the latent states, they may have learned to predict the latent state of the investor better, and subsequently transferred what they learned to the condition where no explicit cues to the latent state were given. Conversely, when the participants started the experiment without the emotional cues, their returns increased once these cues were later provided. This indicates that higher returns, a proxy for trustworthiness are associated with the trustee's ability to accurately predict the latent state of the investor. The effect size of the interaction between Order and Condition (partial eta-squared: $\eta_p^2 = 0.10$) is comparable to the effect sizes of an intervention focused on gratitude intervention (Drazkowski, Kaczmarek, & Kashdan, 2017) and higher than the effect sizes shown in an intervention priming participants with he concepts of friend or foe with respect to the co-player (Burnham, McCabe, & Smith, 2000). This result has implications for experiments where decreased trust and trustworthiness were observed in certain patient populations, such as patients with Borderline Personality Disorder (King-Casas et al., 2008; Unoka, Seres, Áspán, & Nikoletta, 2009). Potentially, the lower returns associated with the absence of emotional cues might be linked to lower social inference ability such as the actions of the investor are not translated accurately into their correct emotional disposition in this population.

When using the latent state as a covariate in our model we also found that participants displayed behavior consistent with coaxing the opponent into cooperation once emotional cues had been displayed. Indeed, participants who were shown the emotion cues first sent back higher returns when the investor was "unhappy" (a proxy for low-trust), and transferred this strategy to the subsequent games where the emotion cues were hidden. Participants' strategic use of higher returns in the low-trust state demonstrates a rational approach, aiming to incentivize and reinforce the investor's future cooperation. For those who were not shown the state cues at the start, their returns were significantly higher in the second game of the experiment, indicating that coaxing behavior is dependent on the correct prediction of the latent state. This finding is consistent with those of Stratou et al. (2015) as the investor in our study displayed both positive emotions to signal cooperative intentions, but also expressed unhappiness that their cooperation was taken advantage of. Furthermore, the coaxing behavior exhibited by the participants when the emotional cue indicated the investor has low trust confirms the findings of Antos, De Melo, Gratch, & Grosz (2011), whose participants made more concessions when their opponent displayed expressions of anger. This result is also consistent with the mechanistic explanation for the absence of coaxing when trustees suffer from Borderline Personality Disorder proposed by King-Casas et al. (2008). If indeed coaxing is contingent on the correct prediction of the state, then an inability to mentalise about the state of the opponent, which associated with social dysfunction in BPD patient (Allen & Fonagy, 2006), could indeed offer an explanation to the inability to coax behavior seen in these patient populations.

While these findings offer valuable insights, it's important to acknowledge the potential limitations in directly generalizing these results to human-human interactions. Artificial agents may not fully capture the nuances of human emotional expression and interpretation, which could influence trust dynamics in real-world settings. For instance, since the investor starts in a medium-trust state, if the human participant keeps rewarding the trust and the investor keeps making a small net return, it is unlikely the investor would move to a non-cooperative state. This prevents us from measuring the ability of participants to repair a breakdown of cooperation fully if it does not happen in the first place. We intend to run future experiments where the investor is programmed to break down cooperation at predetermined rounds to look at whether participants coax back cooperative behavior and how we can encourage that sort of coaxing behavior through cognitive interventions. Another limitation is the absence of emotion measurement of the participants' reaction to both the investments and the emotional cues. There may be heterogeneity in the way people react to this information and this could affect how they perceive their opponent's actions and what returns they send back. In future studies, we intend to incorporate an "online" emotion elicitation tool to measure the trustee's reaction as the investor's action is revealed. This would allow us to explore whether the self-reported emotions are affected by the condition or the order of showing the investor's emotion.

Acknowledgements

We are grateful to Andreas Hula, Tobias Nolte and the team at Virginia Tech for assisting us in obtaining data for human dyadic play in the repeated Trust Game. I. Guennouni was supported by the UK Engineering and Physical Sciences Research Council under grant EP/S515255/1. The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

References

Allen, J. G., & Fonagy, P. (2006). *The Handbook of Mentalization-Based Treatment*. John Wiley & Sons.

Angie, A. D., Connelly, S., Waples, E. P., & Kligyte, V. (2011). The influence of discrete emotions on judgement and decision-making: A meta-analytic review. *Cognition And Emotion*, *25*(8), 1393–1422. http://doi.org/10.1080/02699931.2010.550751

Antos, D., De Melo, C., Gratch, J., & Grosz, B. (2011). The Influence of Emotion Expression on Perceptions of Trustworthiness in Negotiation. *Proceedings of the AAAI Conference on Artificial Intelligence*, *25*(1), 772–778. http://doi.org/10.1609/aaai.v25i1.7939

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. http://doi.org/10.3758/s13428-019-01237-x

Averbeck, B. B., & Duchaine, B. (2009). Integration of social and utilitarian factors in decision making. *Emotion*, *9*(5), 599–608. http://doi.org/10.1037/a0016509

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, *10*(1), 122–142. http://doi.org/10.1006/game.1995.1027

Burnham, T., McCabe, K., & Smith, V. L. (2000). Friend-or-foe intentionality priming in an extensive form trust game. *Journal of Economic Behavior & Organization*, *43*(1), 57–73. http://doi.org/10.1016/S0167-2681(00)00108-6

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.

Drazkowski, D., Kaczmarek, L. D., & Kashdan, T. B. (2017). Gratitude pays: A weekly gratitude intervention influences monetary decisions, physiological responses, and emotional experiences during a trust-related social interaction. *Personality and Individual Differences*, *110*, 148–153. http://doi.org/10.1016/j.paid.2017.01.043

Dunn, J. R., & Schweitzer, M. E. (2005). Feeling and Believing: The Influence of Emotion on Trust. *Journal of Personality and Social Psychology*, *88*(5), 736–748. http://doi.org/10.1037/0022-3514.88.5.736

Eckel, C., Wilson, R., Aksal, F., Bowles, R., Dziepak, T., Gilles, R., … Smith, V. (2003). The Human face of game theory: Trust and reciprocity in sequential games. *Trust and Reciprocity: Interdisciplinary Lessons for Experimental Research*.

Harlé, K. M., & Sanfey, A. G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. *Emotion*, *7*(4), 876–881. http://doi.org/10.1037/1528-3542.7.4.876

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., & Montague, P. R. (2008). The Rupture and Repair of Cooperation in Borderline Personality Disorder. *Science*, *321*(5890), 806–810. http://doi.org/10.1126/science.1156902

Krumhuber, E., Manstead, A. S. R., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion*, *7*(4), 730–735. http://doi.org/10.1037/1528-3542.7.4.730

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*(3), 393–404.

Scharlemann, J. P. W., Eckel, C. C., Kacelnik, A., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, *22*(5), 617–640. http://doi.org/10.1016/S0167-4870(01)00059-9

Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T., & Bonnet, K. (2010). Emotional expressivity as a signal of cooperation. *Evolution and Human Behavior*, *31*(2), 87–94. http://doi.org/10.1016/j.evolhumbehav.2009.09.006

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M. S., Højsgaard, S., … Christensen, R. H. B. (2022, April). Afex: Analysis of Factorial Experiments.

Stratou, G., Hoegen, R., Lucas, G., & Gratch, J. (2015). Emotional signaling in a social dilemma: An automatic analysis. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)* (pp. 180–186). Xi'an, China: IEEE. http://doi.org/10.1109/ACII.2015.7344569

Tortosa, M. I., Strizhko, T., Capizzi, M., & Ruz, M. (2013). Interpersonal Effects of Emotion in a Multi-Round Trust Game. *Psicologica: International Journal of Methodology and Experimental Psychology*, *34*(2), 179–198.

Unoka, Z., Seres, I., Áspán, N., & Nikoletta, B. (2009). Trust Game Reveals Restricted Interpersonal Transactions in Patients With Borderline Personality Disorder. *Journal of Personality Disorders*, *23*, 399–409. http://doi.org/10.1521/pedi.2009.23.4.399

Visser, I., & Speekenbrink, M. (2021, May). depmixS4: Dependent Mixture Models - Hidden Markov Models of GLMs and Other Distributions in S4.