Generating Random Sequences For You: Modeling Subjective Randomness in Competitive Games

Arianna Yuan (xfyuan@stanford.edu) and **Michael Henry Tessler** (mtessler@stanford.edu)
Department of Psychology, Stanford University

Abstract

Generating truly random sequences is hard. When participants are engaged in a competitive game (e.g., Matching Pennies), the sequences they generate are surprisingly *more random* than when given explicit instructions to generate random sequences (Rapoport and Budescu, 1992). To explore this phenomenon, we formalized two probabilistic models of Theory of Mind reasoning about subjective randomness. One model (the *Fair-Coin* model) assumes participants predict their opponents' choices by implicitly assuming that their opponents intend to generate binary sequences that simulate the outcome of tossing a fair coin. The other model (the *Markov* model) assumes participants believe that their opponents intend to generate sequences that simulate the outcome of a Markov process with transition probability equal to 0.5. We find that Theory of Mind models of both the Fair-Coin and the Markov definitions of subjective randomness are able to characterize the calibrated subjective randomness that occurs when participants are playing an iterated competitive game (Matching Pennies), but the Markov Model is better than the Fair-Coin Model in simulating the situation where participants need to specify their choice sequences in advance of the game. The current study suggests that the *calibrated* subjective randomness in competitive games can be explained by the online evaluation of sequence randomness with Theory of Mind reasoning.

Keywords: subjective randomness; Theory of Mind; matching pennies; probabilistic models

Introduction

People are relatively poor at generating random sequences (Bakan, 1960). They produce "subjectively random" sequences by switching between heads and tails, but they switch too often (Lopes & Oden, 1987). However, participants are able to generate sequences that are more "truly random" when feedback is available (Neuringer, 1986), as in competitive games (Rapoport & Budescu, 1992). Many theories have been proposed to account for participants' failure to generate random sequences (Griffiths & Tenenbaum, 2003; Hahn & Warren, 2009), but the phenomenon of *calibrated subjective randomness* in the context of feedback has received comparatively little attention by formal models (however, see West and Lebiere (2001); Lee, Conroy, McGreevy, and Barraclough (2004) for some proposals). Here, we propose that calibrated subjective randomness may result from basic Theory of Mind reasoning about others' models of subjective randomness.

The situation we will examine is the competitive game Matching Pennies. In this game, two agents (the "Matcher" and the "Nonmatcher") each make a binary choice (e.g., 0 or 1) secretly. The Matcher wins if their choices are the same and the Nonmatcher wins if their choices are different. Mathematically, the optimal strategy is to choose the two alternatives (0 or 1) with equal probability. Rapoport and Budescu found that participants generated more truly random sequences when playing in the Matching Pennies game.

To understand why feedback helps calibrate subjective randomness, we must first understand the origin of subjective randomness. Several researchers propose that subjective randomness is the result of instructional biases (e.g., when participants are instructed to generate random sequences, they are encouraged to produce sequences that appear to be orderless; Ayton, Hunt, & Wright, 1989). However, even when participants are not explicitly given such instructions, the nature of the task prompts participants to generate sequences that are more "representative" of the output of a random process (Kahneman & Tversky, 1974). Griffiths and Tenenbaum formalized this idea into a probabilistic model of subjective randomness (Griffiths & Tenenbaum, 2001).

Although these theories explain the nature of subjective randomness, they provide no mechanistic explanation as to why this bias is calibrated in the presence of feedback. In the current investigation, we propose that Theory of Mind reasoning can explain participants' behavior in experimental conditions both with and without feedback. The main goal of this paper is to model and explain the data from Rapoport and Budescu (1992). Below we first review the relevant experiments and results. Next, we introduce two computational models, and explore and compare the model predictions in different experimental conditions. Finally, we discuss the implications and limitations of our models.

Calibrated Subjective Randomness

Rapoport and Budescu (1992) ran a subjective randomness experiment with three conditions. In the **Dyad Condition**, participants were paired to form dyads; each dyad played 150 trials of Matching Pennies, generating a response on each trial. The **Single Condition** was the same as the Dyad condition except that the paired dyads were asked to specify their choices in advance of the 150 rounds and were told that the responses would be matched on a trial-by-trial basis to determine the winner. In the **Randomization Condition**, participants were instructed to generate a sequence of 150 random binary responses to simulate the outcome of tossing an unbiased coin in a non-interactive context.

The key results are the patterns of sequential dependencies. The distributions of sequences of length k (i.e., k-tuples; $k = 2,3,4$) were not uniform as expected under a "true" random generating process. The authors calculated the frequencies of k-tuples (e.g., 3-tuple [0 1 1], 4-tuple [0 0 0 0]), and found that in the Randomization Condition, participants were more likely to generate [0 1 0 1] and [1 0 1 0] than [0 0 0 0] and [1 1 1 1]. They used two statistics to indicate the extent to which the distributions deviate from the outcome of a truly random generating process: the mean absolute deviation (MAD) from expectation and the standard deviation of the observed proportions around their expectation (SD):

$$\mathrm{MAD} = \sum_{j=1}^{2^k} |p_j - 1/2^k|/2^k, k = 2,3,4.$$

$$\mathrm{SD} = \sum_{j=1}^{2^k} [(p_j - 1/2^k)^2/(2^k-1)]^{1/2}, k = 2,3,4.$$

Here p_j stands for the probability of individual k-tuples.

Rapoport and Budescu (1992) found that MAD and SD in the Randomization Condition were the largest, followed by the Single Condition, and finally the Dyad Condition. Participants deviated from what would be expected with truly random sequences the most in the Randomization Condition and the least in the Dyad Condition.

To model these results, we propose that in the Dyad Condition, participants use their opponents' previous choices to predict their opponents' choices in the current trial, assuming that their opponents intend to generate sequences that are "subjectively random". They then generate their responses accordingly (i.e., Matchers try to match, Non-matchers try to mismatch). In addition, we posit that participants also have a desire to have their own sequences be subjectively random. In the Single Condition, participants consider that their opponents know that they are likely to generate subjectively random sequences and adjust their choices accordingly (though without feedback). Finally, in the Randomization Condition, participants simply generate sequences that are subjectively random.

Note that one of the critical components in our model is how participants define "subjective randomness". Different definitions or models of subjective randomness are possible. In a Bayesian setting, the inference of whether or not a sequence is random will depend on the specification of the alternative hypotheses (i.e., what counts as "non-random"). For example, participants may imagine "random" to mean a Markov process with transition rate of 0.5. In this case, "non-random" corresponds to a Markov processes with transition rates other than 0.5, which will generate sequences with too many or too few alternations between 0 and 1, though the total counts of 0s and 1s would be approximately equal. If instead, participants imagine "random" to mean an unbiased coin, ''"non-random" corresponds to tossing a biased coin, which would generate sequences with many 1s or many 0s. Given these different possibilities, we formalize both when simulating the calibrated subjective randomness effect reported in Rapoport and Budescu (1992). All models were implemented in the probabilistic programming language WebPPL (Goodman & Stuhlmüller, 2014), and model code can be found online[1].

[1] https://web.stanford.edu/~xfyuan/psych204Code.

The Fair-Coin Model

Model Description

Randomization Condition The Randomization Condition corresponds to the same experimental scenario described in Griffiths and Tenenbaum (2001). We incorporate their model of subjective randomness into a probabilistic model of communication described in Shafto, Goodman, and Frank (2012). The integrated model assumes that when participants are instructed to generate a random sequence, they try to *convince* the experimenter that the "weight of the coin" is 0.5. This will result in sequences that are more representative of a random sequence, such as [0 1 0]. We denote a sequence of length k to be S_k, which can be viewed as a random variable, and a specific instance of it to be s_k ($k = 2,3,4$). For instance, S_3 can take values like [0 1 0]. We obtain the probability $\mathrm{P}(S_k = s_k)$ by assuming that participants attempt to convince the experimenter that the sequence is randomly generated. The goal of the model is to maximize the probability that the listener (i.e., the experimenter) would think the sequences are generated by a fair coin. The model returns the probability of specific k-tuples such as $\mathrm{P}(S_2 = [01])$.

Dyad Condition In the Dyad Condition, choices are made incrementally. The model assumes participants generate responses based on the previous choice made by their opponent and themselves. Concretely, participant A first simulates different alternatives (0 or 1) her opponent (participant B) might choose in the current trial, and then combine B's previous responses with the current possible responses. Participant A then predicts B's current response according to the probability that the combined sequence is judged as random. Mathematically, the probability of choosing 0 given the previous responses can be calculated using equation (1).

$$\begin{aligned}
&\mathrm{P}(R_k = 0|S_{k-1} = s_{k-1}) \\
=&\frac{\mathrm{P}(R_k = 0|S_{k-1} = s_{k-1})}{\mathrm{P}(R_k = 0|S_{k-1} = s_{k-1}) + \mathrm{P}(R_k = 1|S_{k-1} = s_{k-1})} \\
=&\frac{\mathrm{P}(R_k = 0 \wedge S_{k-1} = s_{k-1})}{\mathrm{P}(R_k = 0 \wedge S_{k-1} = s_{k-1}) + \mathrm{P}(R_k = 1 \wedge S_{k-1} = s_{k-1})} \\
=&\frac{\mathrm{P}(S_k = (s_{k-1}, 0))}{\mathrm{P}(S_k = (s_{k-1}, 0)) + \mathrm{P}(S_k = (s_{k-1}, 1))}
\end{aligned} \quad (1)$$

With equation 1 and the probability $\mathrm{P}(S_k = s_k)$ derived in the Randomization Condition we can compute $\mathrm{P}(R_k = 0|S_{k-1} = s_{k-1})$. For instance, assuming that a player's most recent two choices are [0 1], the probability that he/she would choose 0 in the current trial is given by equation (2), where $\mathrm{P}(S_3 = [010])$ and $\mathrm{P}(S_3 = [011])$ are computed from the model predictions in the Randomization Condition.

$$\begin{aligned}
&\mathrm{P}(R_3 = 0|S_2 = [01]) \\
=&\frac{\mathrm{P}(R_3 = 0|S_2 = [01])}{\mathrm{P}(R_3 = 0|S_2 = [01]) + \mathrm{P}(R_3 = 1|S_2 = [01])}
\end{aligned} \quad (2)$$

$$
\begin{aligned}
&= \frac{\mathrm{P}(R_3 = 0 \wedge S_2 = [01])}{\mathrm{P}(R_3 = 0 \wedge S_2 = [01]) + \mathrm{P}(R_3 = 1 \wedge S_2 = [01])} \\
&= \frac{\mathrm{P}(S_3 = ([01], 0))}{\mathrm{P}(S_3 = ([01], 0)) + \mathrm{P}(S_3 = ([01], 1))} \\
&= \frac{\mathrm{P}(S_3 = [010])}{\mathrm{P}(S_3 = [010]) + \mathrm{P}(S_3 = [011])}
\end{aligned}
$$

After participant A predicts B's choice in the current trial, A will make a choice according to his/her assigned role, i.e., if A is a matcher, then A will match B's response, otherwise A will choose the opposite response. In addition, participants might also be motivated to generate sequences that are subjectively random so that their choice is not easily predicted by their opponents. We include a weight term w capturing how participants balance these two concerns. The larger the w, the more weight participants put on their opponents' potential choices. In the current simulation, the value of w is set to 0.6. Mathematically, the probability of choosing 0 given previous responses for a matcher is:

$$
\begin{aligned}
&\mathrm{P}(R_k^M = 0 | S_{k-1}^M = s_{k-1}^M \wedge S_{k-1}^{NM} = s_{k-1}^{NM}) \\
=& w * \mathrm{P}(R_k^{NM} = 0 | S_{k-1}^{NM} = s_{k-1}^{NM}) + \\
&(1-w) * \mathrm{P}(R_k^M = 0 | S_{k-1}^M = s_{k-1}^M),
\end{aligned} \tag{3}
$$

and a non-matcher:

$$
\begin{aligned}
&\mathrm{P}(R_k^{NM} = 0 | S_{k-1}^M = s_{k-1}^M \wedge S_{k-1}^{NM} = s_{k-1}^{NM}) \\
=& w * \mathrm{P}(R_k^M = 1 | S_{k-1}^M = s_{k-1}^M) + \\
&(1-w) * \mathrm{P}(R_k^{NM} = 0 | S_{k-1}^{NM} = s_{k-1}^{NM})
\end{aligned} \tag{4}
$$

Using a concrete example to illustrate how equation (3) and (4) should be applied, we assume that the most recent two choices made by the matcher is [0 1], and those two made by the non-matcher is [1 1]. The probability of choosing 0 as a matcher given her and her opponent's previous responses is computed using equation (5) and the one as a non-matcher is computed using equation (6):

$$
\begin{aligned}
&\mathrm{P}(R_3^M = 0 | S_2^M = [01] \wedge S_2^{NM} = [11]) \\
=& w * \mathrm{P}(R_3^{NM} = 0 | S_2^{NM} = [11]) + \\
&(1-w) * \mathrm{P}(R_3^M = 0 | S_2^M = [01]),
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
&\mathrm{P}(R_3^{NM} = 0 | S_2^M = [01] \wedge S_2^{NM} = [11]) \\
=& w * \mathrm{P}(R_3^M = 1 | S_2^M = [01]) + \\
&(1-w) * \mathrm{P}(R_3^{NM} = 0 | S_2^{NM} = [11])
\end{aligned} \tag{6}
$$

The value of the unknowns can be obtained from the results of equation (1). With those probabilities, we calculate the distribution of all the possible k-tuples and compare the model prediction with the empirical data (Figure 1).

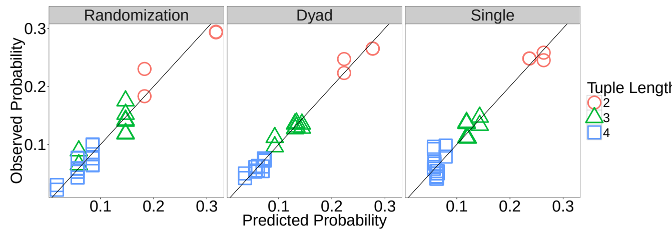

Figure 1: Observed frequencies of k-tuples vs. predicted probabilities of the Fair-Coin Model.

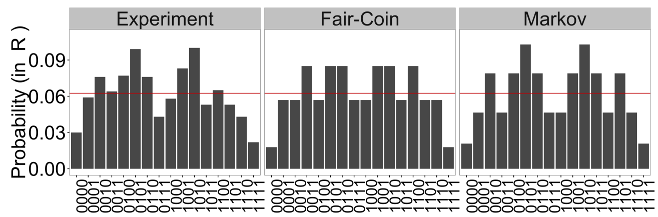

Figure 2: Observed distributions of 4-tuples and model predictions for the Randomization Condition. Left: Empirical data. Middle: Predictions of the Fair-Coin Model. Right: Predictions of the Markov Model. The red line indicates expected probability.

Single Condition For the Single Condition, the model assumes that participants generate responses based on their own previous choices, with the knowledge that their opponent thinks they will generate a random sequence. They extend their previous $k-1$ choices into a k-tuple that is subjectively random. With that subjectively random k-tuple in hand, they make the opposite choice.

Results of the Fair-Coin Model

Randomization Condition Figure 1, Left shows that the model fits the data well, $R^2 = .94$. Figure 2, Middle shows that the model successfully captures the observation that more heterogeneous tuples like [0 1 1 0] are more likely to be generated than less heterogeneous tuples, e.g., [0 0 0 0].

Dyad Condition Although in the Dyad Condition we have different formulae for the matchers and the non-matchers, the simulation results show that the distributions of k-tuples are the same. Therefore, we collapse these two cases (Figure 3, Middle). The model predictions are well aligned with the empirical data (Figure, 1, Middle), $R^2 = .98$. Critically, the model predicts that participants' biased subjective randomness is partially corrected as compared with the Randomization Condition. The MAD and the SD calculated from model predictions in Dyad Condition are much smaller than the ones in the Randomization Condition (Table 1 and 2).

Single Condition In the Single Condition, the formulae for matchers and non-matchers are the same. Therefore, we collapse the two cases. We find that the overall performance

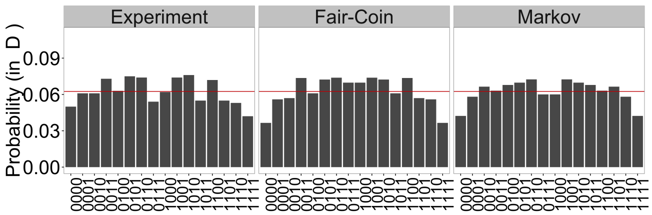

Figure 3: Observed distributions of 4-tuples and model predictions for the Dyad Condition. Plotting conventions are the same as for Figure 2.

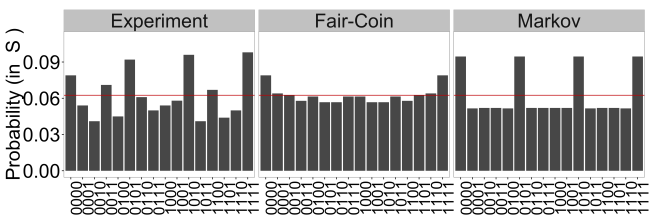

Figure 4: Observed distributions of 4-tuples and model predictions for the Single Condition. Plotting conventions are the same as for Figure 2.

of the model seems to be fine (Figure 1, Right), $R^2 = .94$, though it falls short of capturing the empirical observation that the probability of generating heterogeneous sequences like [0 1 0 1] is high.

The Markov Model

As seen from the model fitting results above, the Fair-Coin Model leaves room for improvement in modeling the Single Condition. Notice that in the empirical data, both sequences that are more representative of the outcome of tossing an unbiased coin (e.g., [0 1 0 1]) and the less representative ones (e.g., [0 0 0 0]) have high probabilities. It is impossible for the Fair-Coin Model to reproduce this effect. We thus explore the possibility that participants adopt a different definition of random sequences (e.g., a Markov process with transition rate of 0.5). This hypothesis has the potential to explain the data in the Single condition because in this case a representative "non-random" sequence would be the outcome of a Markov process with transition rate other than 0.5, thus including too many or too few alternations in the sequences. Indeed, this is what Rapoport and Budescu found. We next explain how this model simulates each condition in the Rapoport and Budescu experiment.

Model Description

Randomization Condition Similar to the Fair-Coin model, the Markov Model assumes that participants try to convince the experimenter that the sequences they give are generated by a random process. However, their notion of a "a random process" is not "tossing an unbiased coin", but rather a generative process with a transition probability $P(R_k \neq R_{k-1})$ equal to 0.5. If there is some bias in the generating process, the transition probability should be less than 0.5, resulting in sequences with fewer alternations, e.g., [0 0 0 0] and [1 1 1 1].[2] As in the Fair-Coin Model, we denote a sequence of length k to be S_k, and a specific instance of it to be s_k ($k = 2,3,4$). Using Bayes' Rule, we obtain the posterior probability of $P(S_k = s_k)$ when participants aim to show the experimenter that the transition probability of the underlying generative process equals to 0.5.

Dyad Condition For the Dyad Condition, the Markov model is very similar to the Fair-Coin Model. It assumes that participants believe that their opponent intends to simulate the outcome of a generative process with transition probability of 0.5. Therefore, they use their opponent's previous choices to predict their opponent's current choice and make a decision according to their prescribed role. At the same time, they are motivated to generate subjectively random sequences so that their own responses are less predictable. Hence, they will try hard to simulate the outcome of a generative process with transition probability of 0.5. Mathematically, the probability of choosing 0 given the previous responses can be calculated using the same equation (1), but now $P(S_k = s_k)$ is obtained from the Markov Model for the Randomization Condition rather than the Fair-Coin Model for the Randomization Condition. We then use Equation (3) and (4) to calculate the probability of choosing 0 as a matcher or a non-matcher conditioned on their own and their opponents' previous responses. w was set to be 0.7 in the Markov Model; since the Markov Model and the Fair-Coin Model have different assumptions, there is no reason that the weight w assigned to the predicted opponents' responses (the Theory of Mind reasoning component) should be equal in these two models.

Single Condition Similar to the Fair-Coin Model, for the Single Condition, the Markov model assumes that participants generate responses based on their own previous choices. Particularly, they know that their opponents think that they intend to generate random sequences. Therefore, in each trial they would make a response so that the opposite of it combined with his/her previous responses will look like an outcome of a Markov process with transition rate equal to 0.5 (see the online code for more detail).

Results of the Markov Model

Randomization Condition Figure 5, Left, shows that the model fits the data well, $R^2 = .95$. Figure 2, Right shows that the model successfully captures the observation that more heterogeneous tuples such as [0 1 1 0] are more likely to be generated than less heterogeneous tuples such as [0 0 0 0]; still, the Markov Model was not statistically significantly

[2]Note that in the Randomization Condition and the Dyad Condition we use this asymmetric prior, whereas in the Single Condition we use a symmetric prior, assuming that a non-random sequence would have either a large transition rate (0.75) or a small transition rate (0.25).

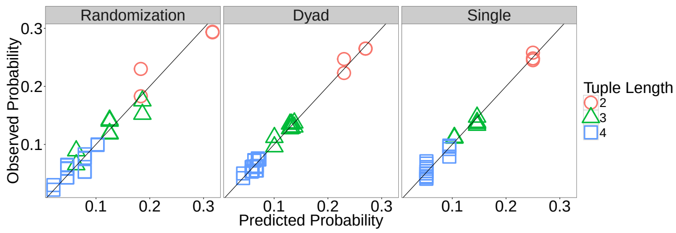

Figure 5: Observed frequencies of k-tuples vs. predicted probabilities of the the Markov Model.

better than the Fair-Coin Model in this condition.

Dyad Condition Although in the Dyad Condition we have different formulae for matchers and non-matchers, the simulation results show that the distributions of k-tuples are the same. Therefore, we collapsed these two cases (Figure 3, Right). The model predictions are well aligned with the empirical data (Figure 5 Middle), $R^2 = .99$. Critically, it predicts that participants' biased subjective randomness is partially calibrated (see the MAD and the SD section).

Single Condition The model fits the data well (Figure 5, Right), $R^2 = .98$. Critically, it captures the findings that both sequences that are more representative of the outcome of tossing an unbiased coin (e.g., [0 1 0 1]) and those that are less representative (e.g., [0 0 0 0]) have higher probabilities. The model works because it assumes that participants avoid producing sequences that are representative of the outcome of a Markov process with transition rate 0.5; thus, participants end up generating sequences with either too many alternations (transition rate larger than 0.5) or too few alternations (transition rate less than 0.5).

Model Comparison From the correlation plot we see that Markov Model seems to fit the data better than the Fair-Coin Model. However, since these two models have different assumptions, traditional statistical tests for model comparison are not applicable. Therefore, we use the R package "cocor" which allows us to directly compare the correlations between the empirical data and the model predictions of these two models (Diedenhofen & Musch, 2014). The results showed that the difference between the two correlations r_{FC} and r_M in the Dyad condition is not significant, Dunn and Clark's $z = -1.49$, $p = .136$. It is also not significant in the Randomization condition, $z = -0.54$, $p = .589$. However, in the Single Condition, the correlation r_{FC} is significantly smaller than the r_M, $z = -3.29$, $p = .001$. Overall the difference between the two correlations r_{FC} and r_M is significant, $z = -2.47$, $p = .014$. In other words, the Markov Model provides a better fit to the data than the Fair-Coin Model. Therefore, in the following section, we only present the MAD and SD calculated from the predictions of the Markov Model.

MAD and SD Consistent with the data, the model predicts the same qualitative results for the MAD and the SD of the three conditions (Table 1 and 2), i.e., the Randomization Condition has the largest deviation from what would be expected with truly random sequences and the Dyad Condition has the smallest one. This suggests that the Markov Model successfully captures the more calibrated subjective randomness in the Dyad Condition.

Table 1: Mean absolute deviation (MAD) of the data and the predictions of the Markov Model. D: Dyad, S: Single, R: Randomization.

	Data			Model		
	D	S	R	D	S	R
2 – tuple	0.0150	0.0043	0.0435	0.0200	0.0000	0.0667
3 – tuple	0.0105	0.0135	0.0273	0.0123	0.0211	0.0309
4 – tuple	0.0087	0.0159	0.0174	0.0068	0.0160	0.0224

Table 2: Standard deviation (SD) around expectations of the data and the predictions of the Markov Model. D: Dyad, S: Single, R: Randomization.

	Data			Model		
	D	S	R	D	S	R
2 – tuple	0.0199	0.0057	0.0534	0.0231	0.0000	0.0770
3 – tuple	0.0141	0.0150	0.0353	0.0155	0.0226	0.0465
4 – tuple	0.0105	0.0194	0.0222	0.0092	0.0191	0.0257

Discussion

Empirical evidence suggests that people generate more truly random sequences in competitive contexts. We explored two probabilistic models to explain the calibrated subjected randomness in a competitive game that was reported in Rapoport and Budescu (1992). We find that Theory of Mind models based on both the Fair-Coin and the Markov formalizations of subjective randomness are able to capture the calibrated subject randomness effects that appear in an iterated competitive game (Dyad Condition vs. Randomization Condition). However, the Markov Model is better than the Fair-Coin Model in explaining the intermediate degree of calibrated subjective randomness that appears in a competitive game where participants must specify their choices ahead of time (the Single Condition).

Why is the Markov Model better than the Fair-Coin Model in simulating the Single Condition? The reason might be that the transition probability of a generative process is more cognitively accessible than "the weight of a coin". When people attempt to generate random sequences, it may be easier to track the transition probability and make sure it approximates 0.5 than to check whether one of the binary responses is made more often than the other. In short, transition probability might be a more convenient heuristic than "the weight of the coin" in evaluating the randomness of sequences.

In addition, it is worth noting that the Markov Model and the Fair-Coin model share the common Theory of Mind reasoning structure. The only difference between these two is the assumption on how people define "random sequences". Kubovy and Gilden (1991) showed that participants attend to multiple numerical properties of the sequence, such as number of alternations, lengths of runs, and imbalance between 0 and 1. The Fair-Coin model focuses on the imbalance between 0 and 1, and the Markov Model focuses on number of alternations and lengths of runs. The results suggest that when online feedback is not available, participants are more likely to rely on number of alternations and lengths of runs to produce unpredictable sequences.

We note some limitations of these models. Both the Fair-Coin Model and the Markov Model assume that participants are probability matching rational agents and generate binary responses in proportion to the interpreted randomness. Therefore, one limitation is that the models cannot predict a player's behavior when his/her opponent does not use the optimal strategy. For example, if matcher "A" plays with a person who chooses "0" more often than "1", A would quickly notice it and choose "0" more (if not always). However, the two models in the current study would not make such predictions because of the assumption that the other agent intends to generate random sequences. Hence, a more complete model may retain uncertainty as to what kind of opponent the participant is playing with. This may also be formalized using a the reinforcement learning algorithm (Lee et al., 2004), and it is worth comparing the assumptions and predictions of the current probabilistic approach with previous reinforcement learning approaches. Another limitation is that we do not explicitly manipulate the number of previous trials the models consider and compare the corresponding performances. However, post-hoc analysis indicates that in both the Fair-Coin model and the Markov model, taking more previous trials into account results in better calibrated subjective randomness, which is consistent with the results of previous connectionist modeling that manipulates the working memory capacity of the models (West & Lebiere, 2001).

In summary, the current investigations suggest Theory of Mind reasoning interacts with participants internal models of subjective randomness in the generation of random sequences in competitive contexts. Future computational approaches should take this into account when modeling subjective randomness.

Acknowledgements

We would like to thank Robert X.D. Hawkins for useful discussions and Noah D. Goodman for his support in teaching Psych 204: Computation and Cognition.

References

Ayton, P., Hunt, A. J., & Wright, G. (1989). Psychological conceptions of randomness. *Journal of Behavioral Decision Making*, *2*(4), 221–238.

Bakan, P. (1960). Response-tendencies in attempts to generate random binary series. *The American Journal of Psychology*, *73*(1), 127.

Diedenhofen, B., & Musch, J. (2014). cocor: A comprehensive solution for the statistical comparison of correlations. *PLOS ONE*, *10*(3), e0121945–e0121945.

Goodman, N. D., & Stuhlmüller, A. (2014). *The Design and Implementation of Probabilistic Programming Languages.* `http://dippl.org`.

Griffiths, T. L., & Tenenbaum, J. B. (2001). Randomness and coincidences: Reconciling intuition and probability theory. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 370–375).

Griffiths, T. L., & Tenenbaum, J. B. (2003). From algorithmic to subjective randomness. In *Advances in Neural Information Processing Systems* (pp. 953–960).

Hahn, U., & Warren, P. A. (2009). Perceptions of randomness: Why three heads are better than four. *Psychological Review*, *116*(2), 454–461.

Kahneman, D., & Tversky, A. (1974). Subjective probability: A judgment of representativeness. In *The concept of probability in psychological experiments* (pp. 25–48). Springer.

Kubovy, M., & Gilden, D. (1991). Apparent randomness is not always the complement of apparent order. In *The perception of structure: Essays in honor of Wendell R. Garner.* (pp. 115–127). American Psychological Association.

Lee, D., Conroy, M. L., McGreevy, B. P., & Barraclough, D. J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Cognitive Brain Research*, *22*(1), 45–58.

Lopes, L. L., & Oden, G. C. (1987). Distinguishing between random and nonrandom events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 392–400.

Neuringer, A. (1986). Can people behave "randomly?": The role of feedback. *Journal of Experimental Psychology: General*, *115*(1), 62–75.

Rapoport, A., & Budescu, D. V. (1992). Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology: General*, *121*(3), 352–363.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others the consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, *7*(4), 341–351.

West, R. L., & Lebiere, C. (2001). Simple games as dynamic, coupled systems: randomness and other emergent properties. *Cognitive Systems Research*, *1*(4), 221–239.