Modeling the dynamics of suspense

Zhi-Wei Li, Neil R. Bramley and Todd M. Gureckis
Center for Neural Science and Psychology, New York University

Question
What makes some games more engaging than others?
What makes certain moments in a game, movie, or story more engaging than others?

Theoretical background
Ely et al (2015) proposed:

- Engagement is driven by suspense and surprise.
- Suspense and surprise are derived from the dynamics of belief change. They formalized suspense as:
  \[ \text{Suspense} = E_x [(\mu_{t+1} - \mu_t)^2] = \sum_p p(s) (\mu_{t+1} - \mu_s)^2 \]
  \( \mu_t \): belief of outcome \( s \): what happens at the next step
- Alternative interpretation: suspense = expected surprise for the next step.
- They also derived guidance on how to design a game rule to maximize suspense.
- What’s lacking: empirical evidence

The following models are based on Ely et al’s formulation, where the belief update is bayesian.

Experiment 1
Task: Watching beach volleyball games

Subjects are randomly assigned to either condition:
- Rule 1 "consecutive 4": whichever team wins 4 pts consecutively will win this round.
- Rule 2 “first 4 w. advantage point”: whichever team wins 4 points will win this round. Need to win by 2 if tied at the 3-3.

Suspense model predictions:
Four sequences showing point-by-point differences in predicted suspense depending on the game rules

Suspense self reports:

- Betting on the outcome increases overall suspense

Summary
- We see large individual differences in suspense reports.
- Aggregate results are noisy and only weakly consistent with theory.
- Don’t have enough data since watching volleyball games takes time.

Experiment 2 (ongoing)
Task: Playing card games

Subjects will experience both rules (allowing within-subject comparison).
- Each round maximum 3 cards to be drawn.
- Rule 1 "stay-below N": whenever your current total reaches N, you lose (even if you haven’t finished drawing 3 cards). N=7 below.
- Rule 2 “end-below N": As long as the final total of 3 cards is smaller than N, you win. N=5 below.

Q1. How good does our model describes the data overall?
We choose 14 game sequences, among which 7 are under rule 1 (green background) and 7 under rule 2 (blue). The order of the two rules are counterbalanced among subjects.

Q2. Does the rule manipulation work? (within-subject suspense difference)
Two “critical sequences” are selected to maximize the predicted suspense difference (y axis) under 2 rules.

Q3. Does suspense predict engagement?
We asked subjects to rate their engagement at the end of each game.

Future directions
- Alternative models:
  - Heuristics for simpler forward prediction, i.e. 1-step ahead only, attending to high value cards.
  - Modeling the inter-subject difference of using the scale?
- Better experiment design and data collection:
  - Explore other indirect measures of suspense to minimise distraction and self report noise, e.g.: willingness to pay, forced choice between games, galvanic skin response, continuous response with suspense dial
  - Other kinds of easily manipulable yet interesting stimuli?