Beyond uncertainty and information bonuses: Fun and empowerment

Eric Schulz
MPI for Biological Cybernetics

Traditional view

Traditional view

Active learning

- Calculate current entropy
- Calculate entropy reduction for different outcomes
- Calculate expectation of these outcomes
- Query expected reduction
- What measure of entropy and how to calculate expectation?

Traditional view

Active learning

- Calculate current entropy
- Calculate entropy reduction for different outcomes
- Calculate expectation of these outcomes
- Query expected reduction
- What measure of entropy and how to calculate expectation?

Exploration

- Calculate mean and variance
- Put means and variances into a utility function
- Treat an option's uncertainty positively
- Choose highest utility
- How do expected values and uncertainties integrate?

Traditional view

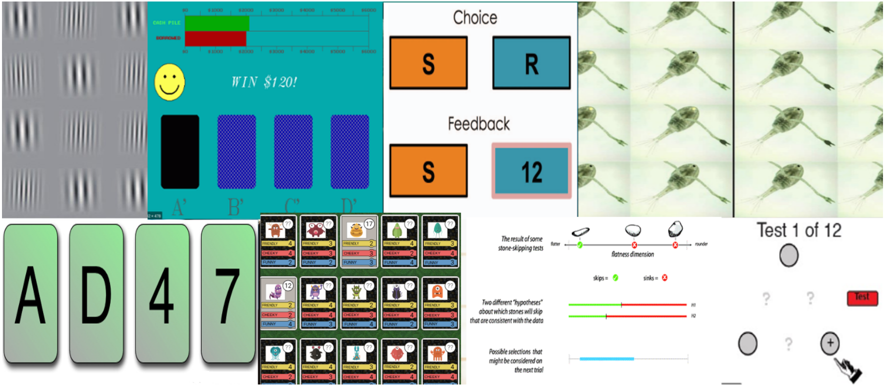


Figure: Paradigms that have been traditionally used.

Imagined future

Model improvement as fun

- Assess current world model
- Calculate what would lead to best improvement
- Model improvement is learning progress
- Learning progress is fun
- What influences how much fun people have?

Exploration as empowerment

- Assess current world model
- Calculate what could enable further exploration
- Options that enable further exploration are empowering
- Maximize empowerment
- How to test empowerment in humans?

Imagined future

Figure: Paradigms that we would like to study (things that Franziska finds fun).

Fun as model improvement

Model improvement as fun

- $r_t = g(r_{external}, r_{internal})$
- $r_{internal} = q(M_{t-1}, \mathcal{D}_{1:t}) - q(M_t, \mathcal{D}_{1:t})$
- Learning progress is maximized when a task is medium complex
- Fun could depend on both internal and external rewards
- Related to curiosity, proximal development, and Goldilocks effect

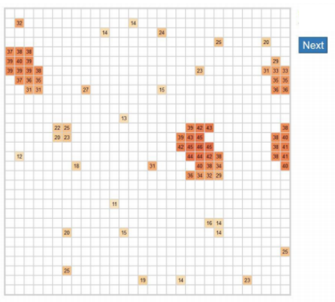
Jürgen Schmidhuber

Grid Search

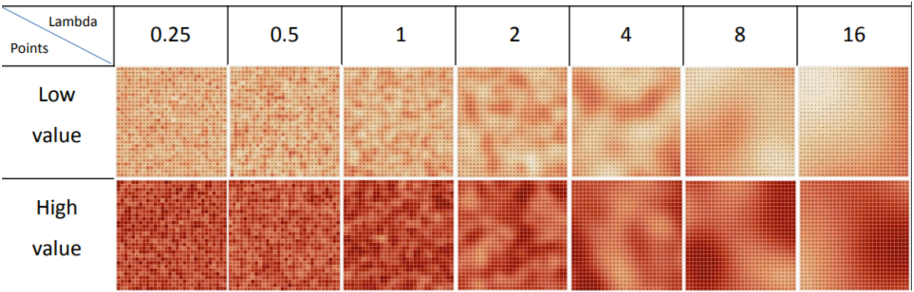
Description

- Spatially correlated multi-armed bandit paradigm (Wu et al. 2018)
- Free exploration for 10 minutes
- Explore a grid freely and click “Next” to go to next grid
- No monetary reward for performance
- Measure of fun: clicks per grid

Interface



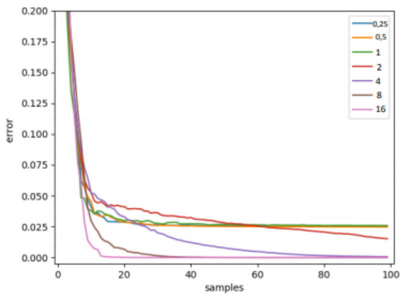
Grid Search



- Different levels of smoothness
- Different average of overall values

Grid Search

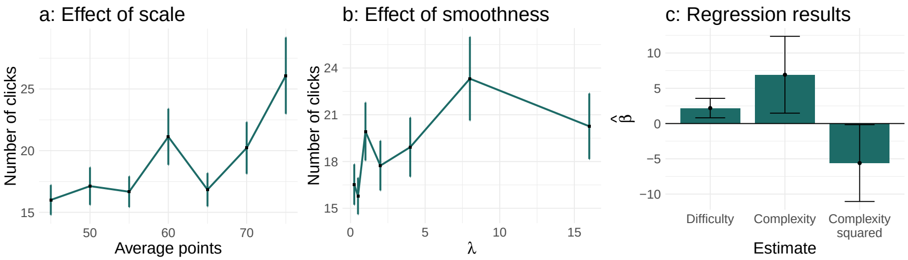
Description



Interface

- Smother functions are easier to learn
- Error for smoother function decays faster
- Inverted u-shape prediction of fun

Grid Search



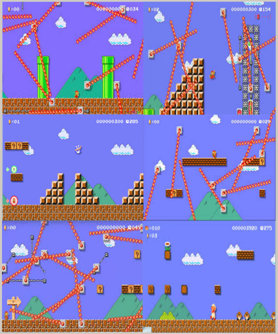
- Higher average makes grids more fun
- Intermediate smoothness is played the longest
- Linear effect of difficulty, quadratic effect of complexity

Super Mario Maker

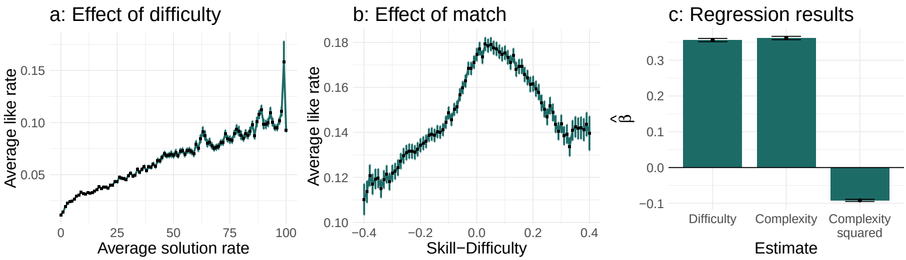
Description

- Video game which allows players to create their own Super Mario levels
- Players can publish their courses for other players to experience, who can play and evaluate the games
- 851,824 players playing 115,032 player-designed levels over 630,116,357 plays in total
- Attempts, completions, and likes

Interface



Super Mario Maker



- Easier levels are liked more
- Levels that can just about be solved are liked more
- Linear effect of difficulty, quadratic effect of complexity

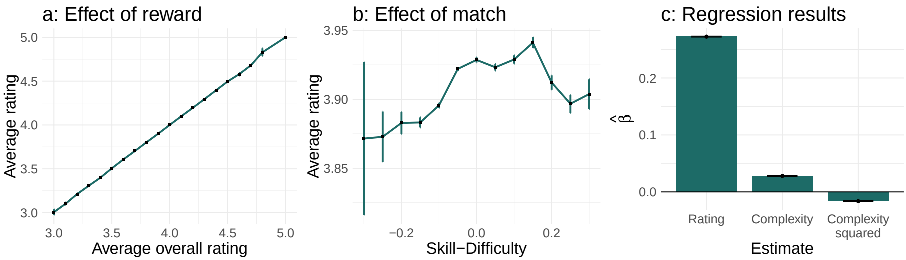
Goodreads

Description

- Social website that allows individuals to search a database of books, annotations, quotes, and reviews
- Users can mark books they want to read, have read, and rate them
- 5,976,479 ratings from 53,424 users evaluating 10,000 books
- Books marked as “want to read”, “have read”, and ratings

Interface

Goodreads



- Average overall rating predicts reader's rating
- Books that match skill level are liked most
- Linear effect of rating, quadratic effect of complexity

Exploration as empowerment

Explore to explore more

- Explore options that can generate more options
- Learning progress is maximize when a task is medium complex
- Fun could depend on both internal and external rewards
- Related to learning progress and minimum surprise exploration

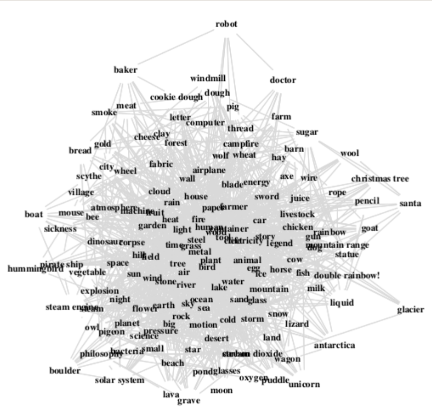
Laura Schulz

Little Alchemy

Screenshot

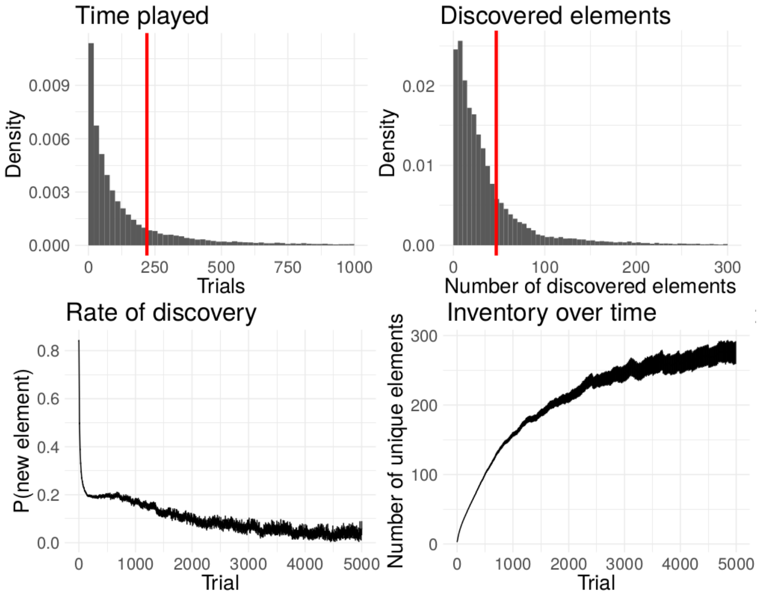
- Start with 4 base elements: water, fire, earth, air
- Combine elements to create new ones

Network of elements

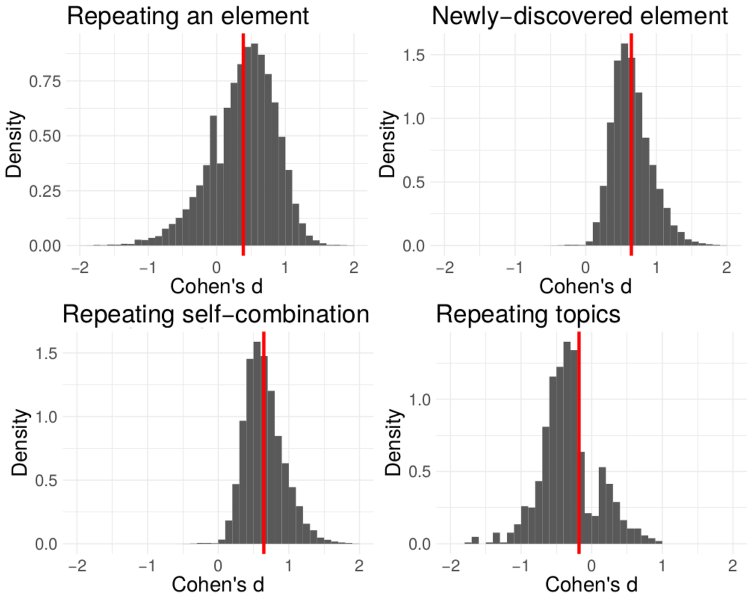


Creating life

Online data set (N=30,415)



Behavioral results



Tiny Alchemy

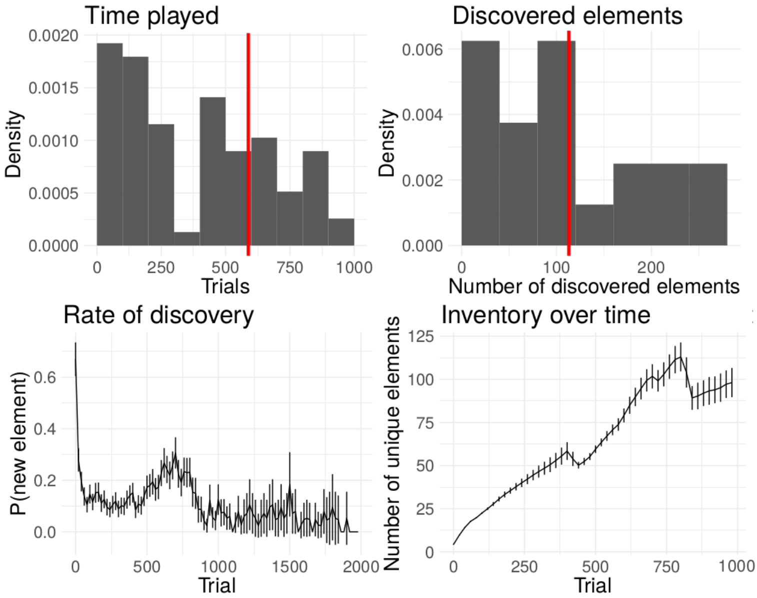
Screenshot

- Javascript version of Little Alchemy
- 520 elements based on Little Alchemy 1 game tree

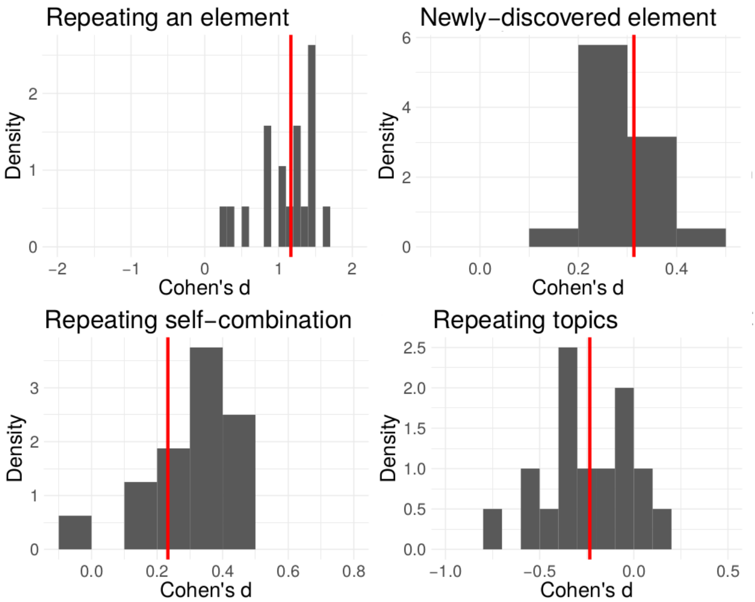
Design

- Participants recruited from MTurk
- $0.10 per discovered item
- Can give up whenever they want
- Maximum play time: 5 hours
- Top rated HIT on turkprime for two days

Mturk data set (N=100)



Behavioral results



Tiny Fractals

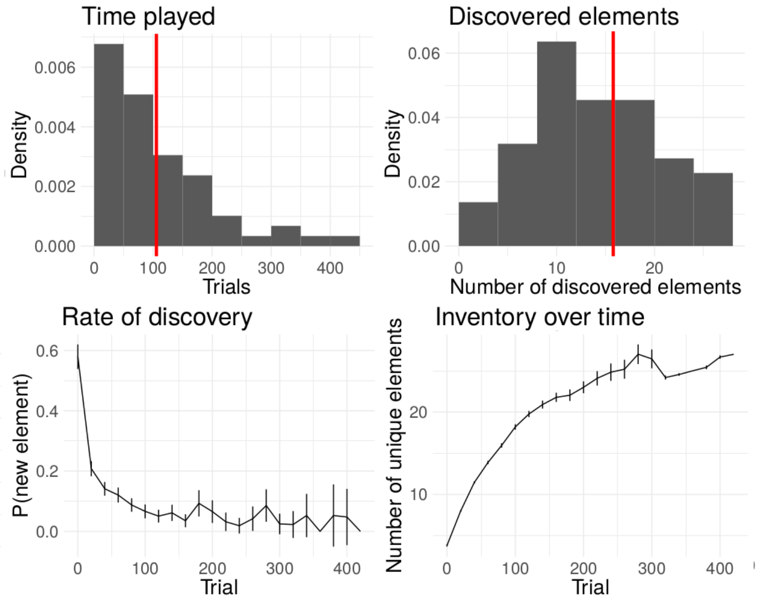
Screenshot

- Javascript version of Little Alchemy
- Elements have been randomly pixelated

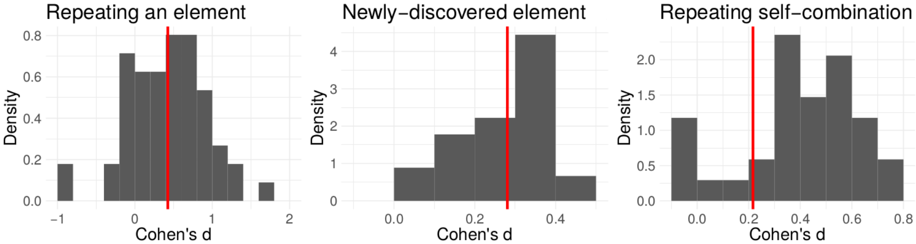
Design

- 100 participants recruited from MTurk
- $0.10 per discovered item
- Can give up whenever they want
- Maximum play time: 5 hours
- Used to strip away the semantics of the original game

Mturk data set (N=98)



Behavioral results

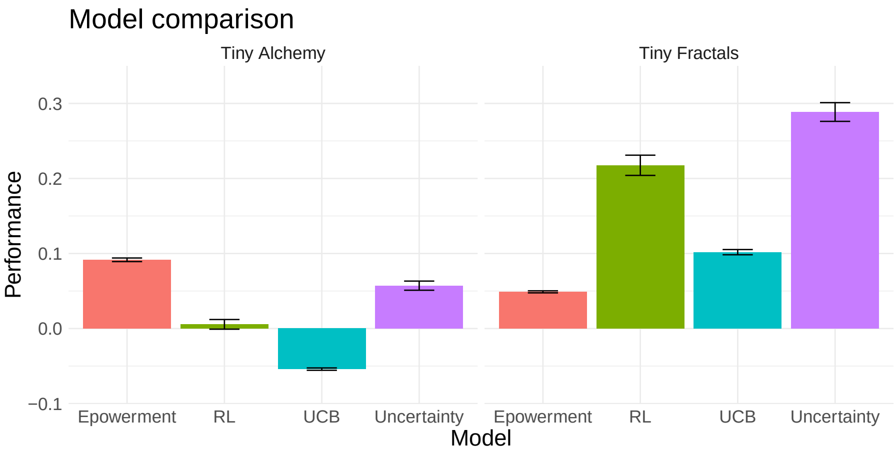


Note: Repeating topics is impossible here.

Model comparison

- Simple RL:
  - If element was part of a successful combination: +1
  - If element was part of a successful combination: -1
  - Utility is sum of two elements' scores
- Uncertainty-guided exploration:
  - Count-based exploration; elements will be picked that have been less often chosen in the past
- Mean and uncertainty:
  - Count-based exploration and past successes and failures
- Empowerment:
  - Utility is how combinable resulting element is with given the inventory
  - Based on actual game tree
  - We are working on other approximations

Model comparison results



Conclusion

Conclusion

- Enriching paradigms and models of human exploration

Conclusion

- Enriching paradigms and models of human exploration
- Fun as learning progress and rate of rewards

Conclusion

- Enriching paradigms and models of human exploration
- Fun as learning progress and rate of rewards
- Empowerment as exploration to explore more

Conclusion

- Enriching paradigms and models of human exploration
- Fun as learning progress and rate of rewards
- Empowerment as exploration to explore more
- Study games and exploration in real world scenarios

Thank you.

Franziska Brändle

Lena Stocks

Charley Wu

Samuel Gershman

Joshua Tenenbaum

- Max Planck Society
- Jacobs Foundation
- The workshop organizers