Learned Impatience?
Dispersed Reinforcement and Time Discounting

David Poensgen (Goethe University Frankfurt)

February 22, 2019

Sloan-Nomis Workshop on the Cognitive Foundations of Economic Behavior

Motivation

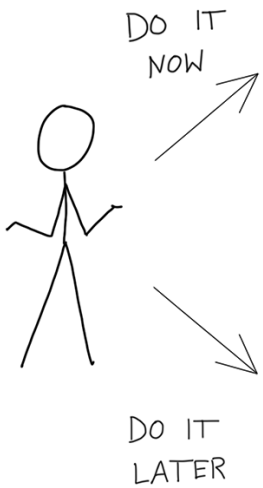
1. Individuals learn from consequences of past actions.
2. Actions often have a series of consequences: some follow soon, some later.
3. How does this ordering affect learning?
   Plausibly: Easiest to learn from soonest consequences.
4. Then: Immediate consequences will be over-weighted.
   Behavior biased towards impatience.

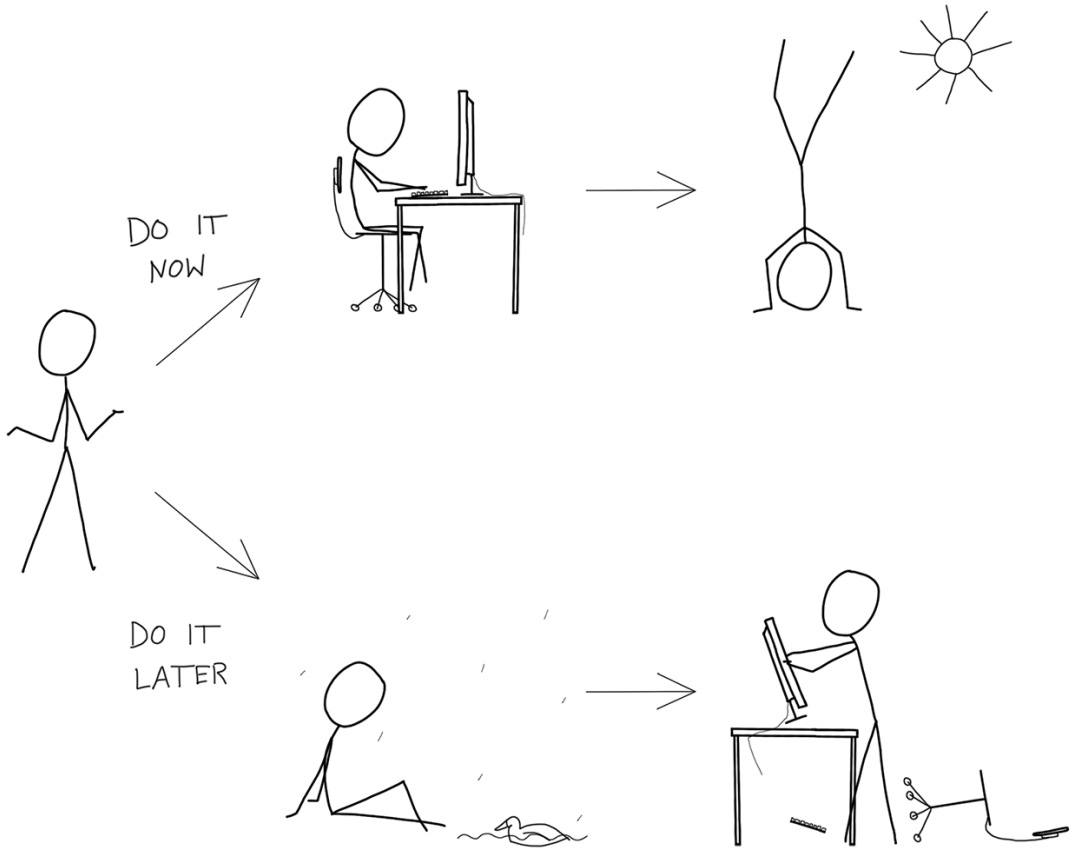
May help explain why myopic behavior is so widespread and persistent.

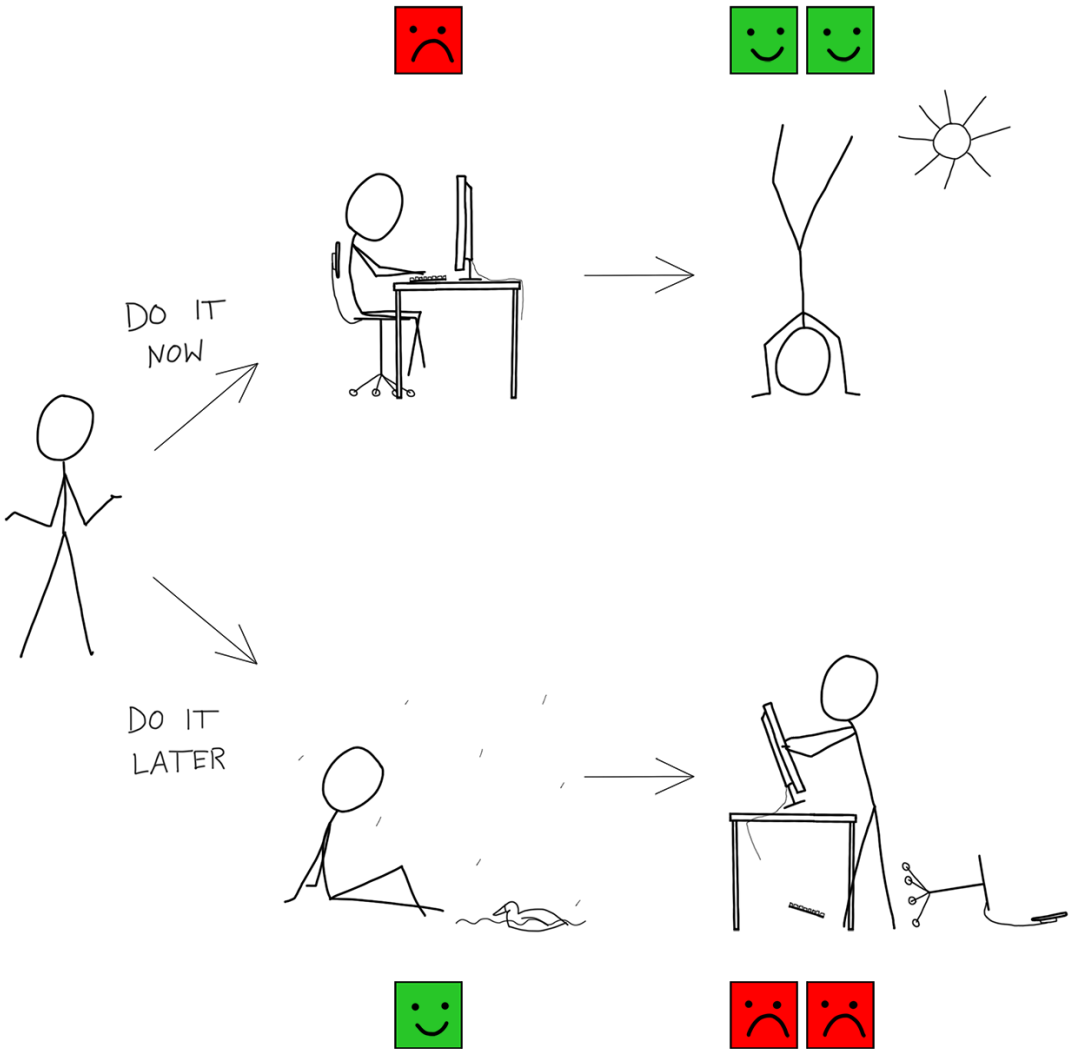
Motivation

1. Individuals learn from consequences of past actions.
2. Actions often have a series of consequences: some follow soon, some later.
3. How does this ordering affect learning?
   Plausibly: Easiest to learn from soonest consequences.
4. Then: Immediate consequences will be over-weighted.
   Behavior biased towards impatience.

May help explain why myopic behavior is so widespread and persistent.


DO IT
NOW
DO IT
LATER


DO IT
NOW
DO IT
LATER


DO IT NOW
DO IT LATER

Motivation

1. Individuals learn from consequences of past actions.
2. Actions often have a series of consequences: some follow soon, some later.
3. How does this ordering affect learning?
   Plausibly: Easiest to learn from soonest consequences.
4. Then: Immediate consequences will be over-weighted.
   Behavior biased towards impatience.

May help explain why myopic behavior is so widespread and persistent.

Motivation

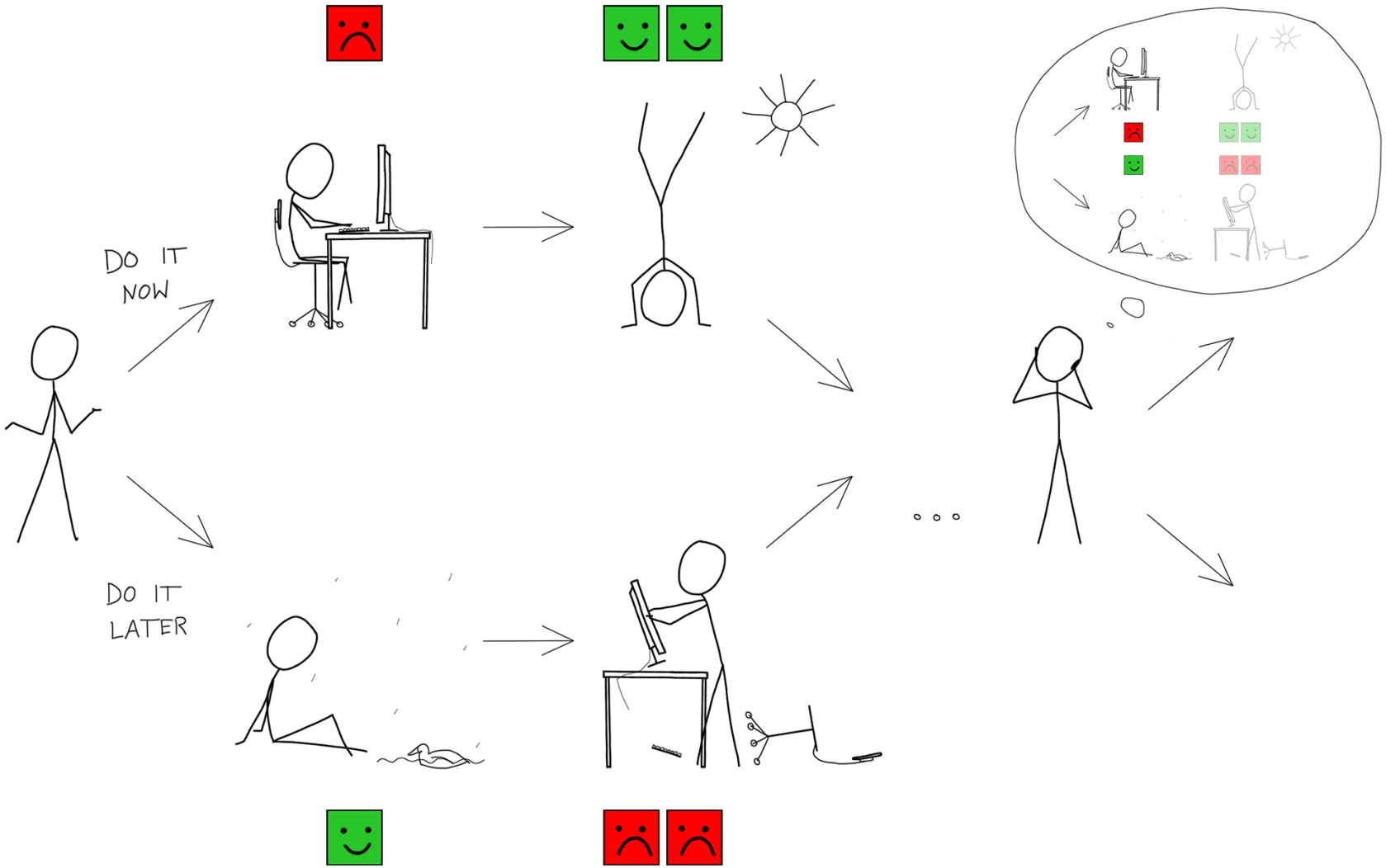
1. Individuals learn from consequences of past actions.
2. Actions often have a series of consequences: some follow soon, some later.
3. How does this ordering affect learning?
   Plausibly: Easiest to learn from soonest consequences.
4. Then: Immediate consequences will be over-weighted.
   Behavior biased towards impatience.

May help explain why myopic behavior is so widespread and persistent.

Motivation

1. Individuals learn from consequences of past actions.
2. Actions often have a series of consequences: some follow soon, some later.
3. How does this ordering affect learning?
   Plausibly: Easiest to learn from soonest consequences.
4. Then: Immediate consequences will be over-weighted.
   Behavior biased towards impatience.

May help explain why myopic behavior is so widespread and persistent.


DO IT NOW
DO IT LATER

Motivation

1. Individuals learn from consequences of past actions.
2. Actions often have a series of consequences: some follow soon, some later.
3. How does this ordering affect learning?
   Plausibly: Easiest to learn from soonest consequences.
4. Then: Immediate consequences will be over-weighted.
   Behavior biased towards impatience.

May help explain why myopic behavior is so widespread and persistent.

Background

- Decreasing effectiveness of reinforcement with delay (e.g. Mazur 2002).
  - Typically not connected to time discounting, but speed of learning.
- Explained via accumulation of noise by Commons, Woodford et al. (1982, 1991).
- Feedback delay modulates neural circuitries involved in learning (Foerde/Shohamy 2011, Foerde et al. 2013, Arbel et al. 2017).
  - Associative learning tasks; singular feedback. Performance not affected.

- Gabaix & Laibson (2017) also link time discounting and information frictions.
  - Formally applicable here; different interpretation on source of noise.

- Melioration theory: Behavior guided by immediate, not overall reinforcement rate (Herrnstein et al.).
  - Important experimental paradigm: "Harvard game" (Review: Prelec 2014).
  - Critique by Sims et al. (2013): Bayesian algorithms need 1000s of trials for solution. Melioration as rational response to task complexity.

Design: Overview

- 6 abstract options (= colors): $\{$  ,  ,  ,  ,  ,  $\}$
- Subjects faced with sequence of 105 binary choices.
- Payoff and feedback mechanism:
  - Each color x associated with a payoff vector (x_1, x_2)
  - Values initially unknown, but can be learned.
  - Choosing x has 2 consequences:

    $x_1 + \epsilon$ points shown and awarded immediately.
    $x_2 + \epsilon'$ points shown and awarded with one round delay.

  - ϵ, ϵ' are disturbances drawn uniformly from $\{1, 2, 3, 4\}$.
  - Total value of x is $x_1 + x_2$
- Goal: Collect as many points as possible.
  - All points rewarded simultaneously after the experiment.
- All rules and mechanisms clearly communicated to subjects.

Design: Overview

- 6 abstract options (= colors): $\{$ ■ , ■ , ■ , ■ , ■ , ■ $\}$
- Subjects faced with sequence of 105 binary choices.

- Payoff and feedback mechanism:
  - Each color x associated with a payoff vector (x_1, x_2)
  - Values initially unknown, but can be learned.
  - Choosing x has 2 consequences:

    $x_1 + \epsilon$ points shown and awarded immediately.
    $x_2 + \epsilon'$ points shown and awarded with one round delay.

  - ϵ, ϵ' are disturbances drawn uniformly from $\{1, 2, 3, 4\}$.
  - Total value of x is $x_1 + x_2$
- Goal: Collect as many points as possible.
  - All points rewarded simultaneously after the experiment.
- All rules and mechanisms clearly communicated to subjects.

Design: Overview

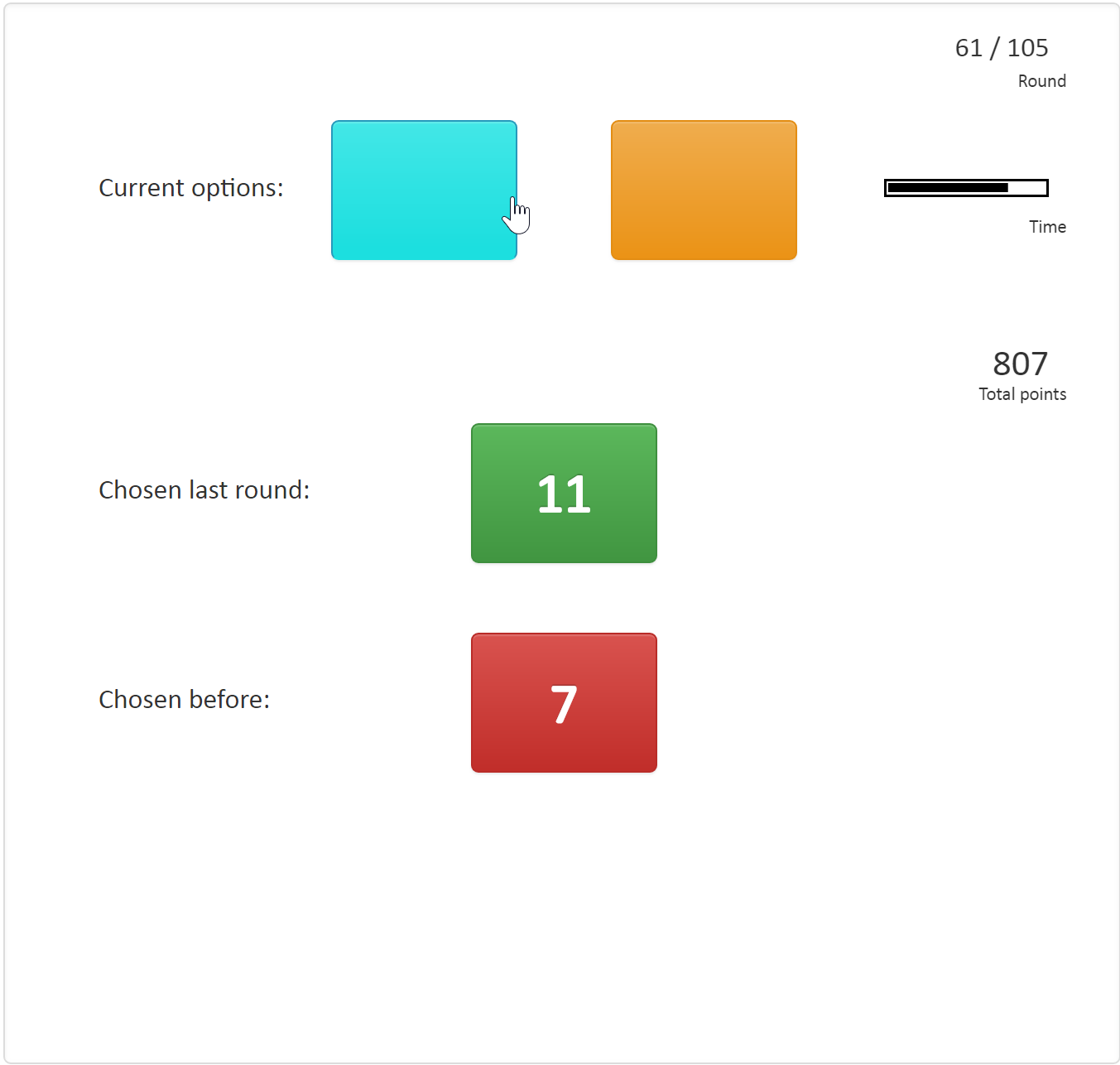
- 6 abstract options (= colors): $\{$  ,  ,  ,  ,  ,  $\}$
- Subjects faced with sequence of 105 binary choices.

- Payoff and feedback mechanism:
  - Each color x associated with a payoff vector (x_1, x_2)
  - Values initially unknown, but can be learned.
  - Choosing x has 2 consequences:

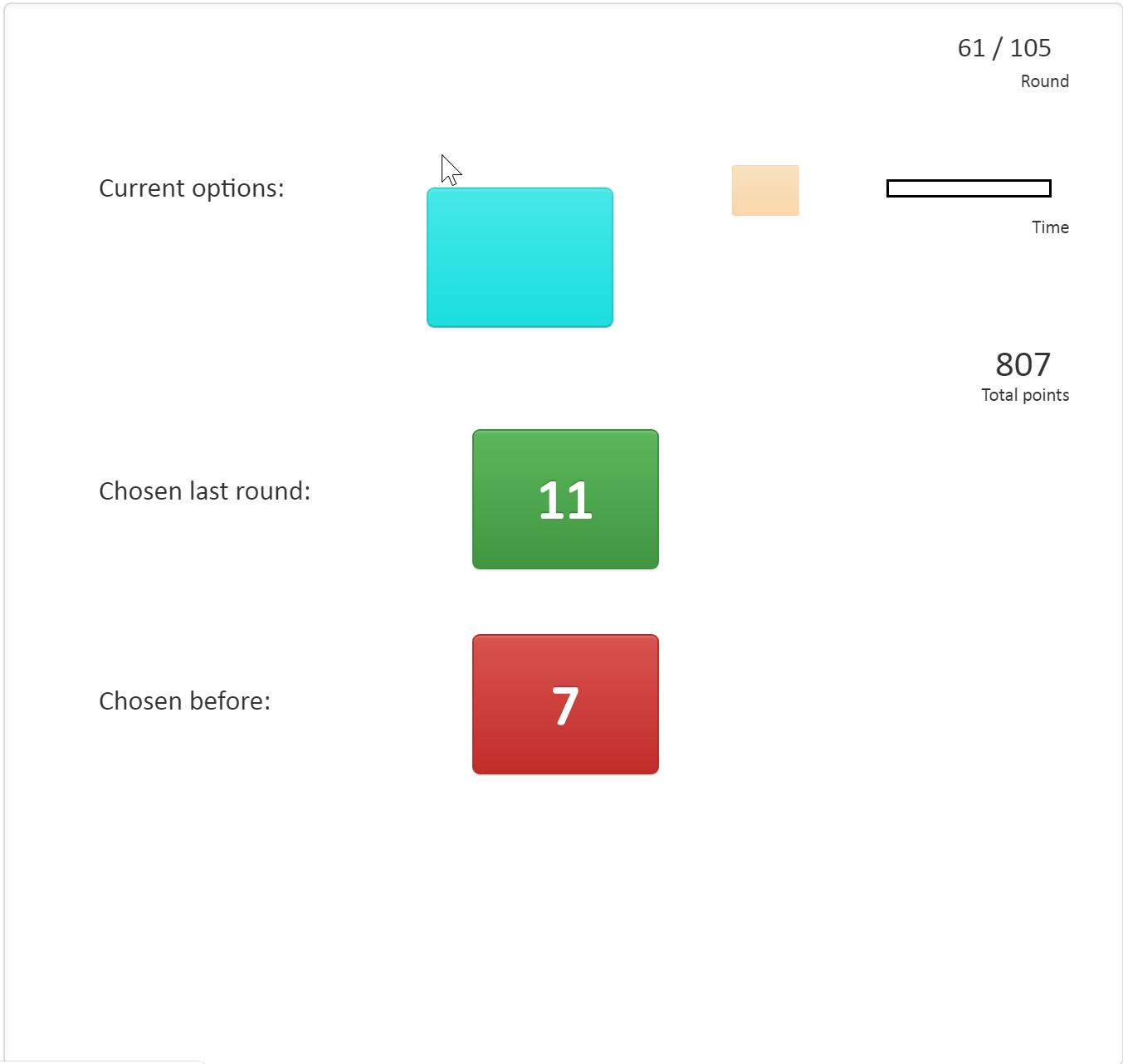
    $x_1 + \epsilon$ points shown and awarded immediately.
    $x_2 + \epsilon'$ points shown and awarded with one round delay.

  - ϵ, ϵ' are disturbances drawn uniformly from $\{1, 2, 3, 4\}$.
  - Total value of x is $x_1 + x_2$
- Goal: Collect as many points as possible.
  - All points rewarded simultaneously after the experiment.
- All rules and mechanisms clearly communicated to subjects.

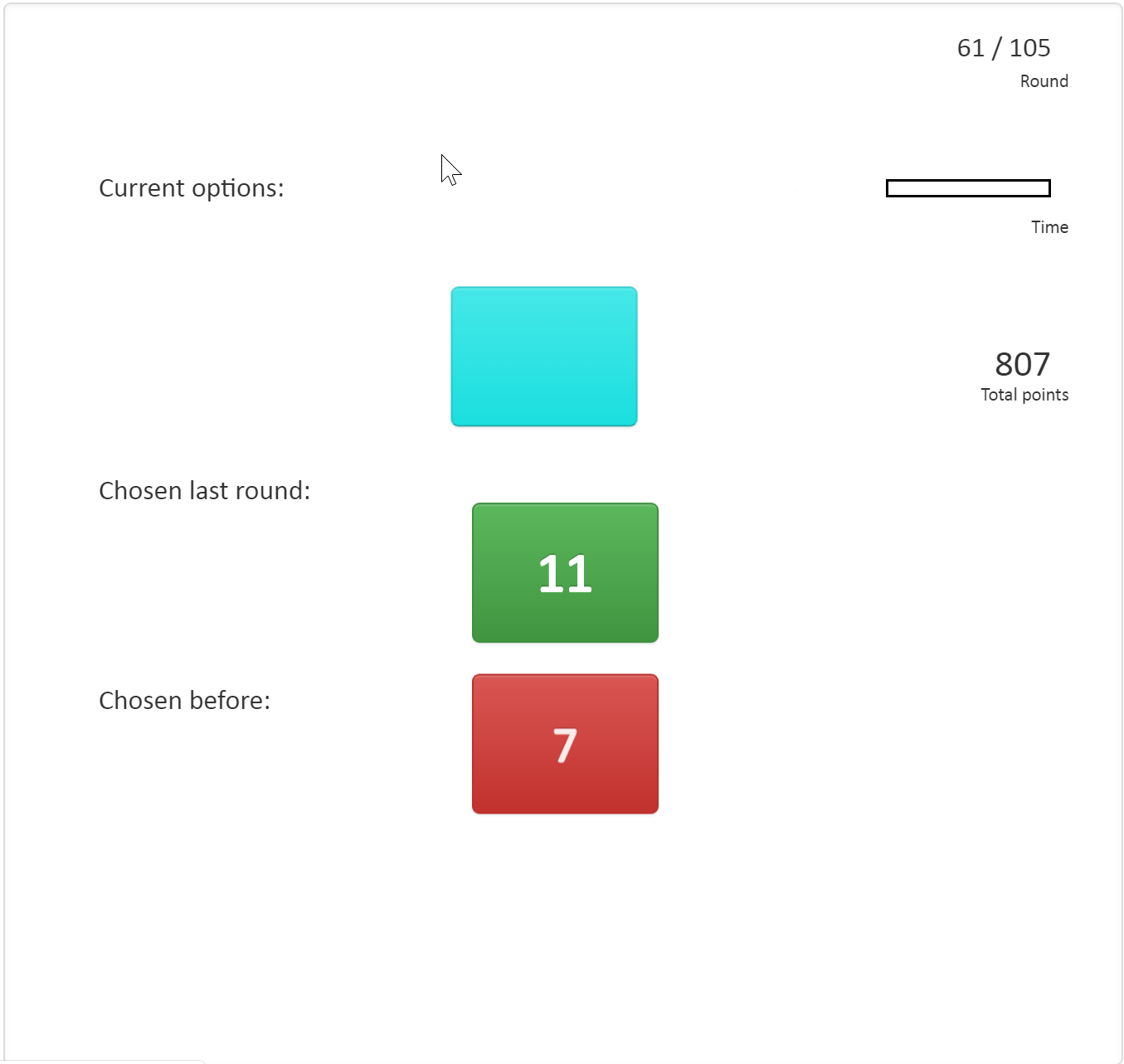
Design: Example Screen



Design: Example Screen



Design: Example Screen



Design: Example Screen

61 / 105
Round

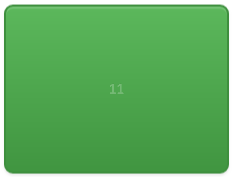
Current options:

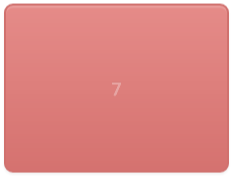
Time

807
Total points

Chosen last round:

Chosen before:





Design: Example Screen

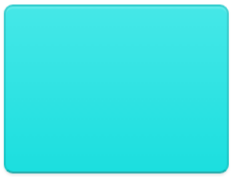
61 / 105
Round

Current options:

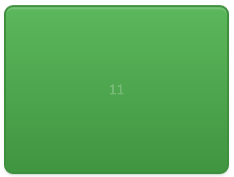
Time

807
Total points

Chosen last round:



Chosen before:



Design: Example Screen

61 / 105
Round

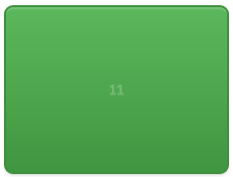
Current options:

Time

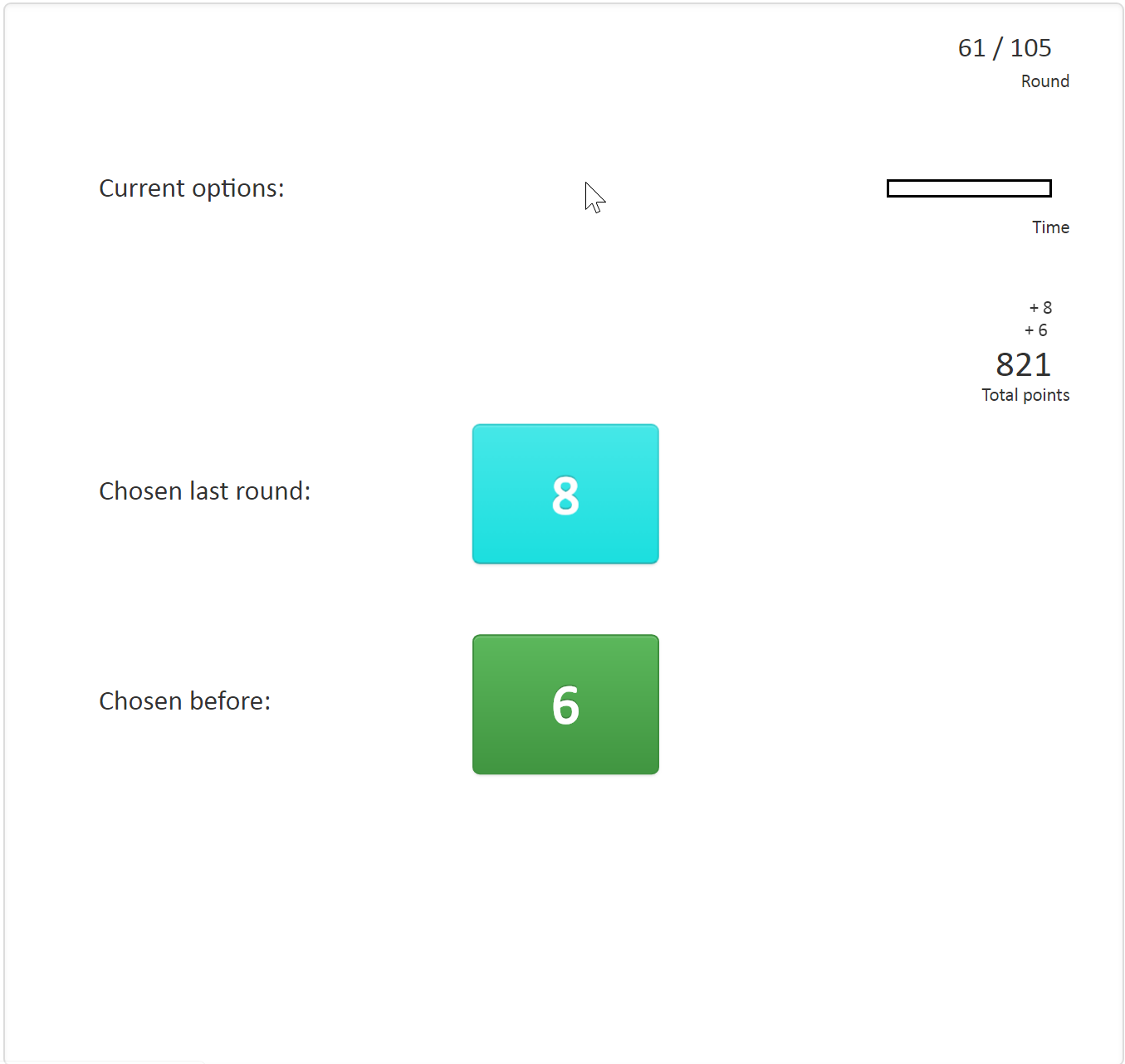
807
Total points

Chosen last round:

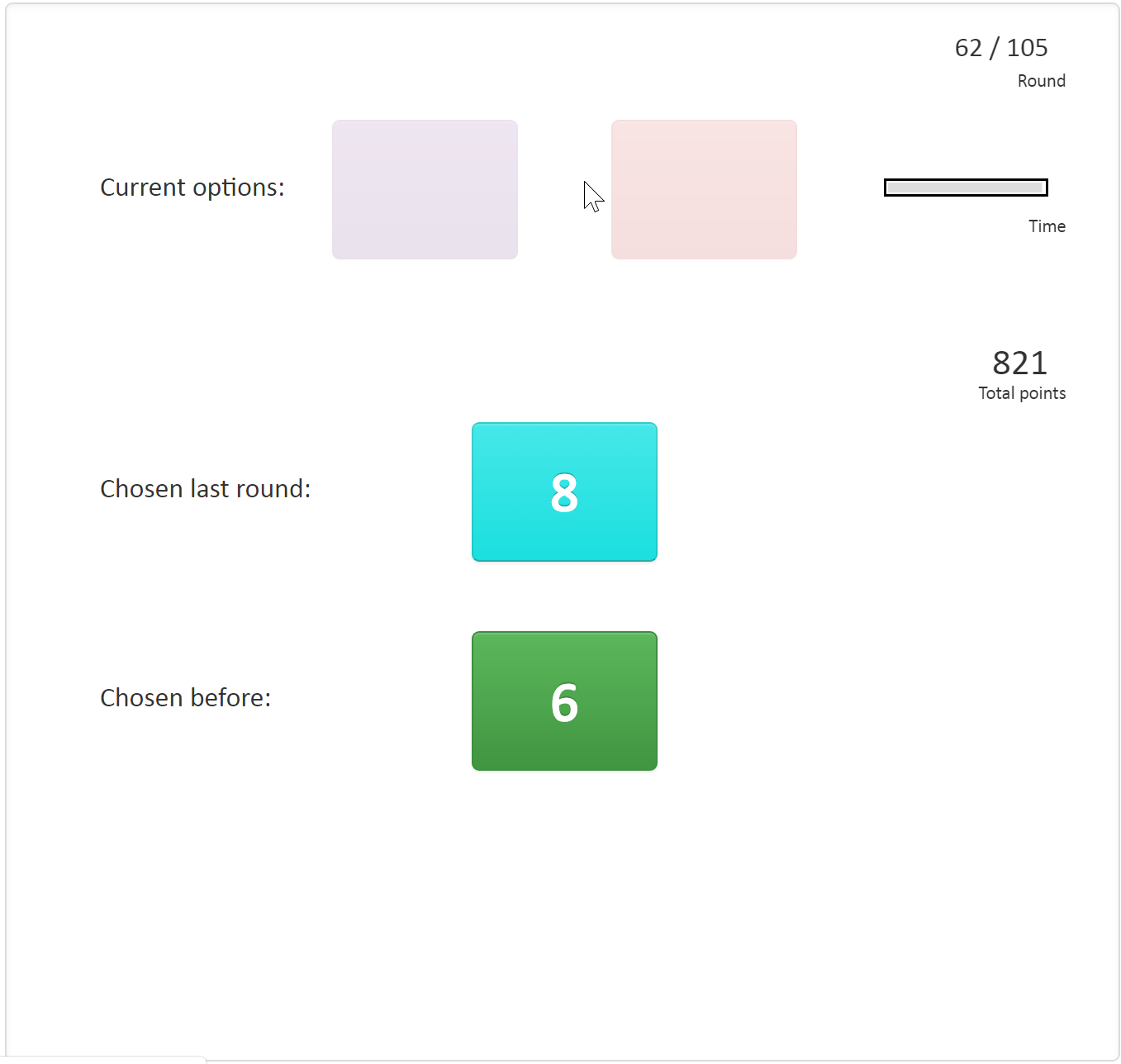
Chosen before:



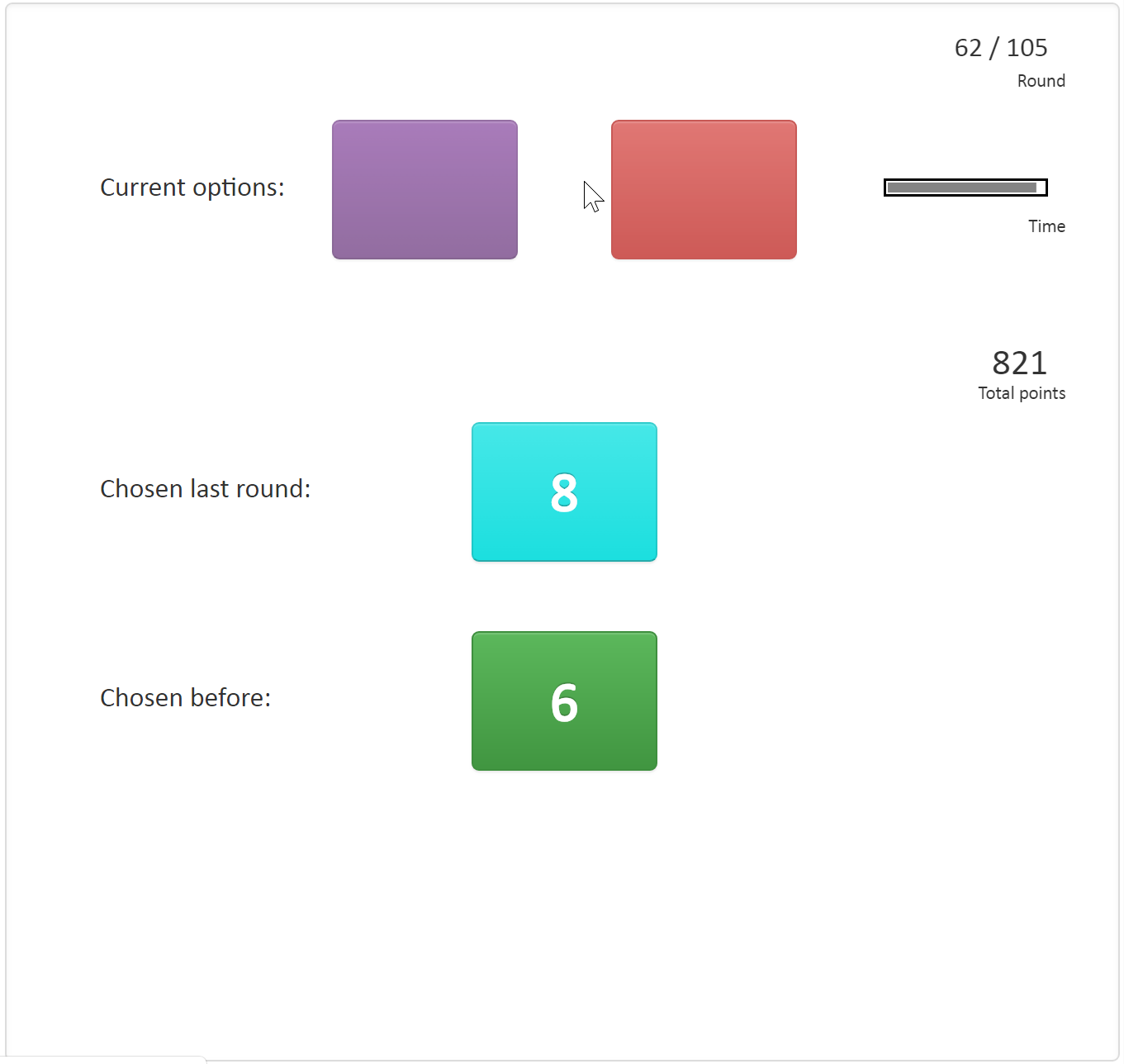
Design: Example Screen



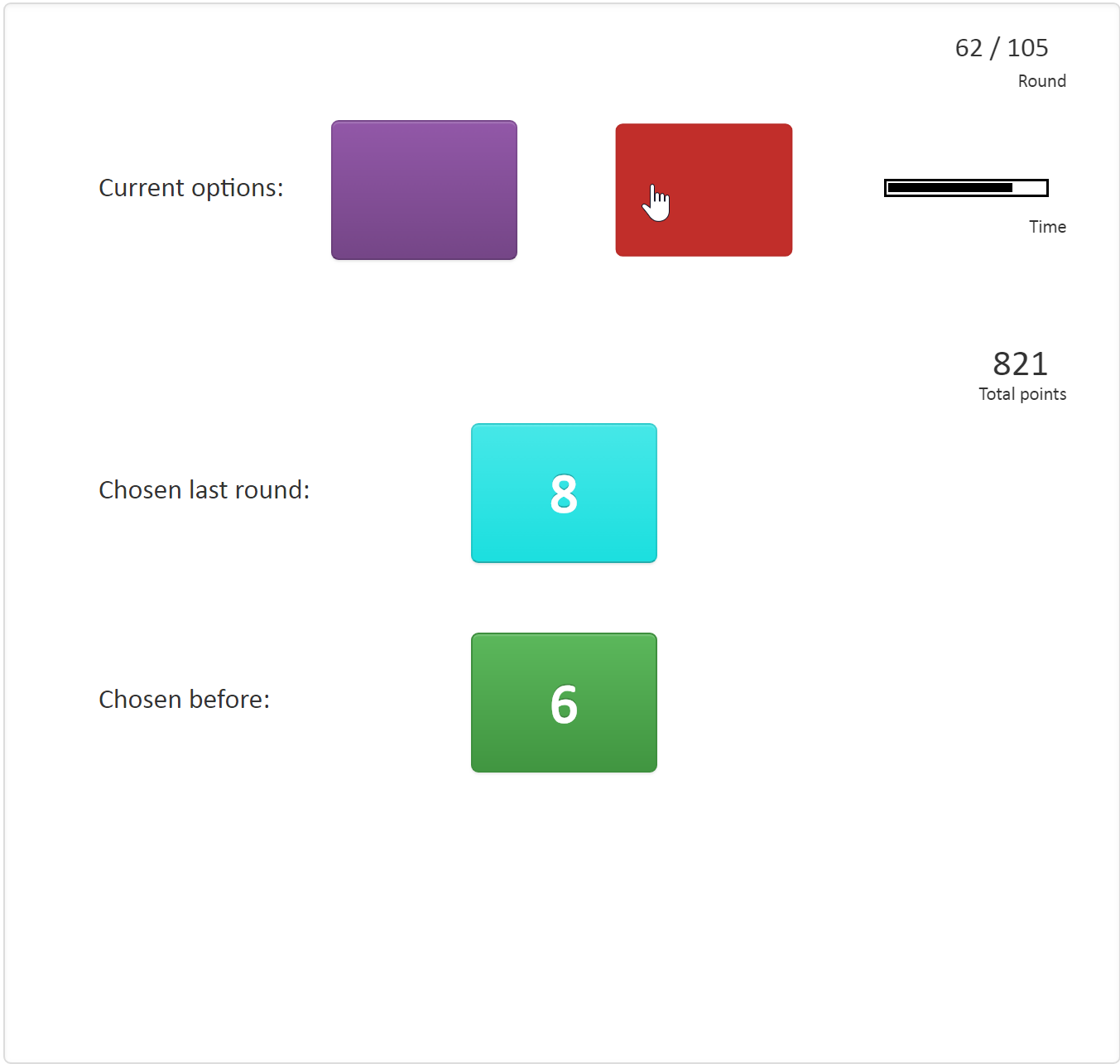
Design: Example Screen



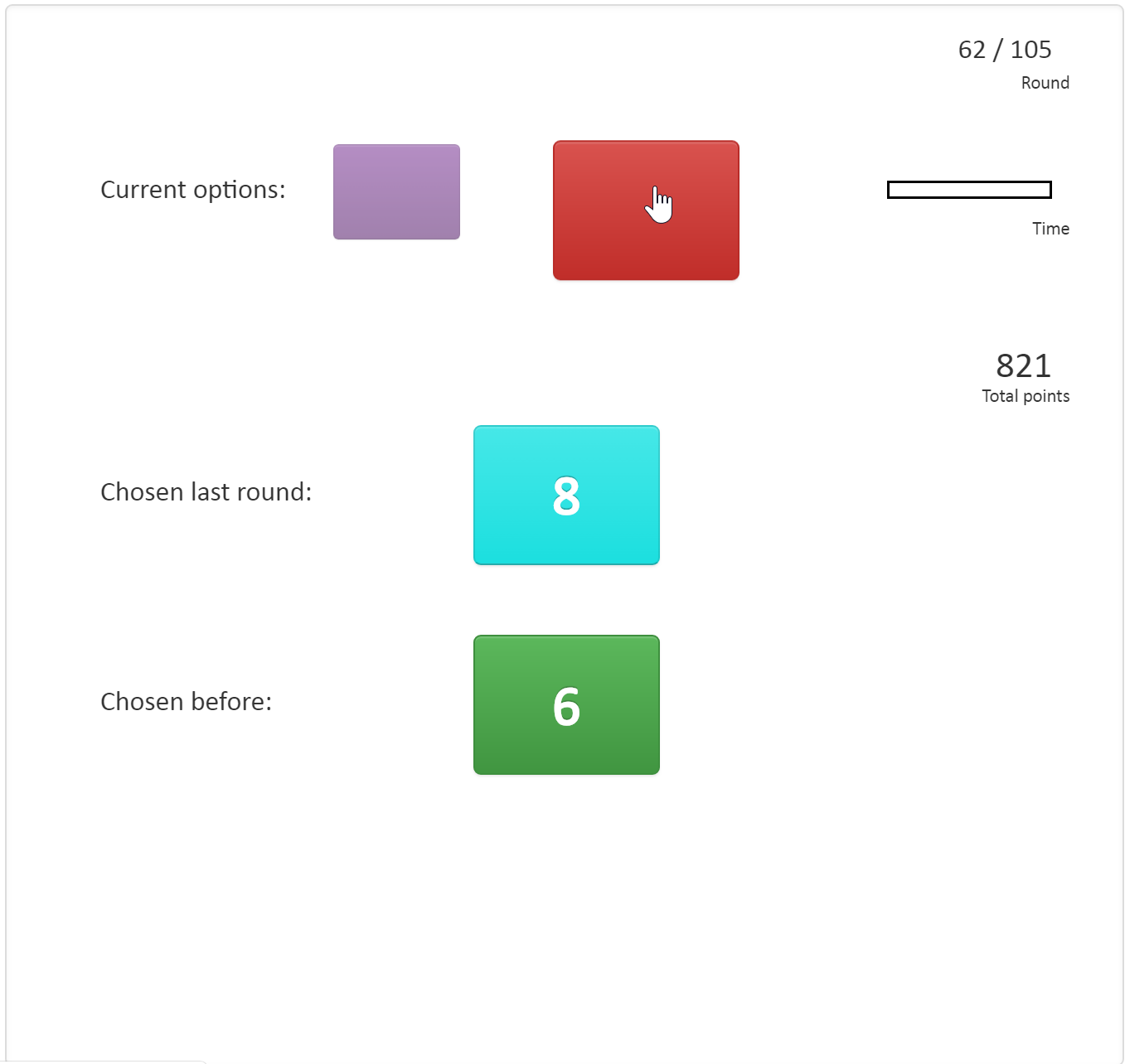
Design: Example Screen



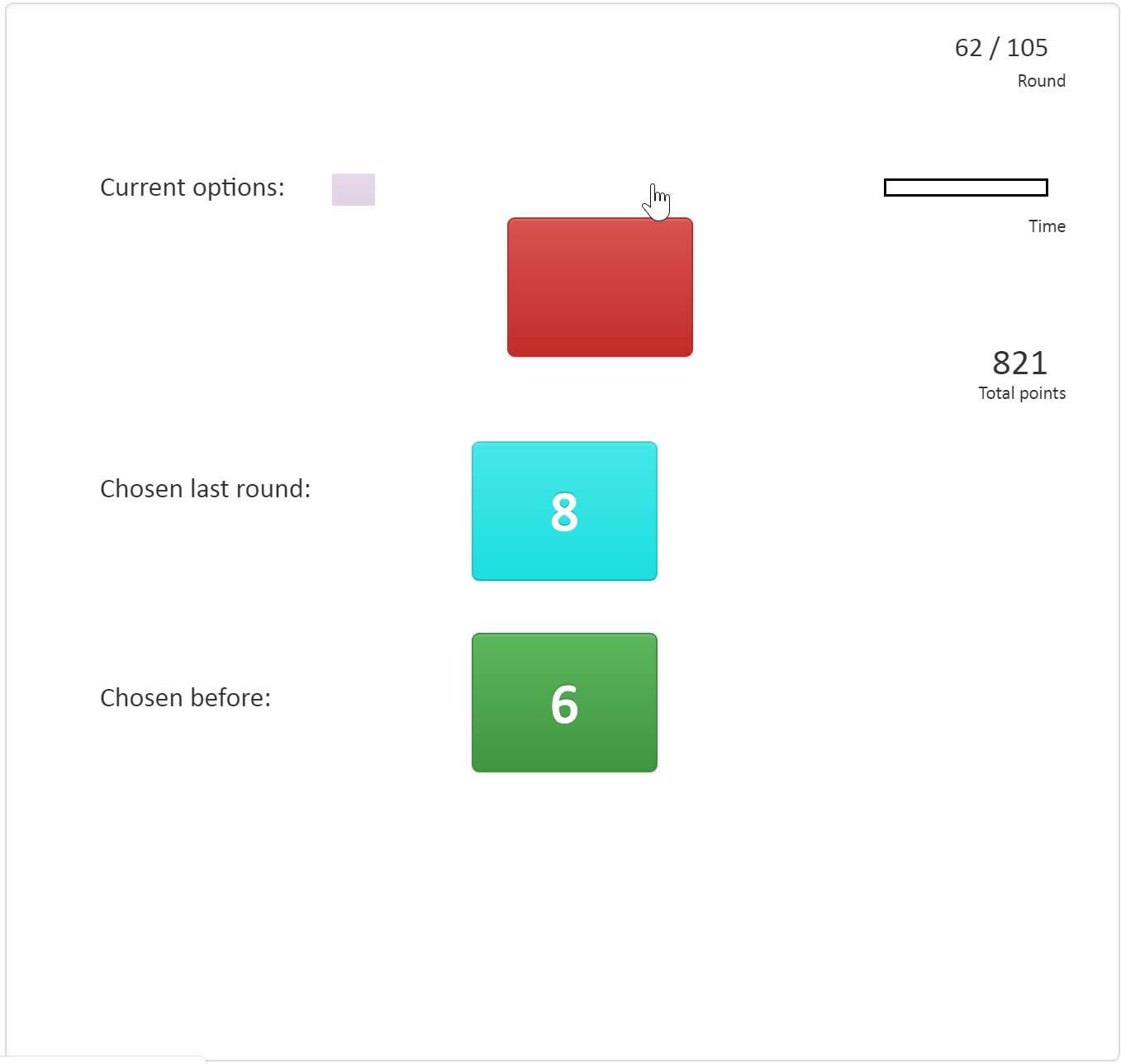
Design: Example Screen



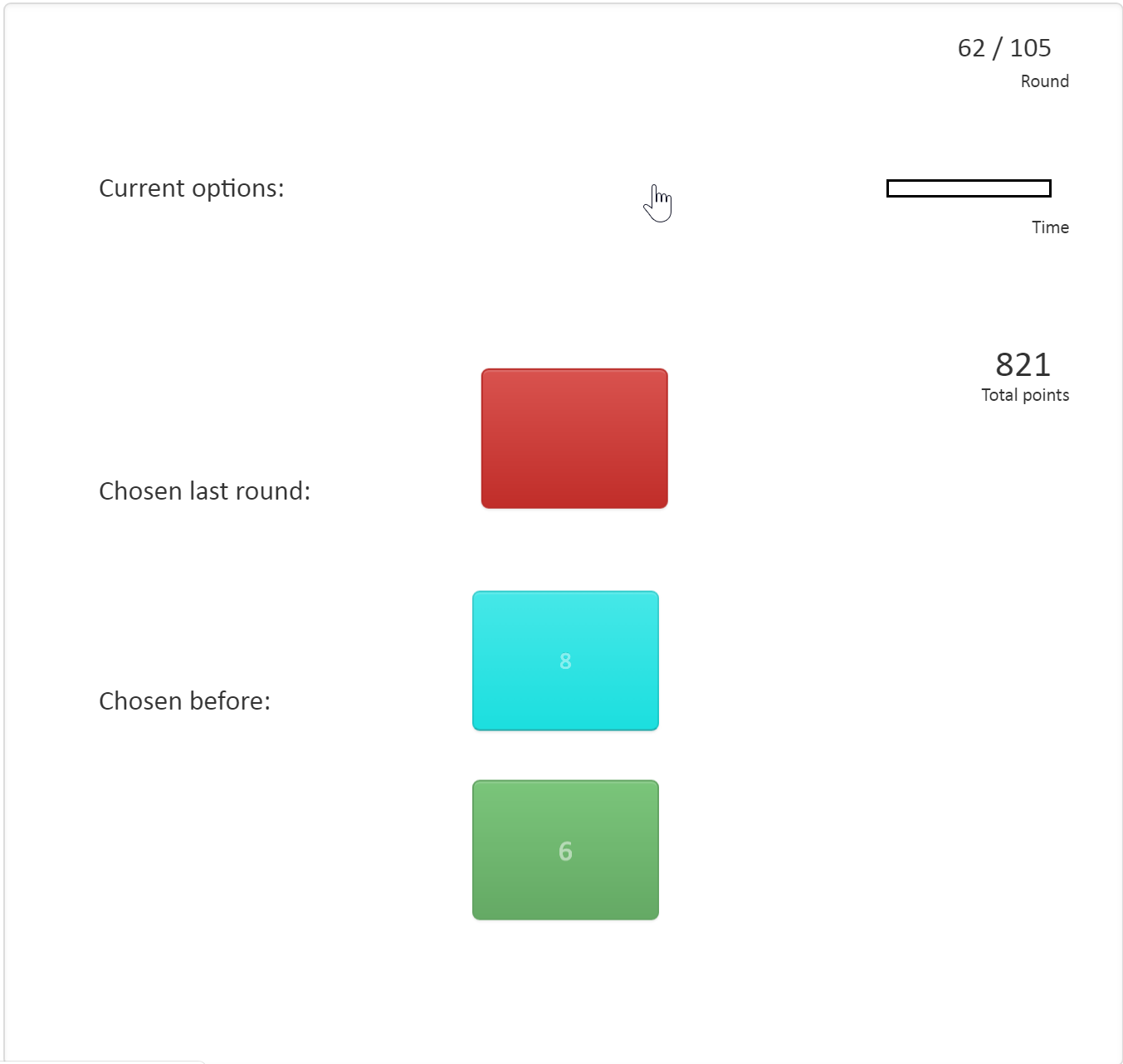
Design: Example Screen



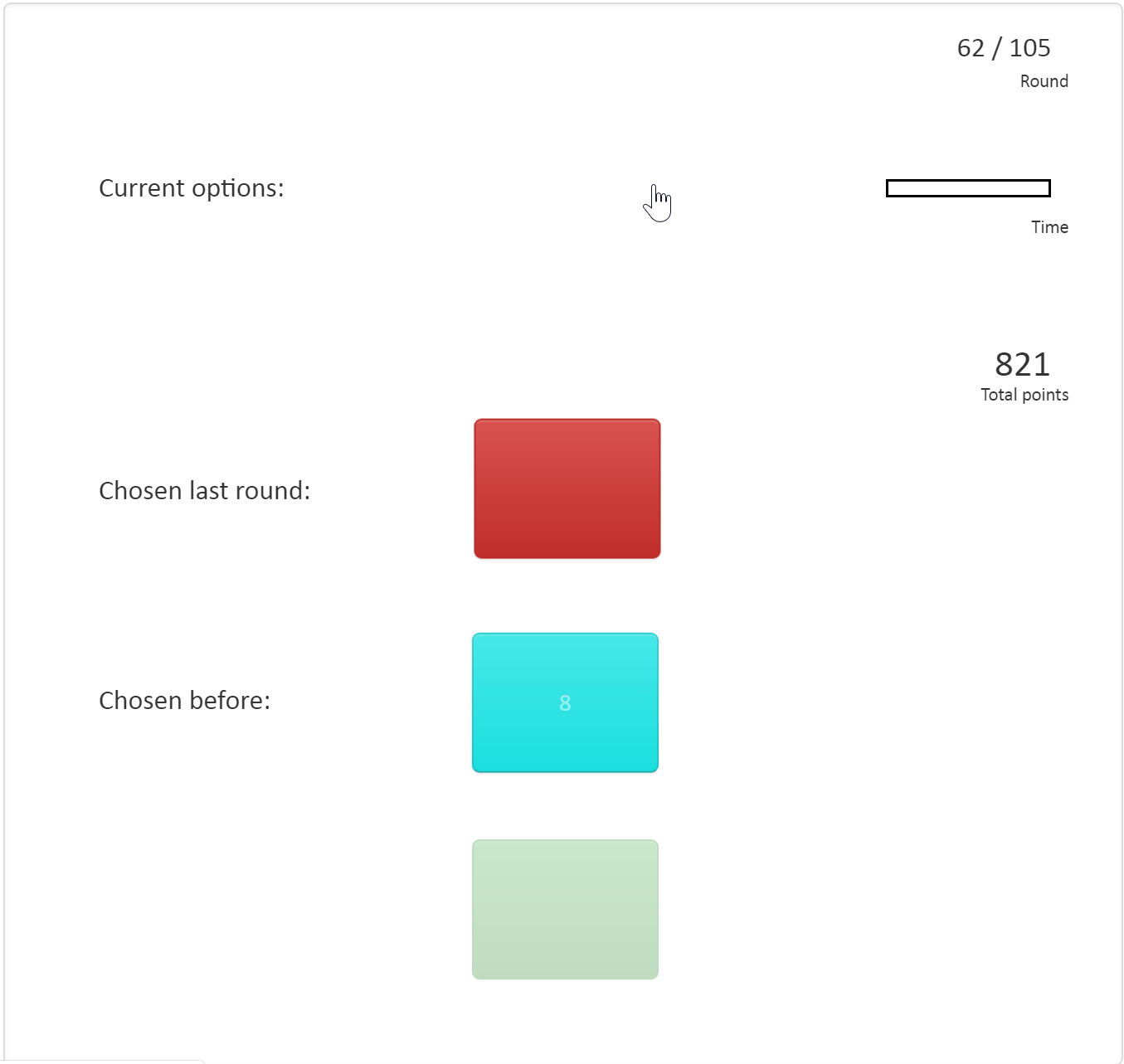
Design: Example Screen



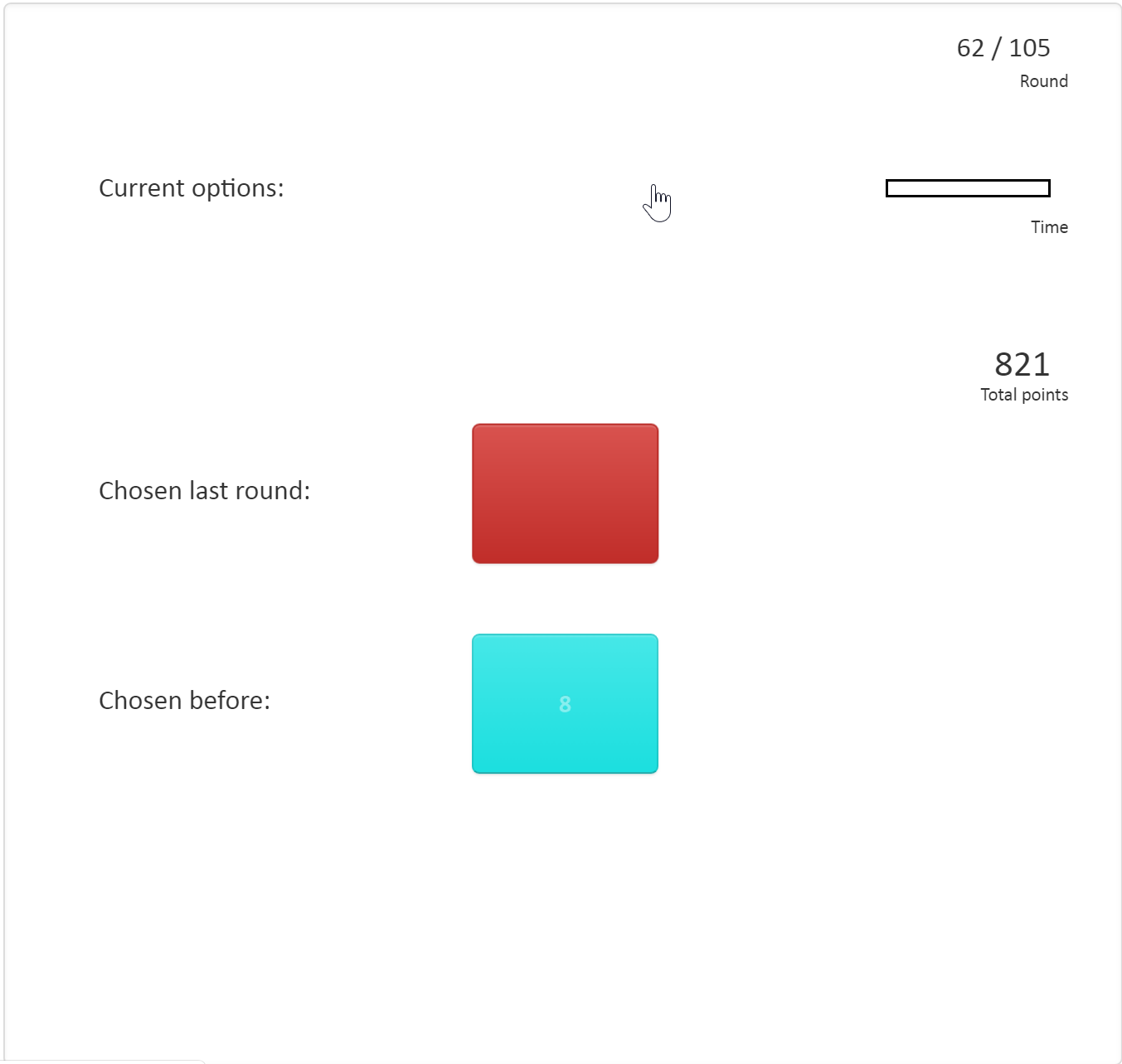
Design: Example Screen



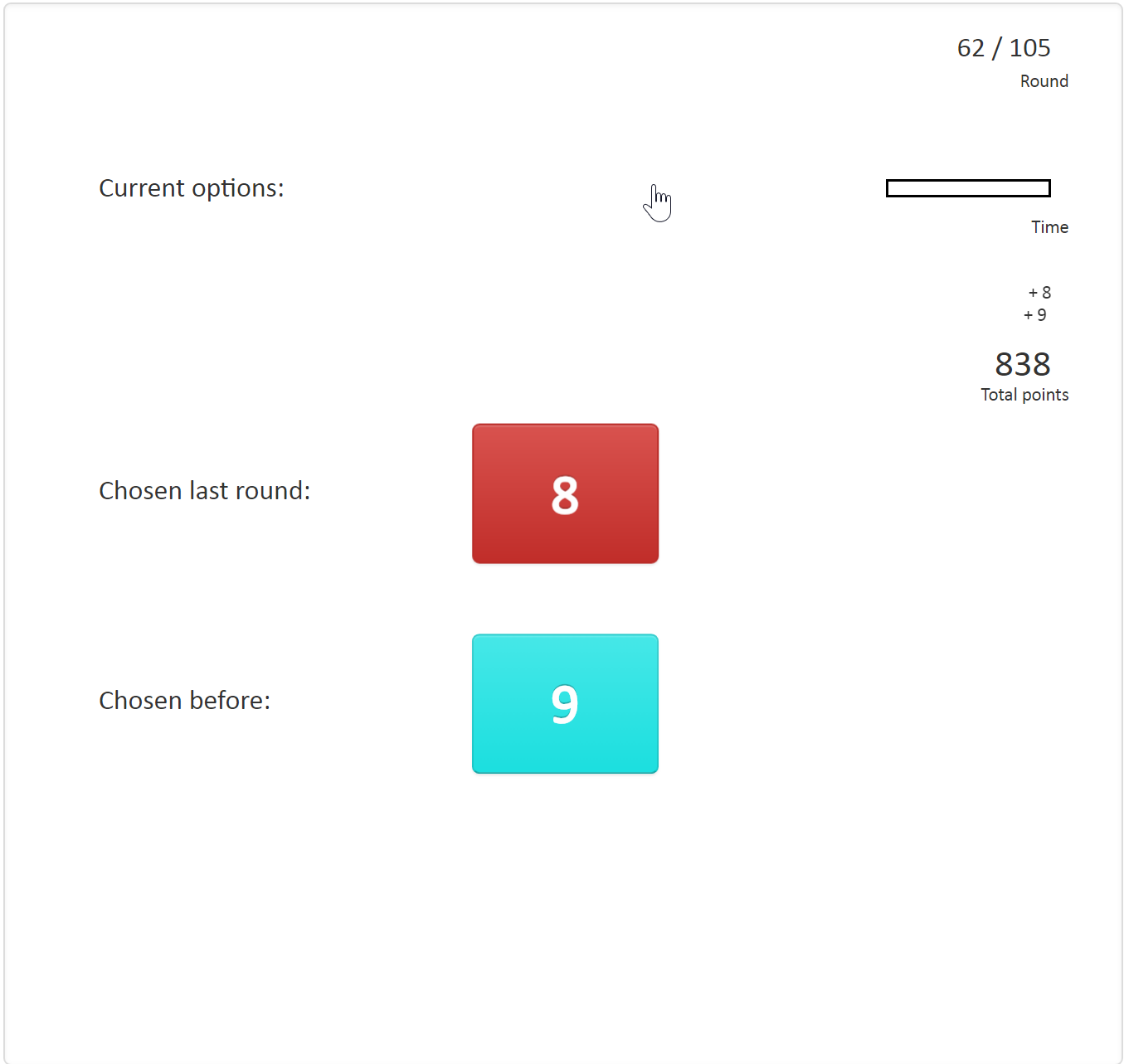
Design: Example Screen



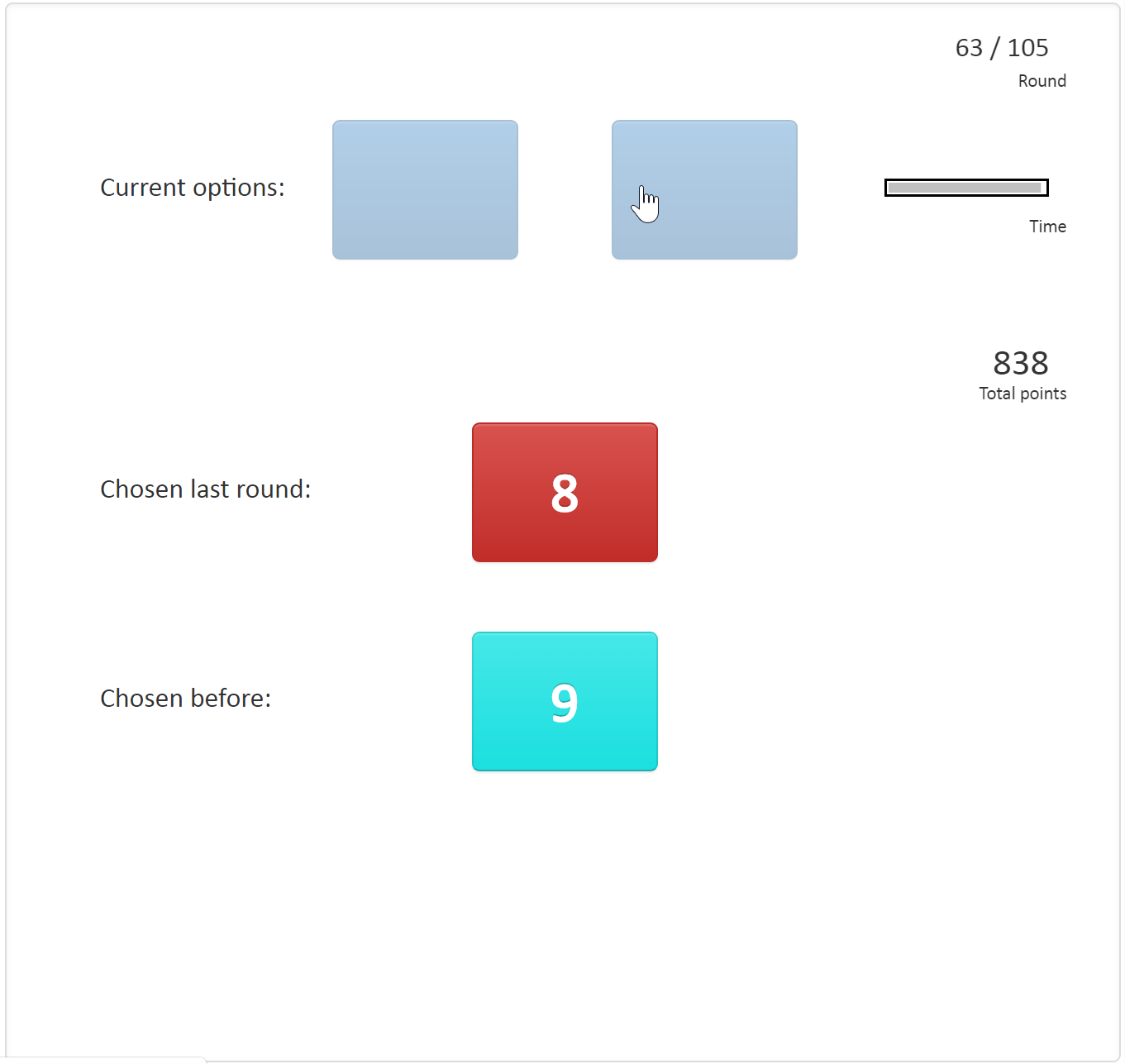
Design: Example Screen



Design: Example Screen



Design: Example Screen



Design: Payoff Vectors

Option		Payoff Vectors (immediate, delayed)	
color e.g.	(total value)	**Group A**	**Group B**
	(18)	$(11, 7)_A$	$(7, 11)_B$
	(16)	$(6, 10)_A$	$(10, 6)_B$
	(14)	$(9, 5)_A$	$(5, 9)_B$
	(12)	$(4, 8)_A$	$(8, 4)_B$
	(10)	$(7, 3)_A$	$(3, 7)_B$
	(8)	$(2, 6)_A$	$(6, 2)_B$

Hypotheses: $(11, 7)_A$ chosen more often than $(7, 11)_B$; $(10, 6)_B$ more than $(6, 10)_A$; ...

$(11, 7)_A$ and $(6, 10)_A$ further apart than $(6, 10)_A$ and $(9, 5)_A$.

Potentially even: $(9, 5)_A$ preferred to $(6, 10)_A$.

Design: Payoff Vectors

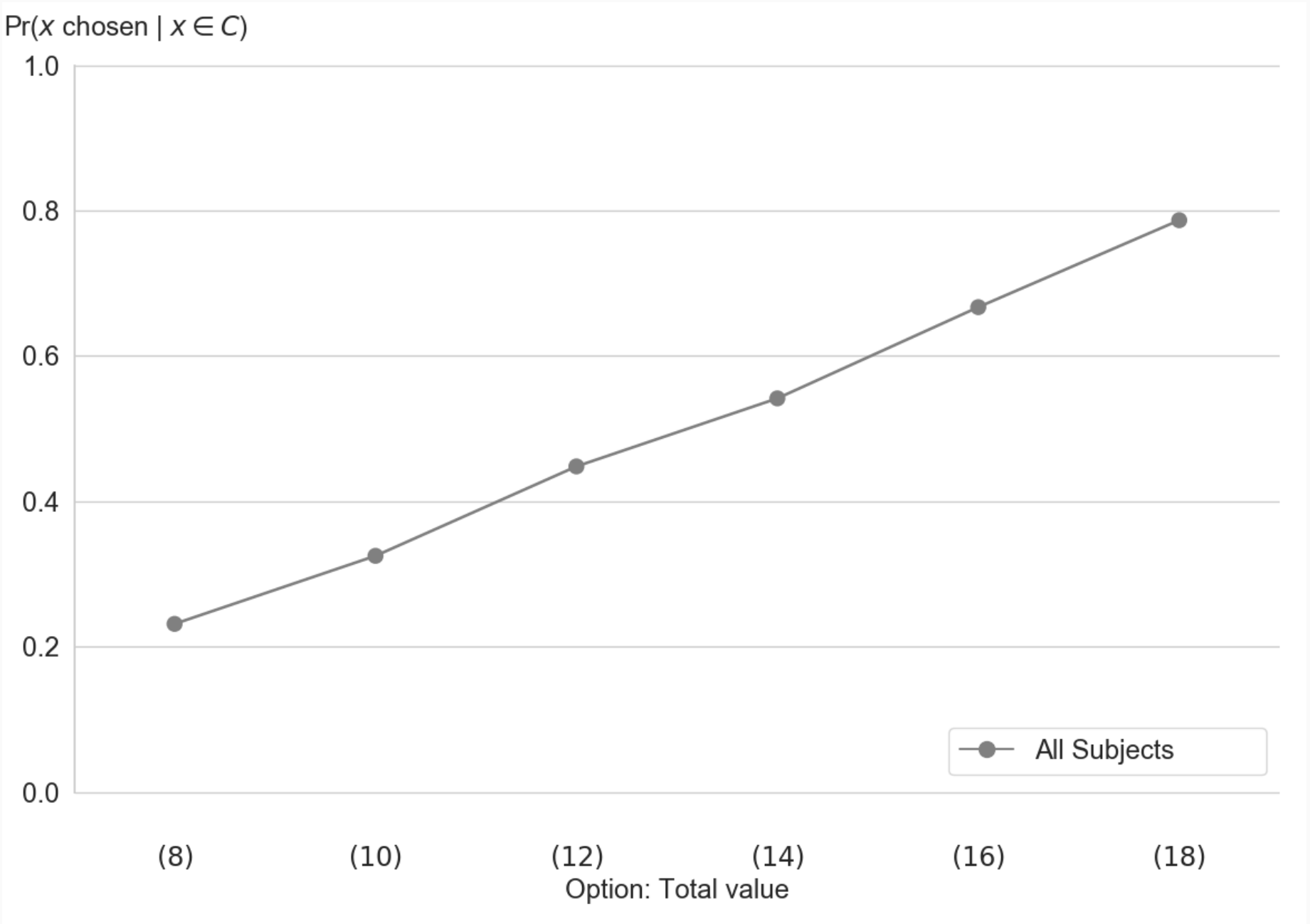
Option		Payoff Vectors (immediate, delayed)	
color e.g.	(total value)	**Group A**	**Group B**
	(18)	$(11, 7)_A$	$(7, 11)_B$
	(16)	$(6, 10)_A$	$(10, 6)_B$
	(14)	$(9, 5)_A$	$(5, 9)_B$
	(12)	$(4, 8)_A$	$(8, 4)_B$
	(10)	$(7, 3)_A$	$(3, 7)_B$
	(8)	$(2, 6)_A$	$(6, 2)_B$

Hypotheses: $(11, 7)_A$ chosen more often than $(7, 11)_B$; $(10, 6)_B$ more than $(6, 10)_A$; ...

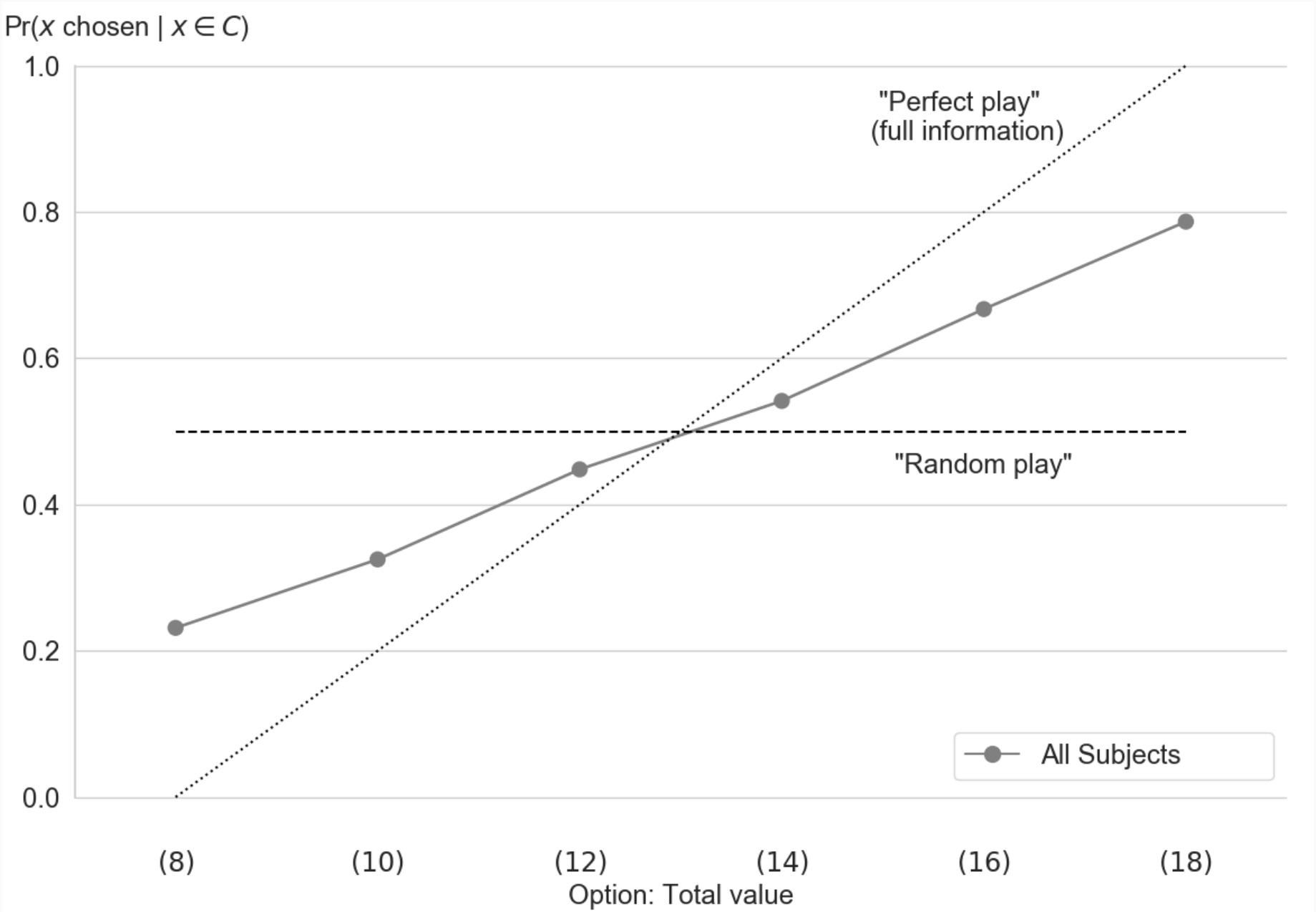
$(11, 7)_A$ and $(6, 10)_A$ further apart than $(6, 10)_A$ and $(9, 5)_A$.

Potentially even: $(9, 5)_A$ preferred to $(6, 10)_A$.

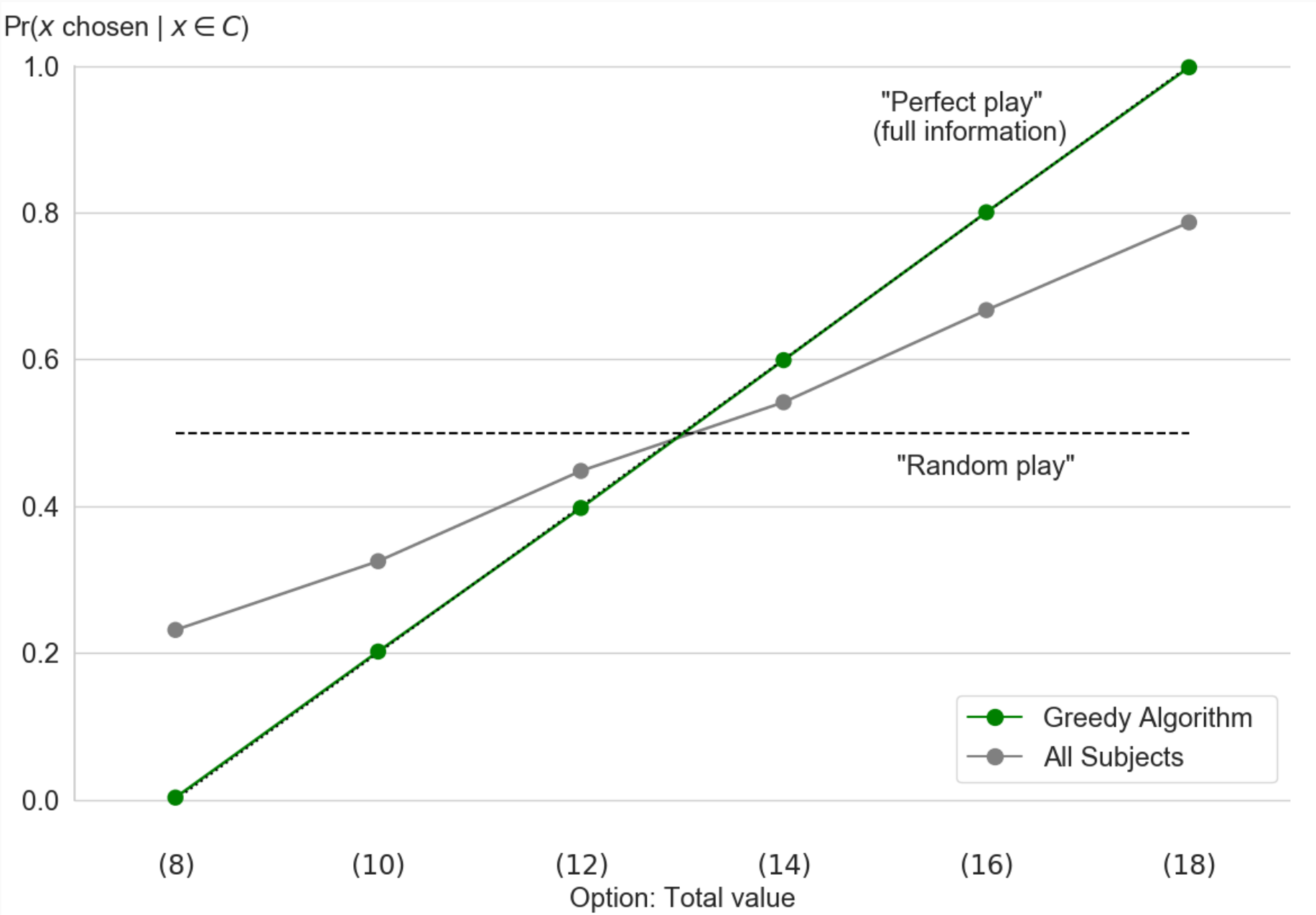
Results: Choice Frequencies



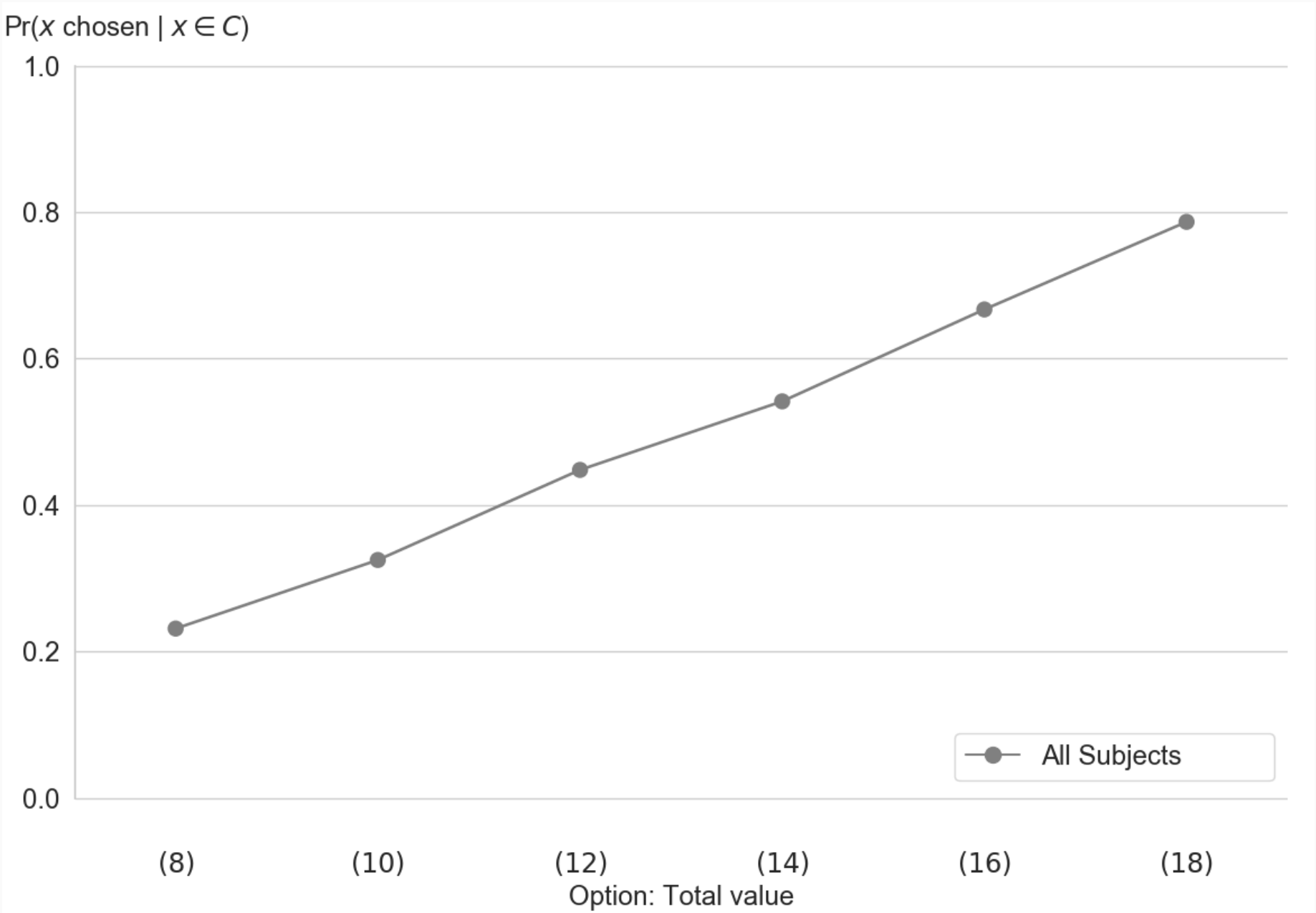
Results: Choice Frequencies



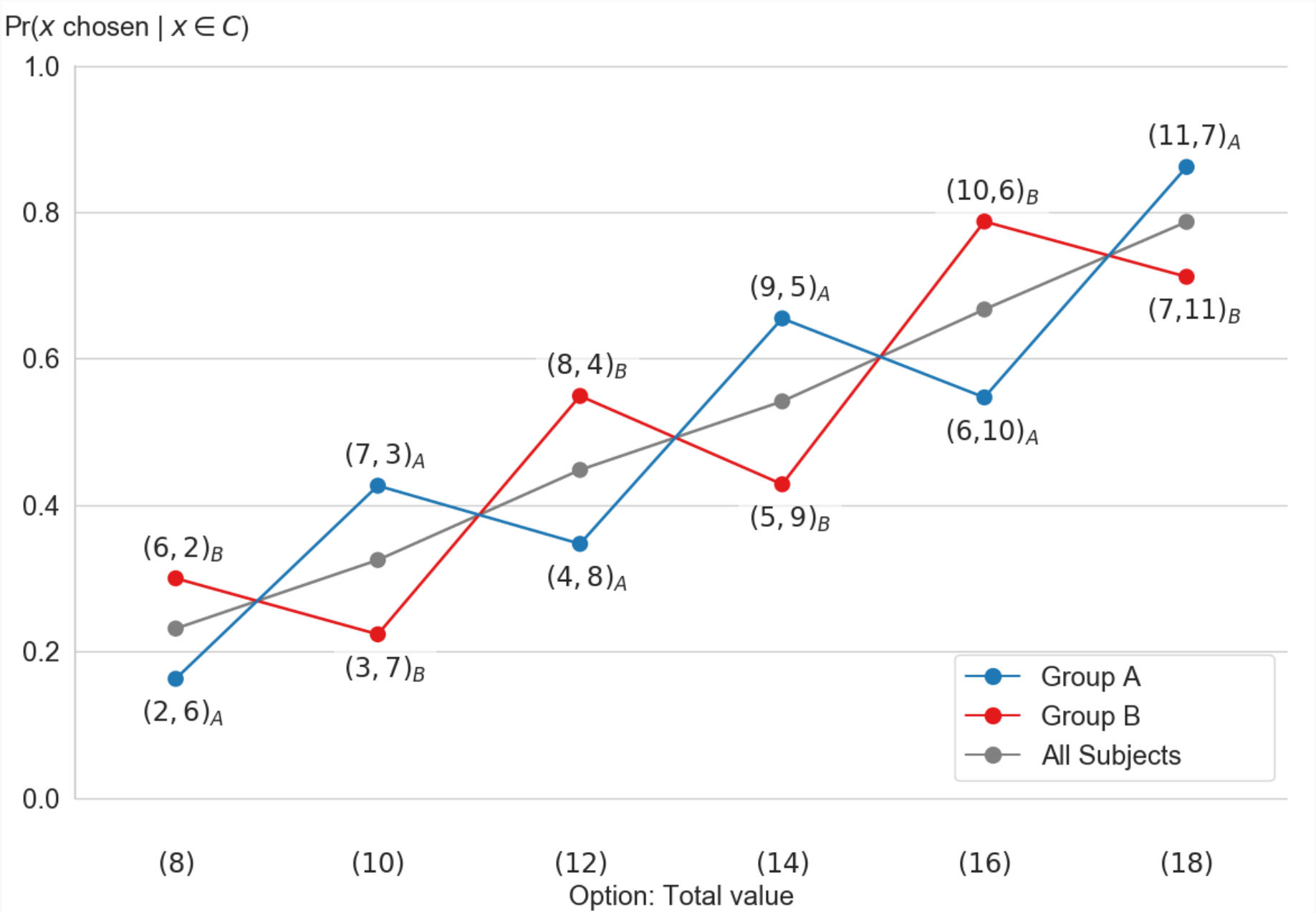
Results: Choice Frequencies



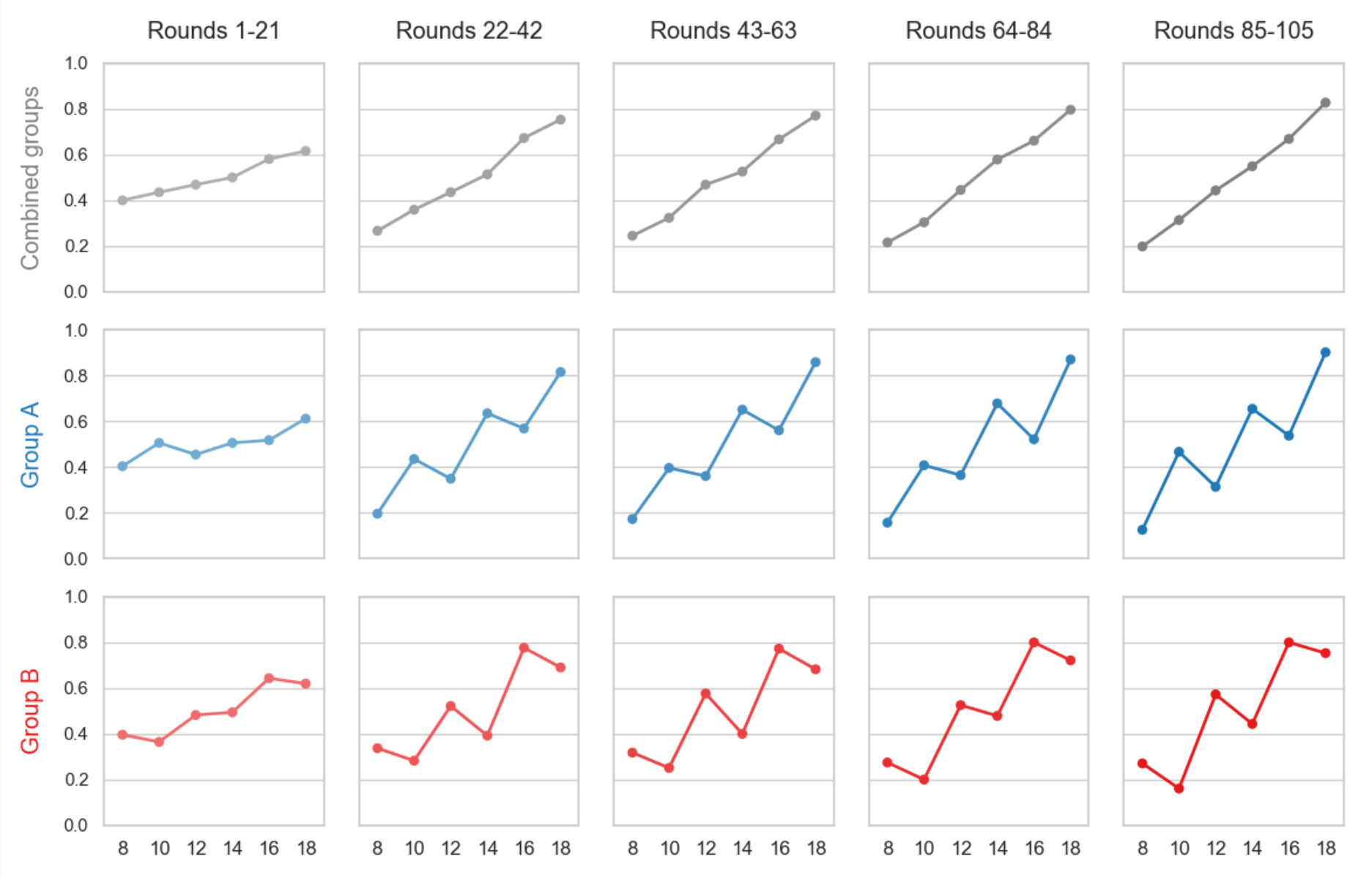
Results: Choice Frequencies



Results: Choice Frequencies



Results: Bias over time



Summary: Further Results

- Estimated latent utility function: $u(x) = x_1 + 0.4x_2$
- Elicited beliefs are in accordance with choice behavior.
- Considerable heterogeneity in degree of biasedness.
  - Correlated to impatience in hypothetical intertemporal choice.
  - (To do: Incentivized choice or field measures of impatience.)
- Treatment: Learning by observation
  - Subjects passively presented with feedback for 63 rounds.
  - Directly afterwards: 42 own decisions.
  - Bias attenuated; low right after the learning phase, then gradually increasing.
  - Suggests emergence of bias is connected to active decision making.

Summary: Further Results

- Estimated latent utility function: $u(x) = x_1 + 0.4x_2$
- Elicited beliefs are in accordance with choice behavior.
- Considerable heterogeneity in degree of biasedness.
  - Correlated to impatience in hypothetical intertemporal choice.
  - (To do: Incentivized choice or field measures of impatience.)
- Treatment: Learning by observation
  - Subjects passively presented with feedback for 63 rounds.
  - Directly afterwards: 42 own decisions.
  - Bias attenuated; low right after the learning phase, then gradually increasing.
  - Suggests emergence of bias is connected to active decision making.

Summary: Further Results

- Estimated latent utility function: $u(x) = x_1 + 0.4x_2$
- Elicited beliefs are in accordance with choice behavior.
- Considerable heterogeneity in degree of biasedness.
  - Correlated to impatience in hypothetical intertemporal choice.
  - (To do: Incentivized choice or field measures of impatience.)
- Treatment: Learning by observation
  - Subjects passively presented with feedback for 63 rounds.
  - Directly afterwards: 42 own decisions.
  - Bias attenuated; low right after the learning phase, then gradually increasing.
  - Suggests emergence of bias is connected to active decision making.

Summary: Further Results

- Estimated latent utility function: $u(x) = x_1 + 0.4x_2$
- Elicited beliefs are in accordance with choice behavior.
- Considerable heterogeneity in degree of biasedness.
  - Correlated to impatience in hypothetical intertemporal choice.
  - (To do: Incentivized choice or field measures of impatience.)
- Treatment: Learning by observation
  - Subjects passively presented with feedback for 63 rounds.
  - Directly afterwards: 42 own decisions.
  - Bias attenuated; low right after the learning phase, then gradually increasing.
  - Suggests emergence of bias is connected to active decision making.

Outlook

- Relation to actual reward discounting – ideally with field measure
- Relation to working memory
  - known to affect reward discounting (Wesley/Bickel 2014)
- Potential explanation: Differential precision in memory
- Investigate this using...
  - response time data
  - more fine-grained belief data
  - variations in timing, payoff vectors