Two stories about distortion in social choice

Ioannis Caragiannis

iannis@cs.au.dk

Two stories…

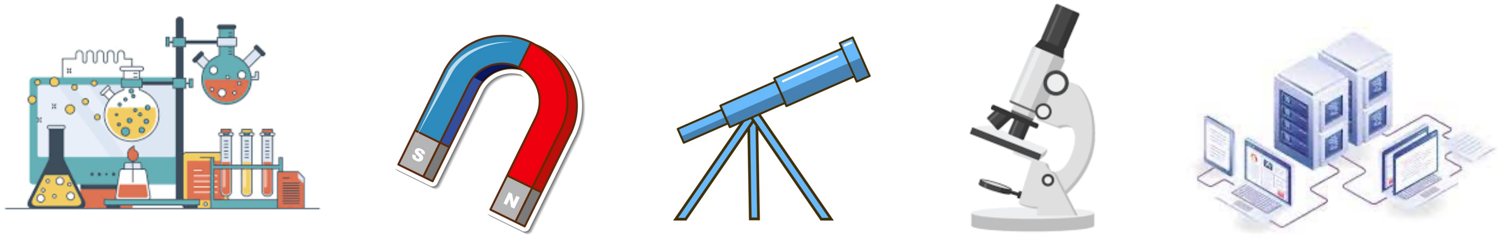
- Based on recent work on the distortion of
  - **impartial culture electorates** (with Karl Fehrs)
  - **sortition** (with Evi Micha and Jannik Peters)

The **distortion** of **impartial culture electorates**

Voting

- Voters (agents)

- Candidates (alternatives)



Voting rules

Voting rules

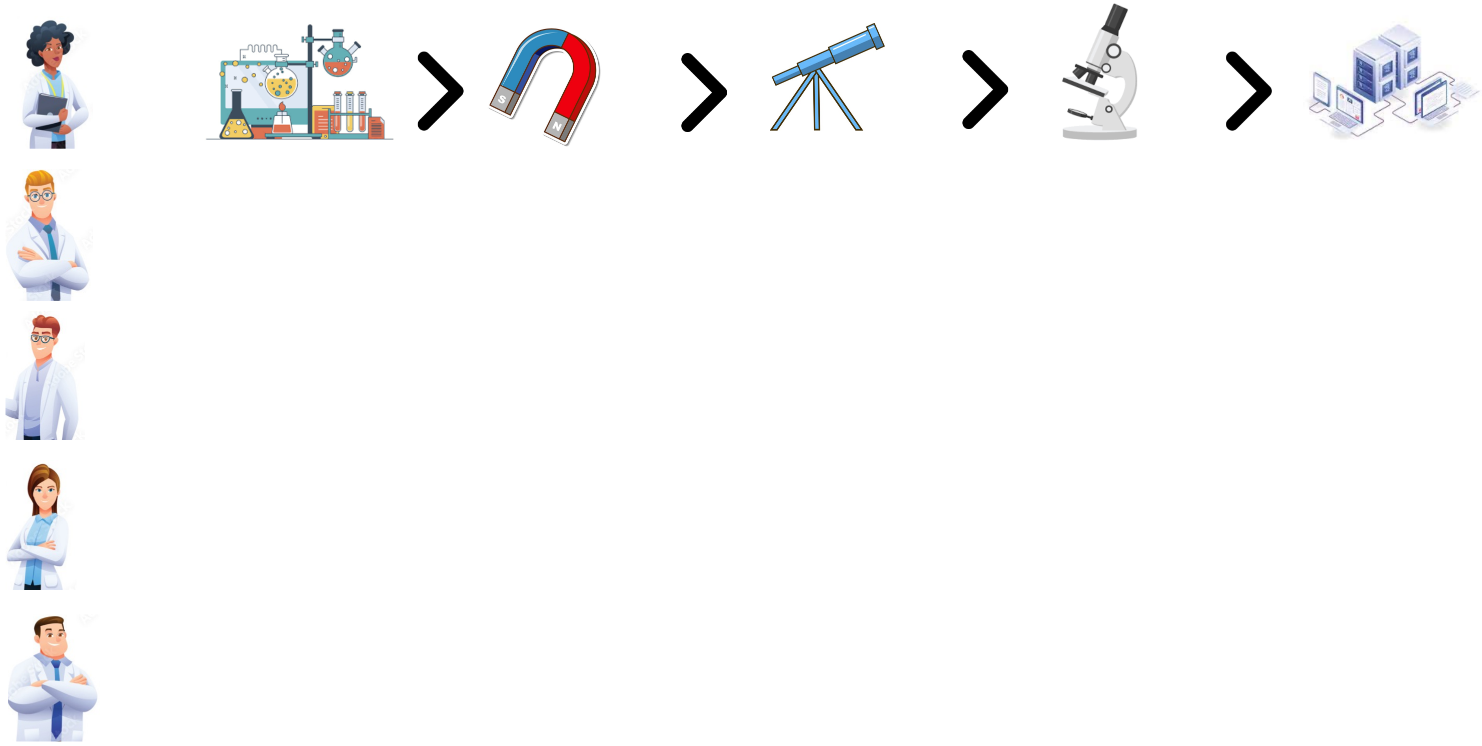
- **Agents**





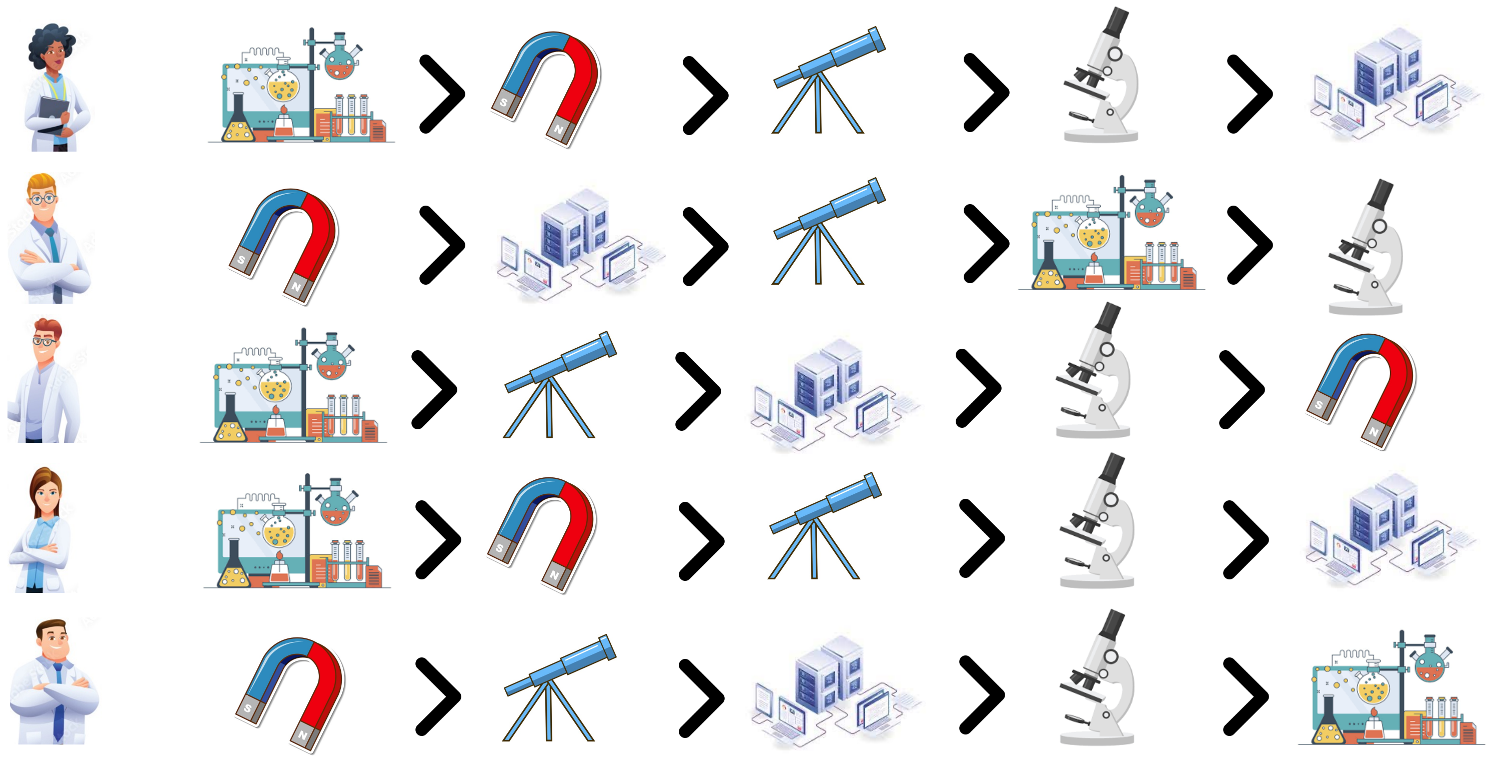
Voting rules

- **Agents** submit **rankings of alternatives**



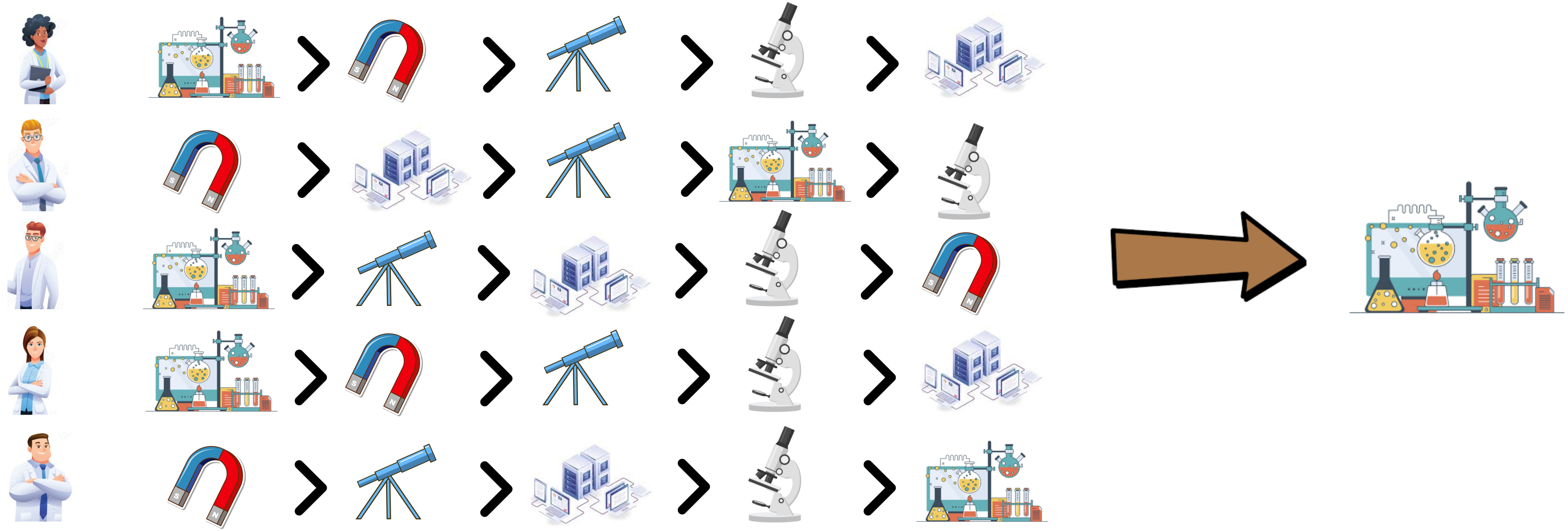
Voting rules

- **Agents** submit **rankings of alternatives**
- A **voting rule** takes such a **profile** of rankings



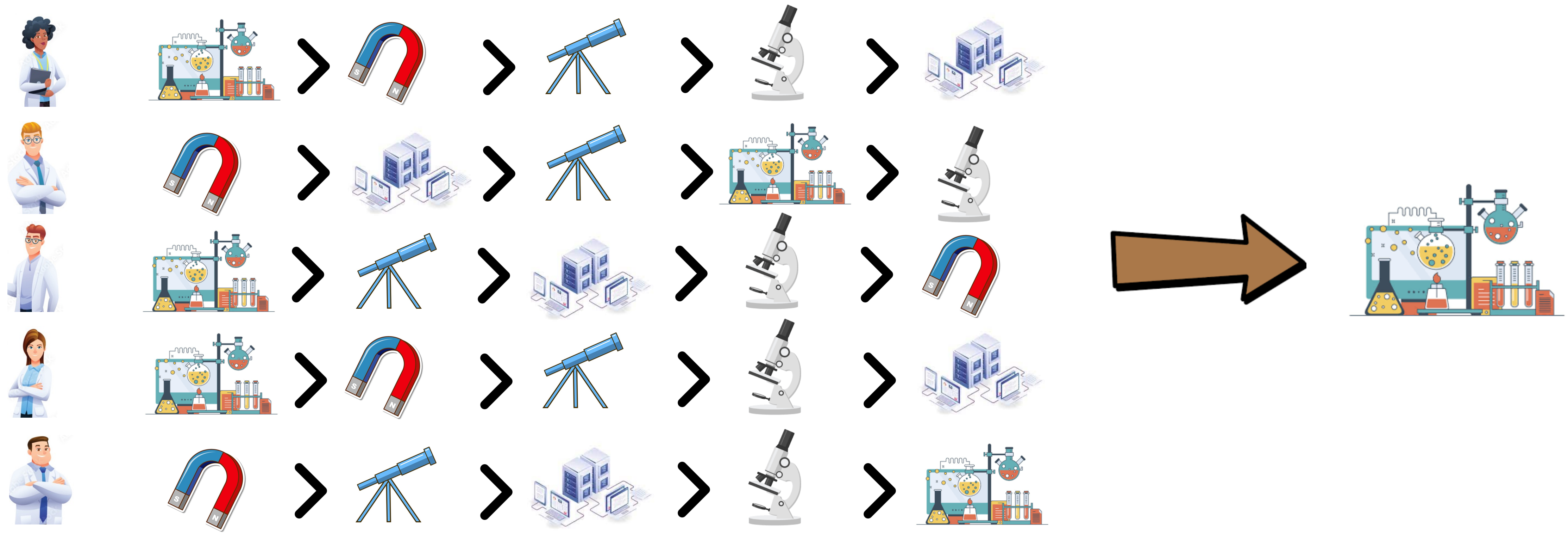
Voting rules

- **Agents** submit **rankings of alternatives**
- A **voting rule** takes such a **profile** of rankings and selects a **winnning alternative**



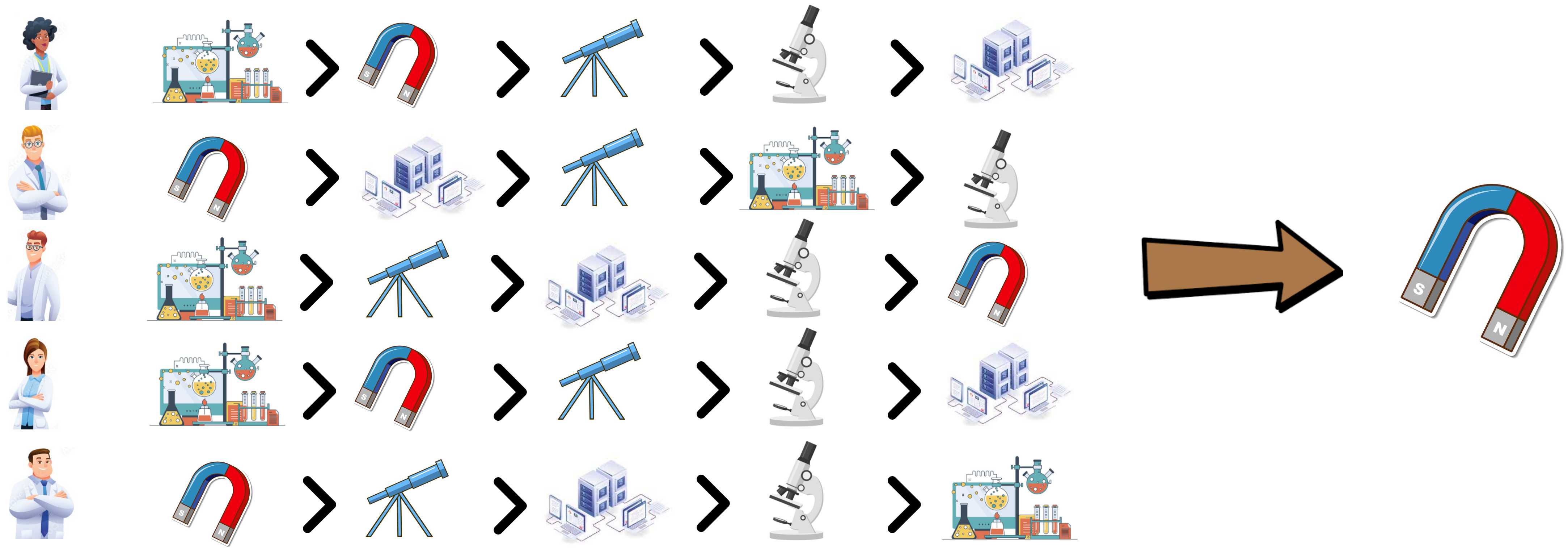
Voting rules

- **Plurality**



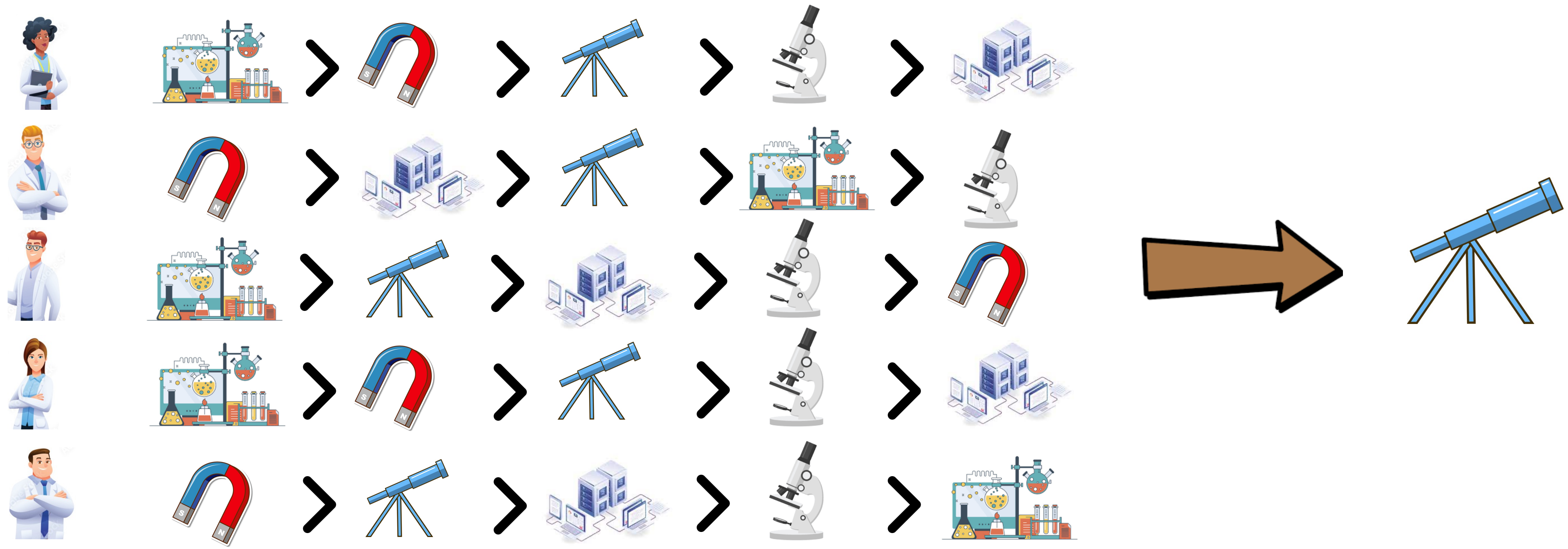
Voting rules

- **Plurality**, **Borda**



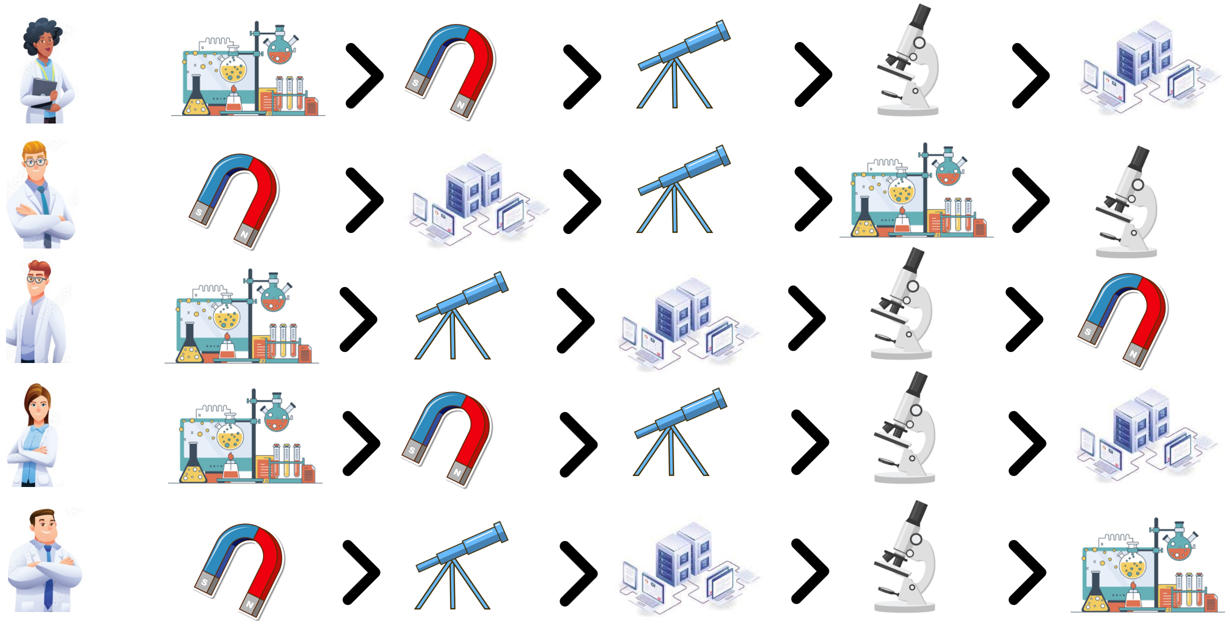
Voting rules

- **Plurality**, **Borda**, **Veto**



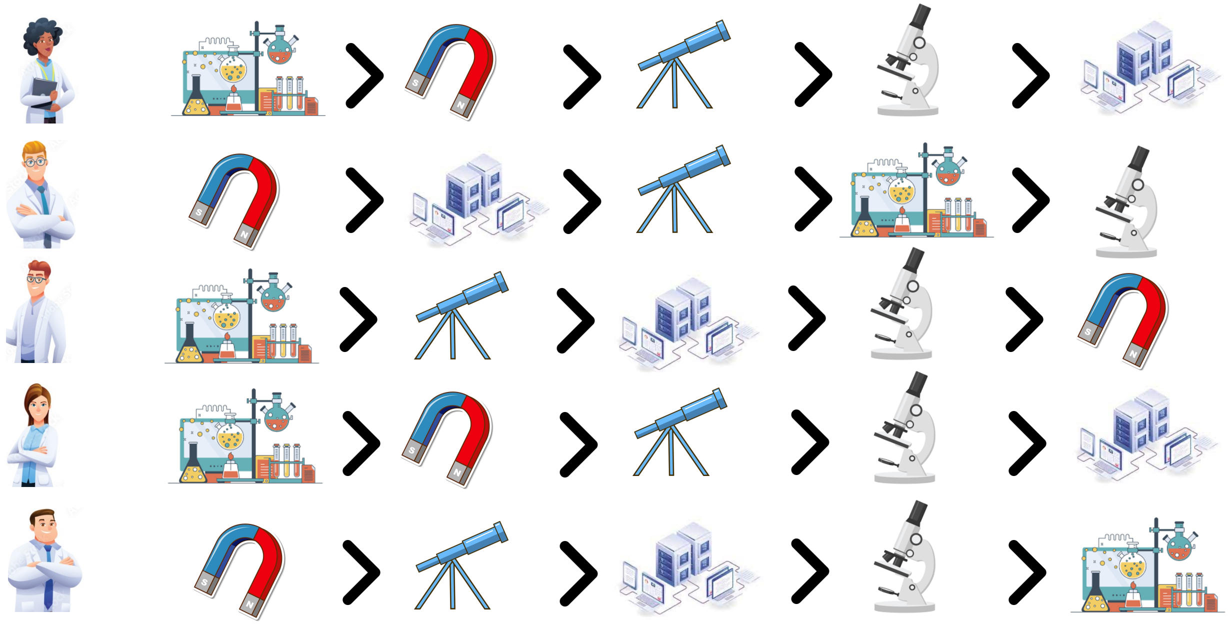
Evaluation of voting rules

- Each agent has hidden **valuations for the alternatives** that are **consistent to her ranking**



Evaluation of voting rules

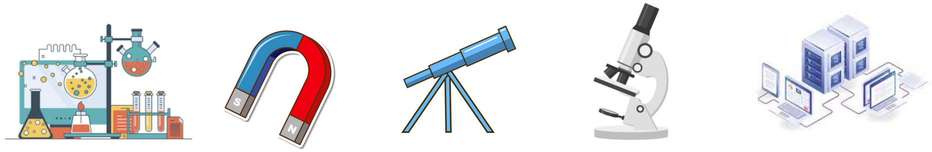
- Each agent has hidden **valuations for the alternatives** that are **consistent to her ranking**



0.22	0.21	0.20	0.19	0.18
0.01	0.49	0.02	0.00	0.48
0.34	0.00	0.33	0.01	0.32
0.22	0.21	0.20	0.19	0.18
0.00	0.34	0.33	0.01	0.32

Evaluation of voting rules

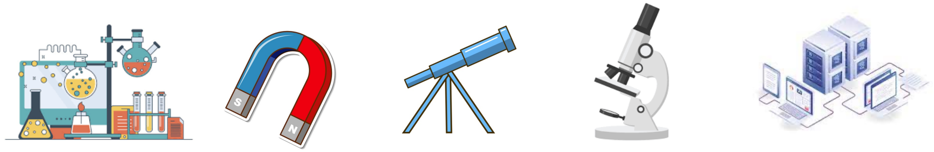
- Each agent has hidden **valuations for the alternatives** that are **consistent to her ranking**
- Quality of an outcome = its **social welfare**



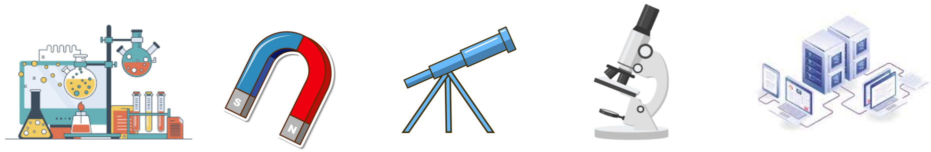
0.22	0.21	0.20	0.19	0.18
0.01	0.49	0.02	0.00	0.48
0.34	0.00	0.33	0.01	0.32
0.22	0.21	0.20	0.19	0.18
0.00	0.34	0.33	0.01	0.32

Evaluation of voting rules

- Each agent has hidden **valuations for the alternatives** that are **consistent to her ranking**
- Quality of an outcome = its **social welfare**



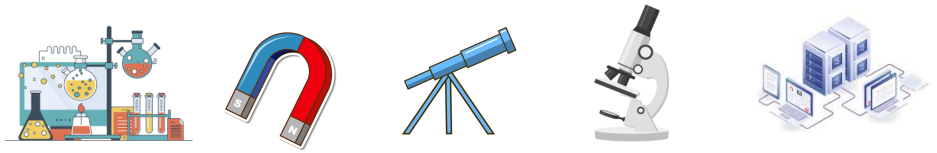
0.79	1.25	1.08	0.40	1.48



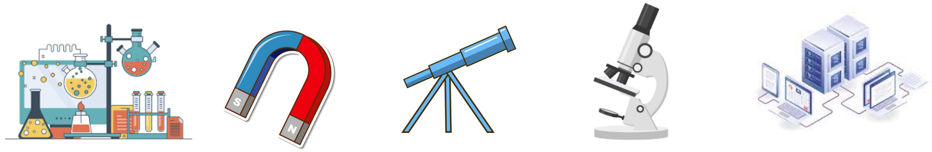
0.22	0.21	0.20	0.19	0.18
0.01	0.49	0.02	0.00	0.48
0.34	0.00	0.33	0.01	0.32
0.22	0.21	0.20	0.19	0.18
0.00	0.34	0.33	0.01	0.32

Evaluation of voting rules

- Each agent has hidden **valuations for the alternatives** that are **consistent to her ranking**
- Quality of an outcome = its social **welfare**



0.79	1.25	1.08	0.40	1.48
Plurality	Borda	Veto		optimal alternative



0.22	0.21	0.20	0.19	0.18
0.01	0.49	0.02	0.00	0.48
0.34	0.00	0.33	0.01	0.32
0.22	0.21	0.20	0.19	0.18
0.00	0.34	0.33	0.01	0.32

Formal definitions

- Set of n **agents** N, set of m **alternatives** A
- Agent $i \in N$ has non-negative **valuation** $v_i(a)$ for each alternative $a \in A$
- **Valuation profile** v (valuations of all agents for all alternatives)
- The **ranking** $\succ_i$ of agent $i \in N$ is **consistent to her valuations**, i.e., $a \succ_i b$ implies that $v_i(a) \geq v_i(b)$.
- **Ranking profile** $P = (\succ_1, \succ_2, \dots, \succ_n)$. Overall, $P \mapsto v$
- The **social welfare** $SW(a, v)$ of alternative $a \in A$ in valuation profile v is its total valuation, i.e., $SW(a, v) = \sum_{i \in N} v_i(a)$
- **Optimal alternative** is the one that maximizes the social welfare
- **Optimal social welfare** $SW^*(v) = \max_{a \in A} SW(a, v)$

Formal definitions (contd.)

- **Voting rule**: a function f that takes as input a ranking profile P and returns an alternative $a \in A$
- The **distortion** $dist(f)$ of a voting rule f is defined as

$$dist(f) = \max_{v} \max_{P \mapsto v} \frac{SW^*(v)}{SW(f(P), v)}$$

maximum over all valuation profiles

maximum over all consistent ranking profiles

how far from optimal is the social welfare of the alternative returned by the voting rule?

Formal definitions (contd.)

- **Voting rule**: a function f that takes as input a ranking profile P and returns an alternative $a \in A$
- The **distortion** $dist(f)$ of a voting rule f is defined as

$$dist(f) = \max_{v} \max_{P \mapsto v} \frac{SW^*(v)}{SW(f(P), v)}$$

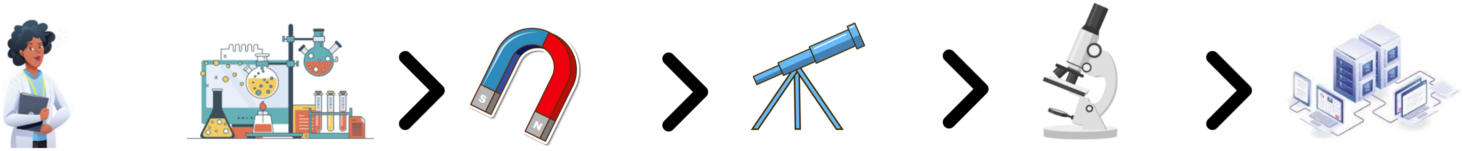
- Procaccia & Rosenschein (2006)
- Boutilier, C., Haber, Lu, Procaccia, & Sheffet (2015)
- Ebadian, Kahng, Peters, & Shah (2024)
- Anshelevich, Filos-Ratsikas, Shah, & Voudouris (2021)

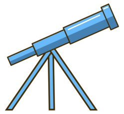
Flavor of distortion results

- **Plurality** has distortion $O(m^2)$, which is optimal among deterministic voting rules
- The best possible distortion among all (possible randomized) voting rules is $\Theta(\sqrt{m})$
- Restrictions: **unit range**, **unit sum valuations**
- Without such restrictions, no distortion bound is possible for deterministic rules, and the **trivial randomized rule** that returns an alternative uniformly at random has best possible distortion

Low-distortion using value queries

- The idea: in addition to the **ranking profile**, the voting rule (mechanism) can make a small number of **value queries** to each agent
- The **query** to agent i for alternative a returns the value $v_i(a)$



- Query: Agent , what is your value for alternative  ? Answer: 0.20

- Amanatidis, Birmpas, Filos-Ratsikas, & Voudouris (2021)
- Main result: **Constant distortion** with $O(\ln^2 m)$ queries per agent
- Feature: **no restrictions** on the valuations

The distortion of impartial culture profiles

- C. & Fehrs (2024)
- A common **probability distribution** F
- **Independently** for each agent-alternative pair, agent $i \in N$ draws a **random value** $v_i(a)$ for alternative $a \in A$ according to p.d. F
- The **consistent profile** $P(v)$ is obtained from v by breaking ties among alternatives uniformly at random
- The **average distortion** of a voting rule/mechanism f on ranking profiles that are consistent to random values drawn from p.d. F

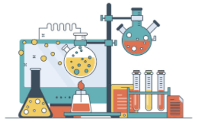
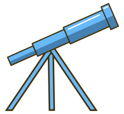
$$avd(f,F) = \frac{\mathbb{E}_{v\sim F}[SW^*(v)]}{\mathbb{E}_{v\sim F}[SW(f(P(v),v),v)]}$$

Binary distributions

- The p.d. F returns 1 with probability p and 0 otherwise

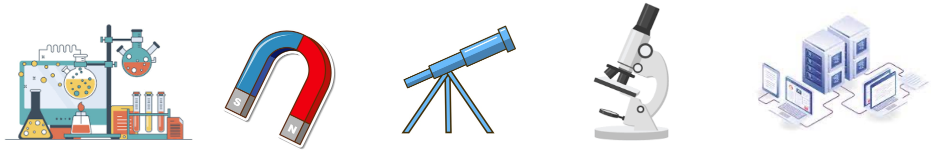
Binary distributions

- The p.d. F returns 1 with probability p and 0 otherwise

					
	0	0	1	1	0
	1	0	1	0	1
	1	0	1	1	1
	0	0	0	1	0
	0	0	1	0	1

Binary distributions

- The p.d. F returns 1 with probability p and 0 otherwise



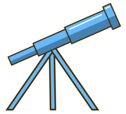
0	0	1	1	0
1	0	1	0	1
1	0	1	1	1
0	0	0	1	0
0	0	1	0	1

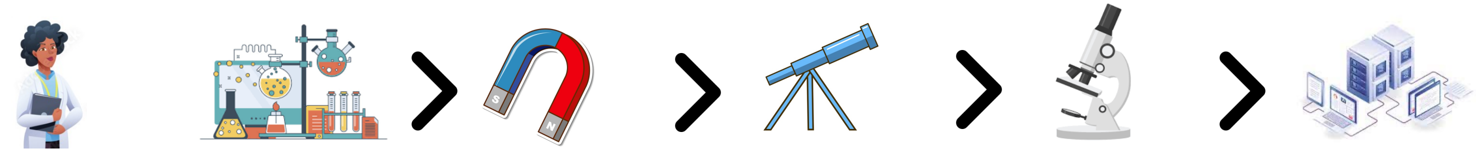
Consistent ranking profile, breaking ties u.a.r.

Average distortion can be high

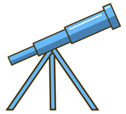
- Consider a binary distribution with $p = \frac{1}{nm}$
- The numerator in the average distortion definition is a **constant**
- W.h.p., there are **very few 1s in the top positions** of the rankings
- Without making queries, **the voting rule cannot guess** which among the top alternatives have these 1s
- Then, the alternative returned will have some 1 in the top position with probability at most $O(1/m)$
- Yields an average distortion of $\Omega(m)$, i.e., **as bad as picking a fixed or a random alternative**

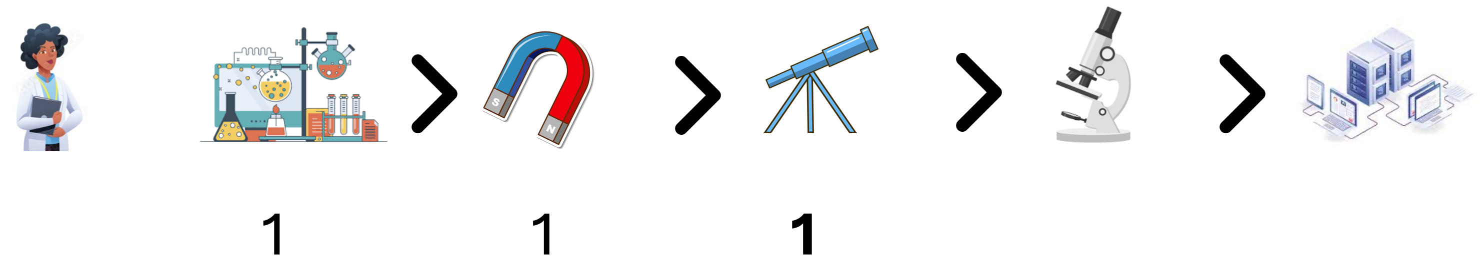
Mechanisms for binary distribution

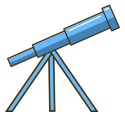
- F (i.e., p) is known to the mechanism
- Assume that we query the value of agent  for alternative 



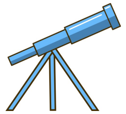
Mechanisms for binary distribution

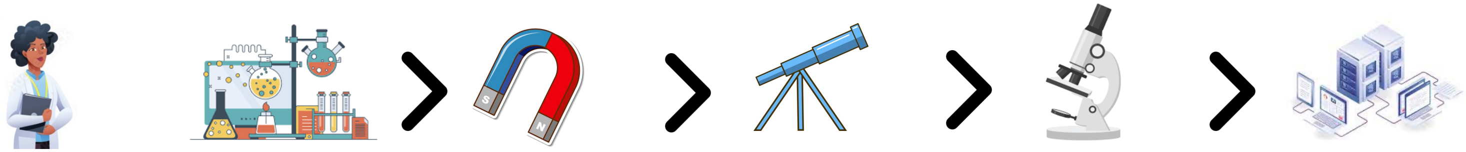
- F (i.e., p) is known to the mechanism
- Assume that we query the value of agent  for alternative 



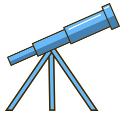
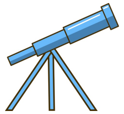
- If the value returned is 1, all alternatives ranked above  have value 1

Mechanisms for binary distribution

- F (i.e., p) is known to the mechanism
- Assume that we query the value of agent  for alternative 



0 0 0

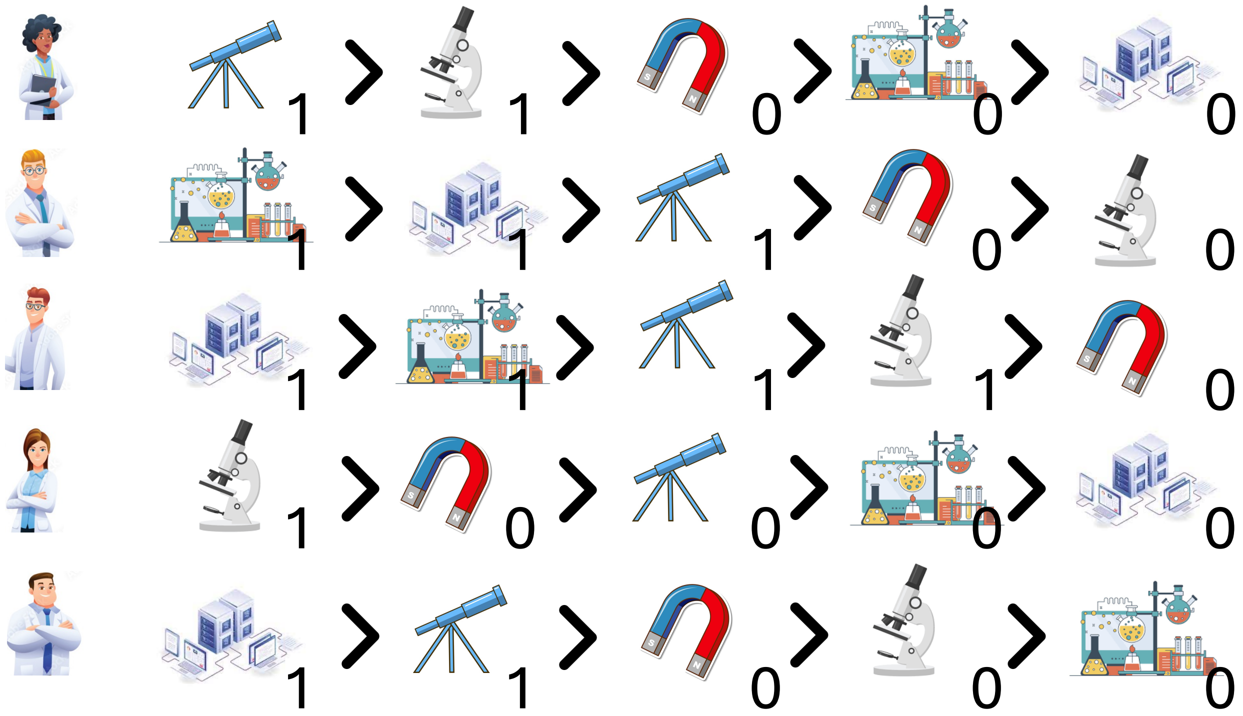
- If the value returned is 1, all alternatives ranked above  have value 1
- Otherwise, all alternatives ranked below  have value 0

Mechanisms for binary distribution

- **One query** per agent
- **Implied social welfare** of alternative a = number of agents who returned 1 to the query and rank a at the queried position or above

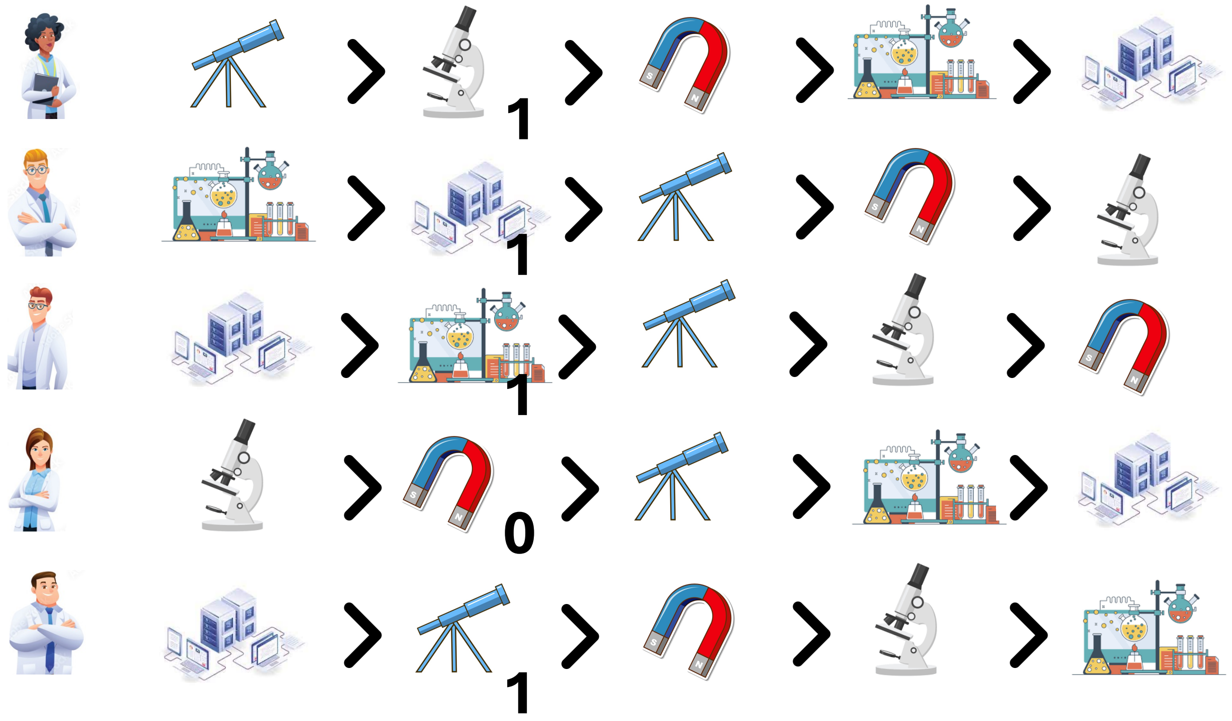
Mechanisms for binary distribution

- **One query** per agent
- **Implied social welfare** of alternative a = number of agents who returned 1 to the query and rank a at the queried position or above



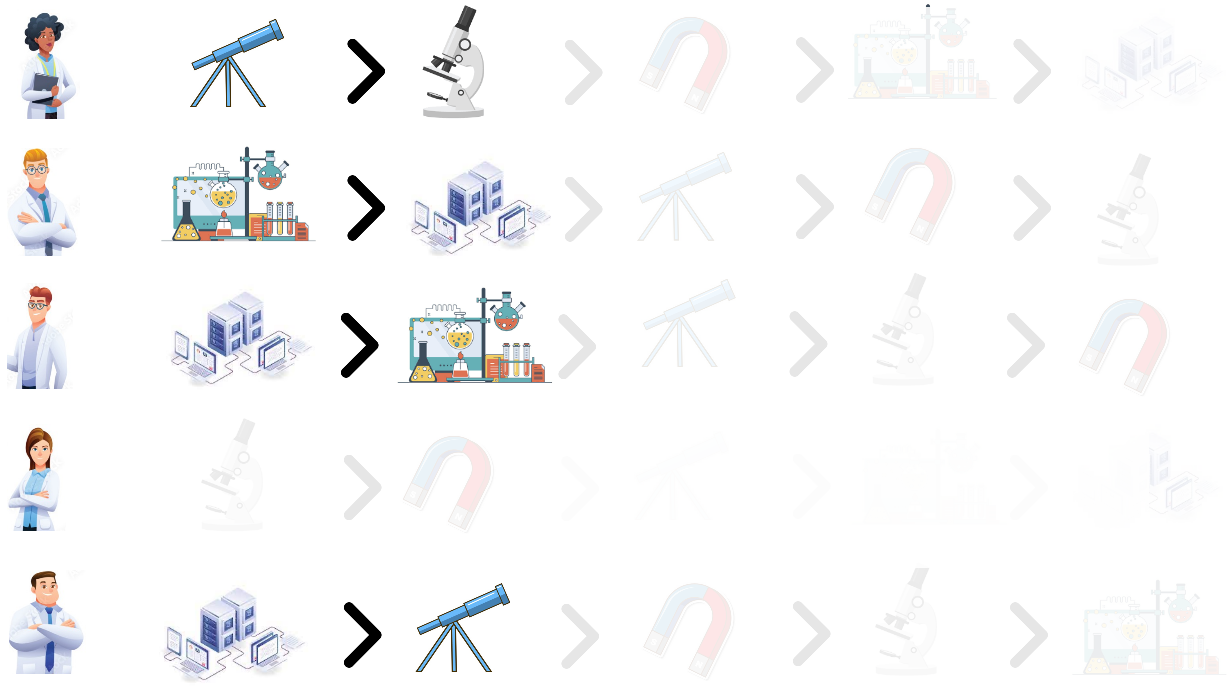
Mechanisms for binary distribution

- **One query** per agent (e.g., at the second position)
- **Implied social welfare** of alternative a = number of agents who returned 1 to the query and rank a at the queried position or above



Mechanisms for binary distribution

- **One query** per agent (e.g., at the second position)
- **Implied social welfare** of alternative a = number of agents who returned 1 to the query and rank a at the queried position or above



Mechanisms for binary distribution

- **One query** per agent (e.g., at the second position)
- **Implied social welfare** of alternative a = number of agents who returned 1 to the query and rank a at the queried position or above

Mechanism Mean

- Input: a profile P consistent to random agent-alternative values drawn from a binary p.d. F with known p
- Query each agent at position $\max\{1, \lfloor pm \rfloor\}$
- Return the alternative with **highest implied social welfare**

- Theorem: Mechanism Mean has **constant average distortion**
  - Analysis distinguishes between cases for small, medium, and large p

The random threshold mechanism RTMean

- Input: a profile P with underlying valuations drawn from a p.d. F
- Uses k **thresholds** parameters $\ell_1 < \ell_2 < \cdots < \ell_k$
- Select an integer t uniformly at random from $[k]$
- Sets $p = \Pr_{z \sim F}[z \geq \ell_t]$
- Simulate an **execution of MEAN on the binary distribution** F_p by
  - making the same value queries as MEAN for F_p, but
  - interpreting the values returned to each query as 1 if above ℓ_t and 0 otherwise
- Return as output the alternative that MEAN selects

The random threshold mechanism RTMean

- Theorem: For every p.d. F with expectation μ and variance σ^2, there are thresholds $\ell_1 < \ell_2 < \cdots < \ell_k$ so that the average distortion of mechanism RTMean is at most $O\left(\ln m + \ln \frac{\sigma^2}{\mu^2}\right)$
- Note: mechanism RTMean makes exactly **one query**

End of the first story

- Improved results for worst-case distortion
  - **Randomized mechanisms** with worst-case distortion $O(\ln m)$ using $O(\ln m)$ queries per agent
  - **Lower bound** on $\Omega(\ln m)$ the number of queries necessary for constant worst-case distortion

- Open problems
  - Is there a mechanism that achieves **constant distortion** for every p.d. F using a **constant number of queries** per agent?
  - What about **unknown p.d.** F?
  - **Tight bound** on #queries for constant worst-case distortion?

The **metric distortion** of **sortition**

Sortition

- The practice of **selecting a citizen’s assembly** (or panel) to decide among alternatives for a given issue
- Desired property: **fairness** in the selection of citizens so that different groups are **represented proportionally**

- Flanigan, Gölz, Gupta, Hennig, & Procaccia (2021)
- Meir, Sandomirskiy, and Tennenholtz (2021)
- Ebadian, Kehne, Micha, Procaccia, & Shah (2022)
- Ebadian & Micha (2023)

Sortition

- The practice of **selecting a citizen's assembly** (or panel) to decide among alternatives for a given issue
- Our focus: **efficiency of decisions** under **fairness constraints** in the selection of the panel
- Tool: a variant of **distortion**

- C., Micha, & Peters (2024)

Our setting

- Citizens (or **agents**) and **alternatives** are **points in a metric space**
- **Panels** are subsets of agents of size k
- The **social cost** $SC(a, P)$ of alternative a for a panel P is defined as the total distance of the alternative to all members of P, i.e.,

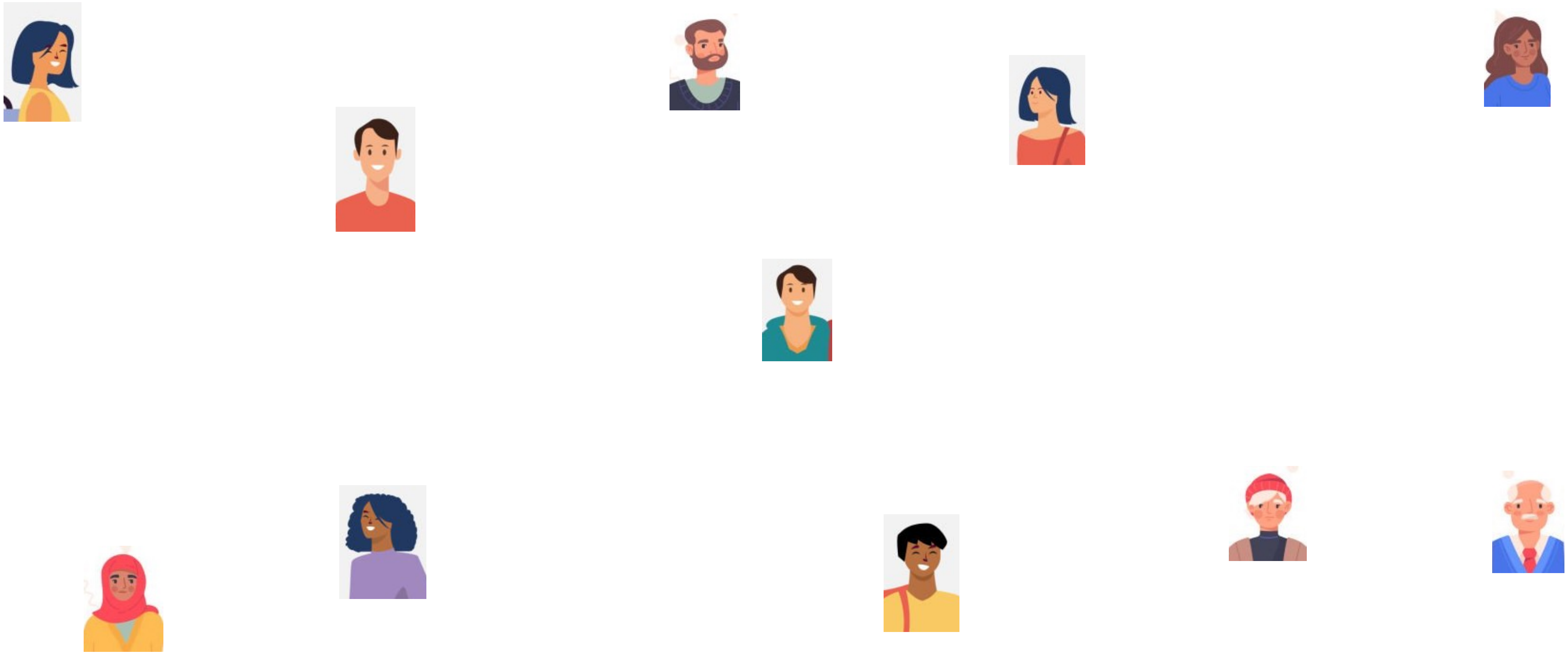
$$SC(a, P) = \sum_{i \in P} d(i, a)$$

- We denote by $SC(a)$ the social cost of alternative a (for all agents)
- The **optimal alternative** is the one with the minimum social cost

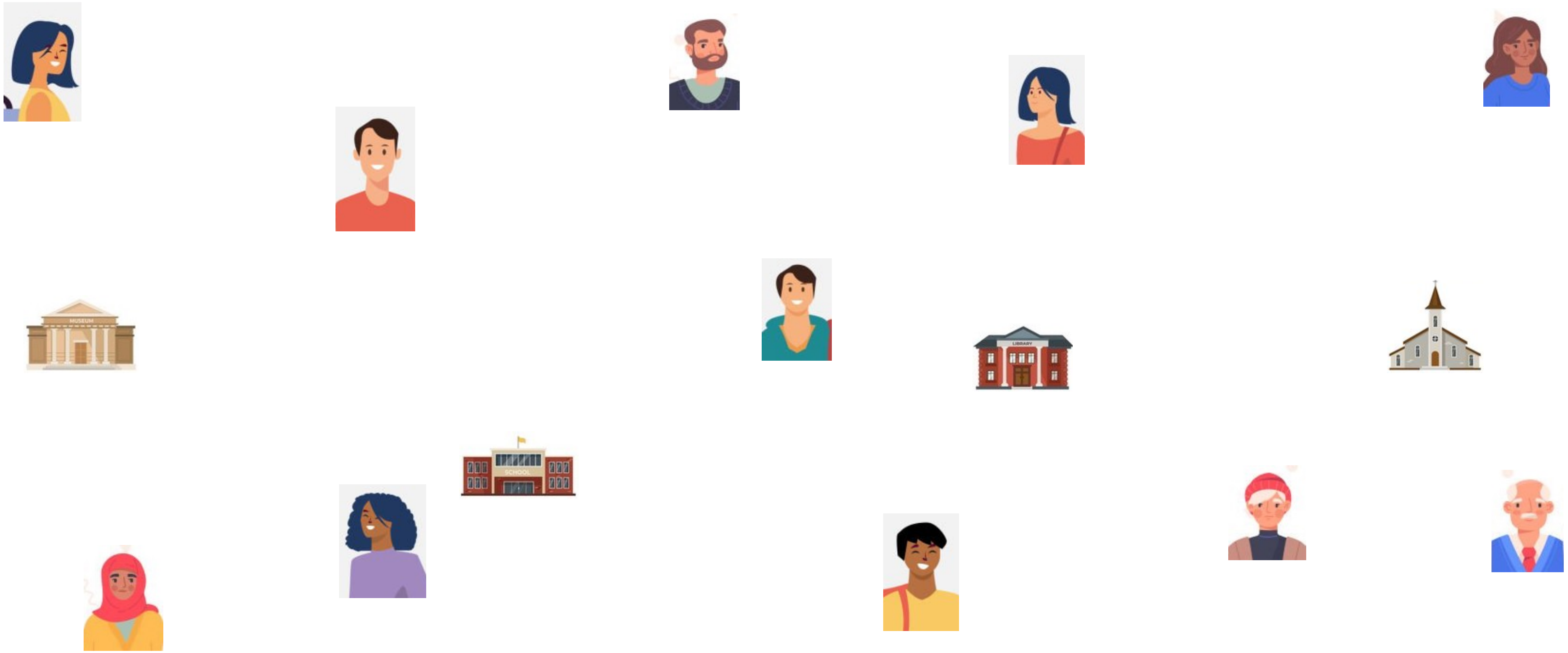
$$SC^*(A) = \min_{a \in A} SC(a)$$

- Among a set of alternatives A for a given issue, the panel P selects the alternative $P(A)$ of **minimum social cost for it**

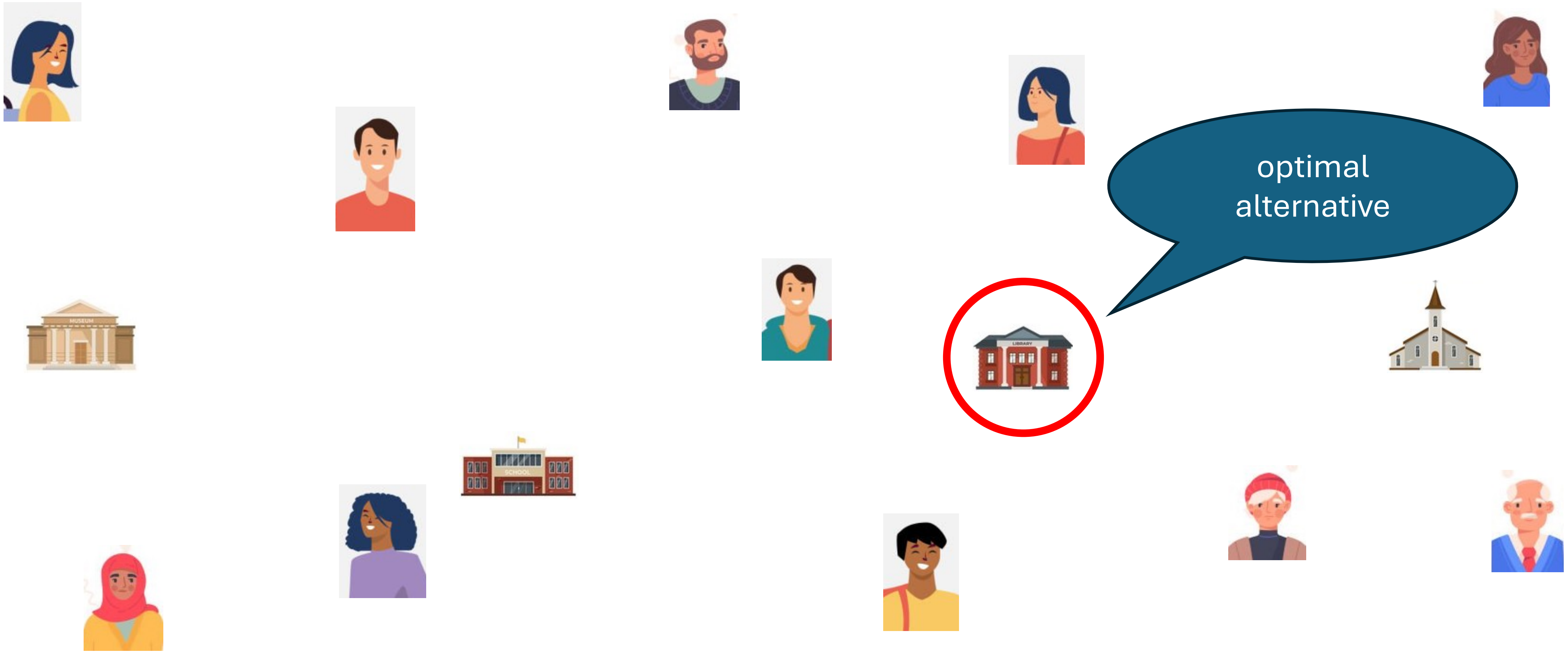
An example



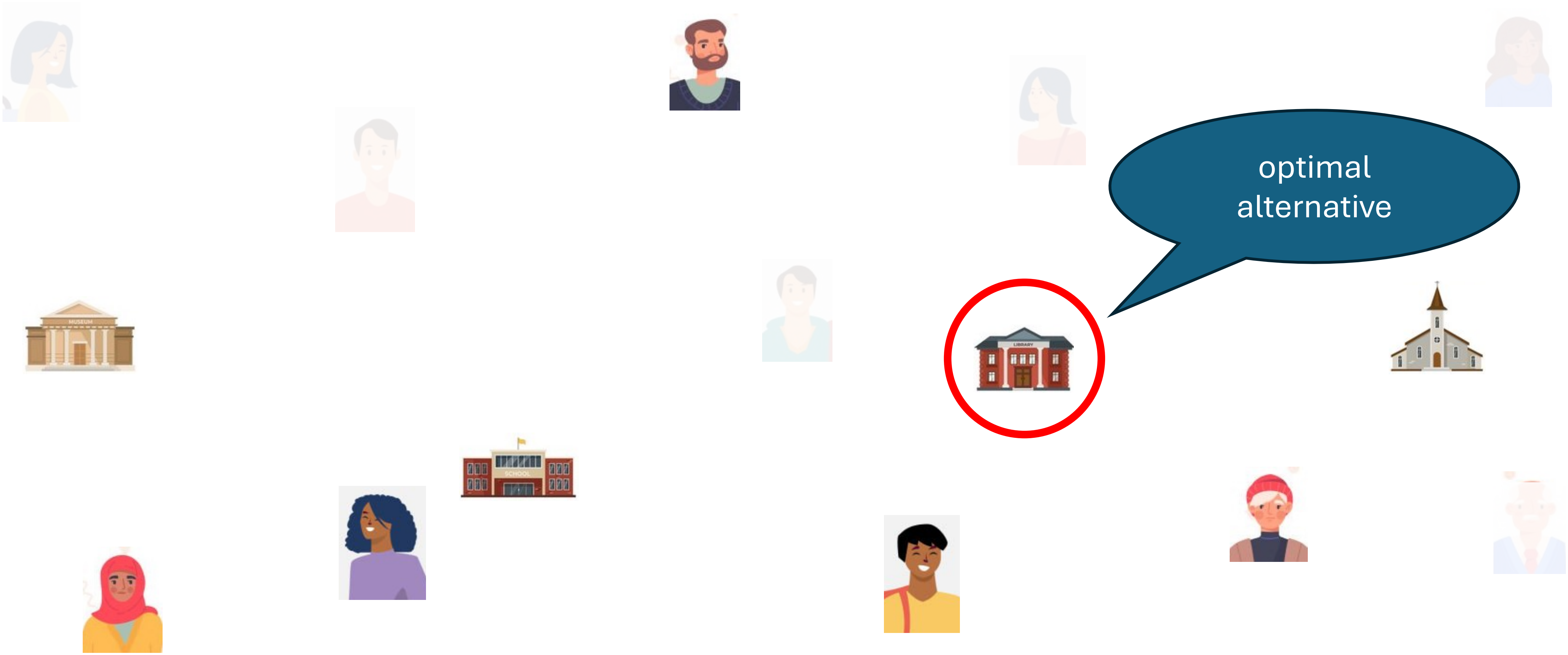
An example



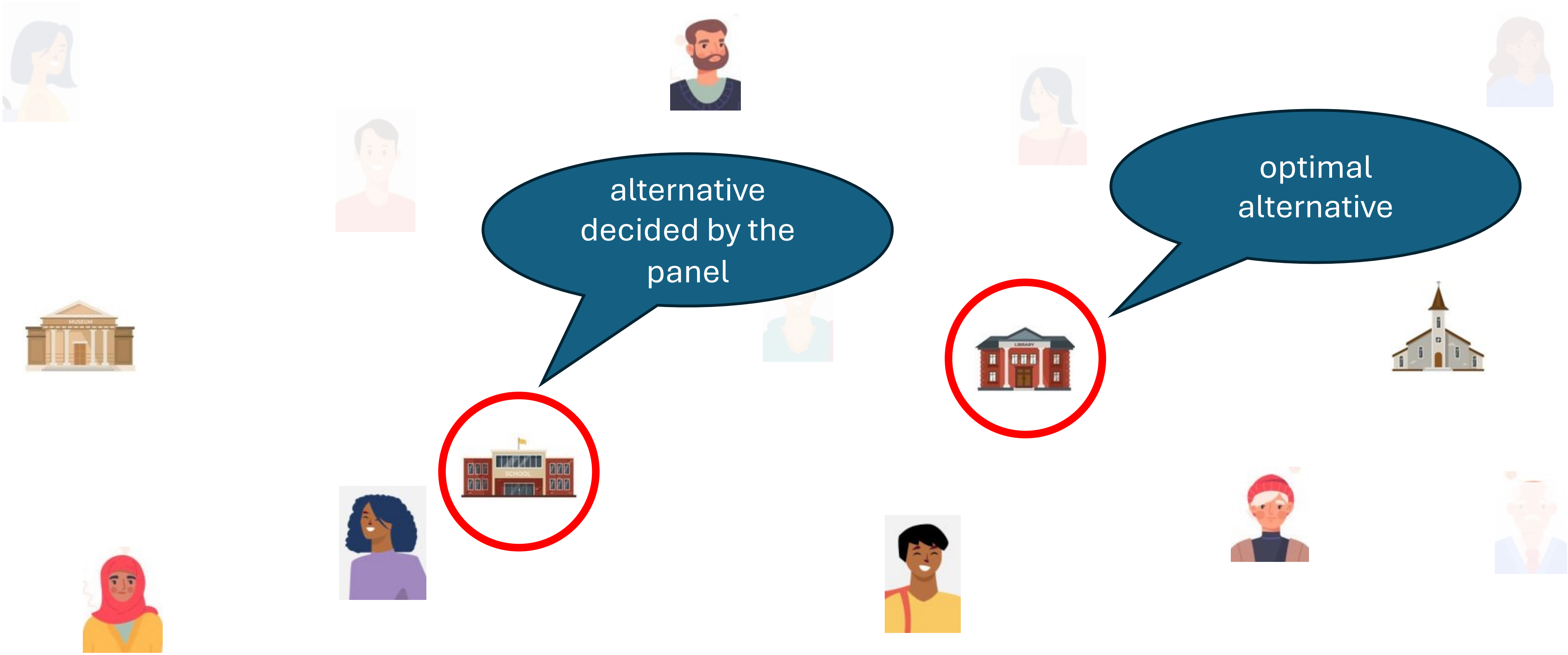
An example



An example



An example



Metric distortion

- Let F be a **randomized panel selection algorithm** (i.e., a probability distribution)
- The **ex ante metric distortion** of F for the set of m alternatives A is

$$eadist(F) = \frac{\mathbb{E}_{P\sim F}[SC(P(A))]}{SC^*(A)}$$

- The **ex post metric distortion** of F for the set of m alternatives A is

$$dist(F) = \frac{\max_{P\sim F} SC(P(A))}{SC^*(A)}$$

Fair selection algorithms

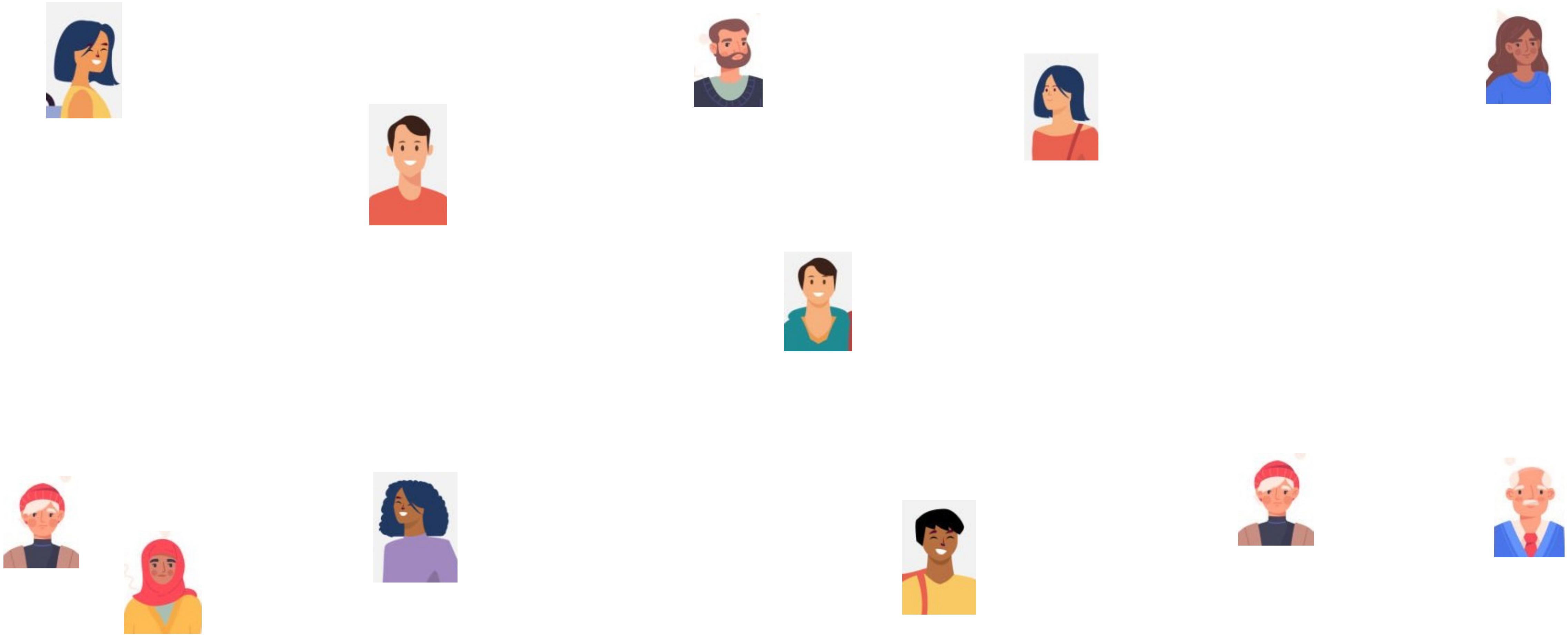
- **Fairness criterion**: the probability that a given agent is selected in the panel is k/n
- E.g., the algorithm that selects uniformly at random among all panels
- **Theorem**: The algorithm that selects uniformly at random among all panels of size $k = O(\varepsilon^{-2} \ln m)$ has ex ante metric distortion $1 + \varepsilon$ for every set of m alternatives
- Proof uses **Azuma's inequality** to handle minor correlations
- Drawback: **ex post distortion can be very high**

Fair greedy capture

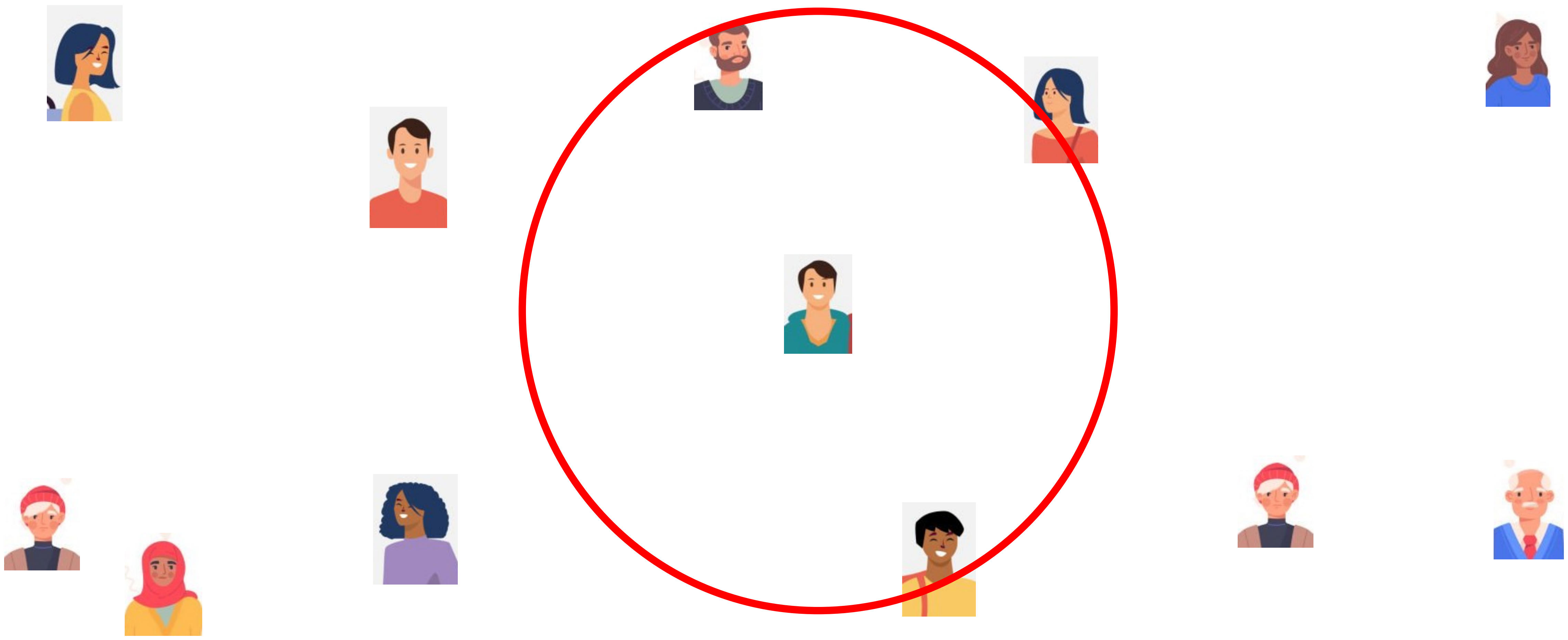
- Ebadian & Micha (2023)
- Variant of the **greedy capture** algorithm of Chen, Fain, Lyu, & Munagala (2019)

- Repeat k times:
  - Start growing balls centered at each agent, until some ball covers n/k new agents
  - Pick uniformly at random one agent from the n/k new ones covered and put her in the panel

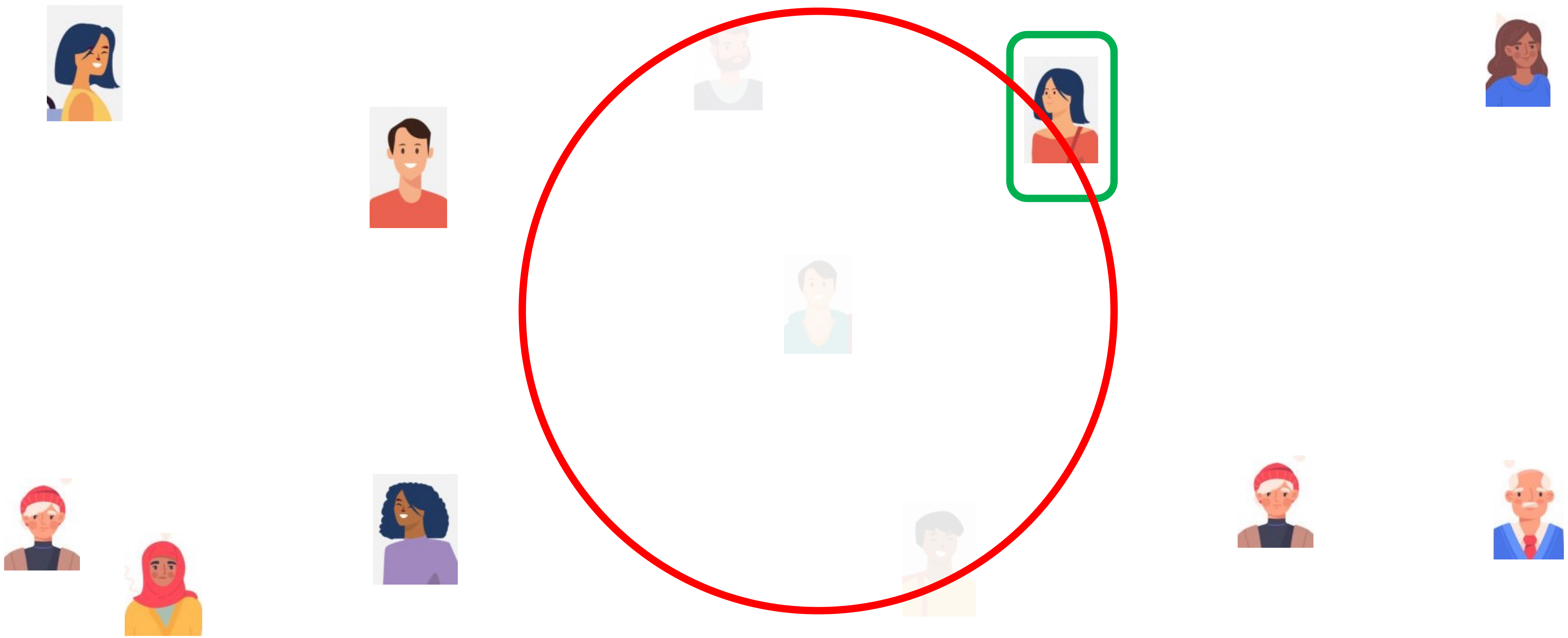
Fair greedy capture: an example ($n = 12, k = 3$)



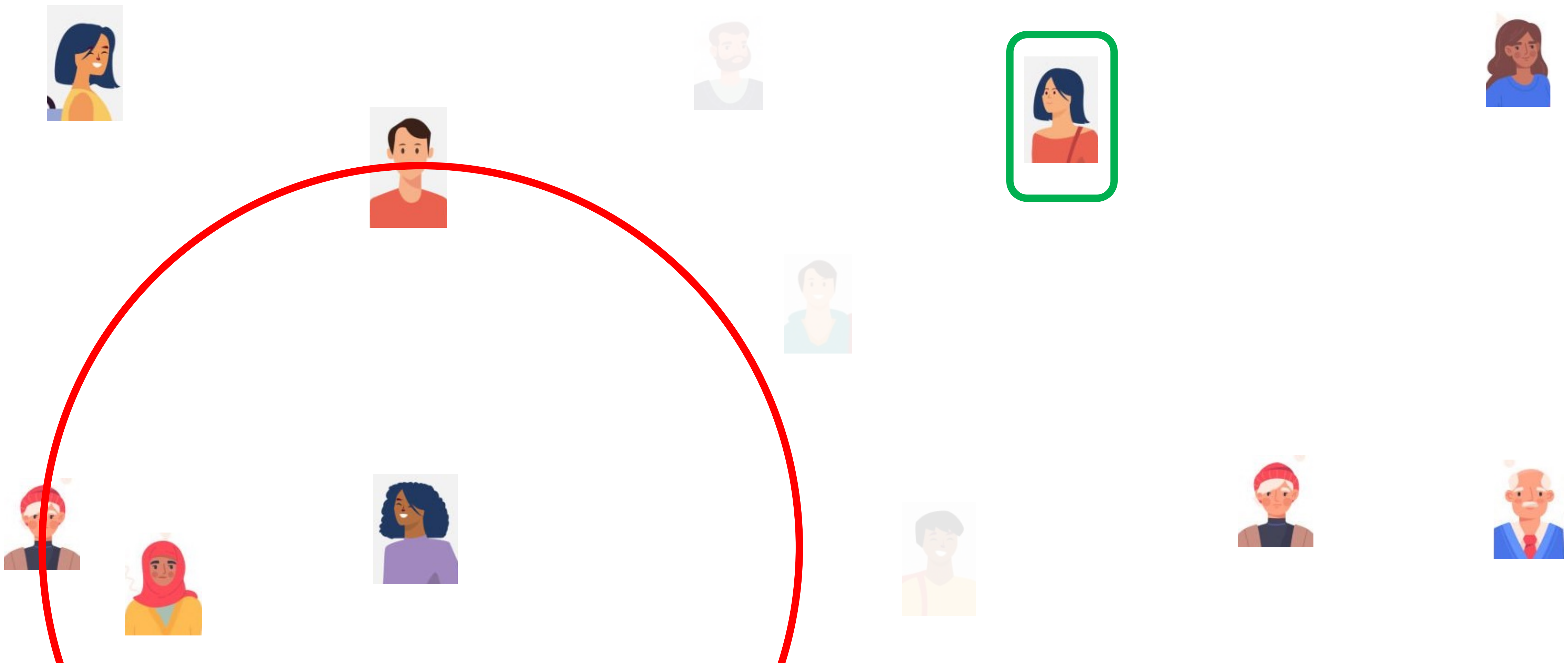
Fair greedy capture: an example ($n = 12, k = 3$)



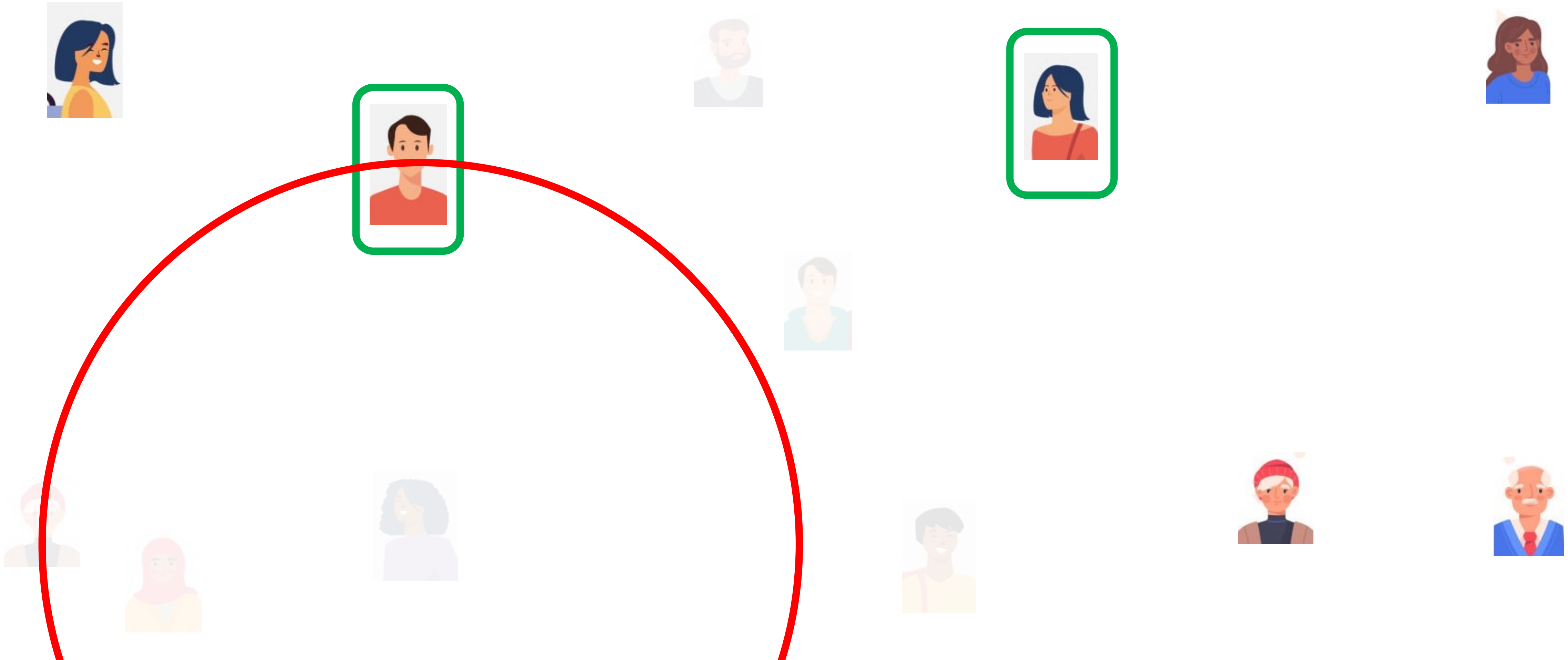
Fair greedy capture: an example ($n = 12, k = 3$)



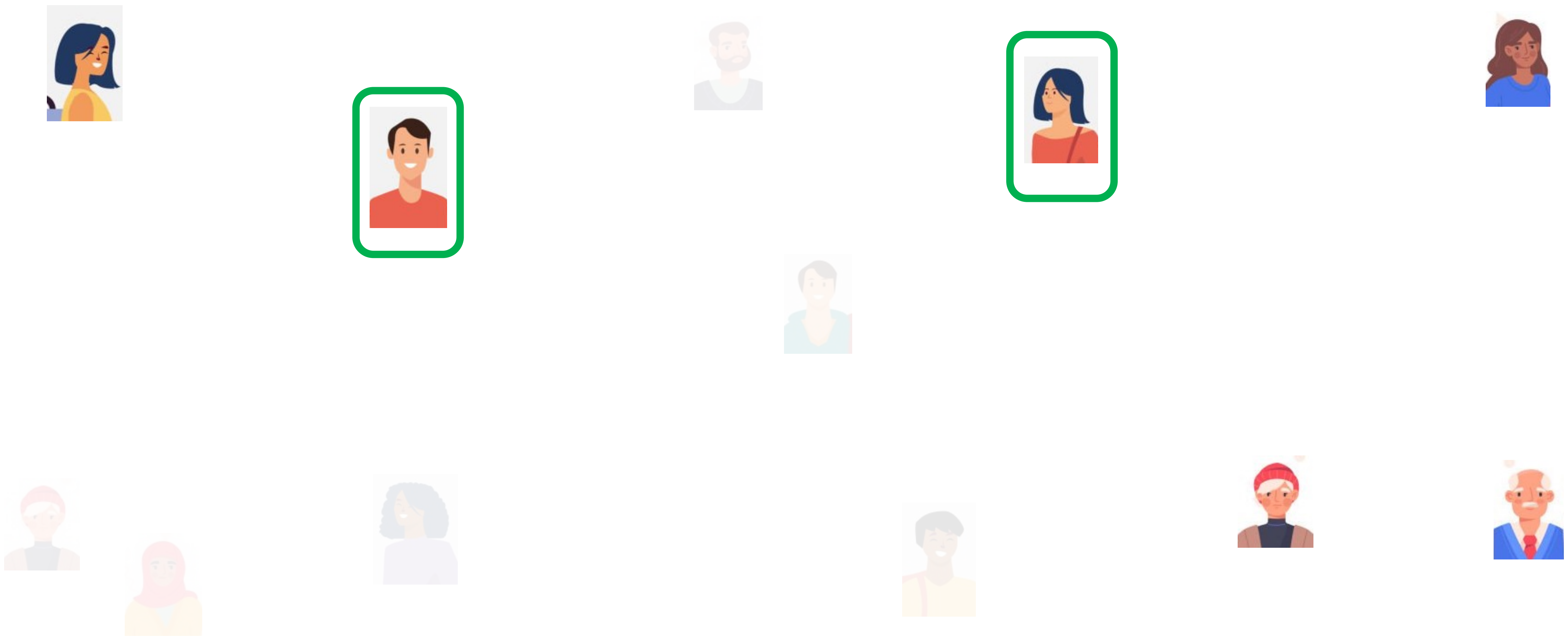
Fair greedy capture: an example ($n = 12, k = 3$)



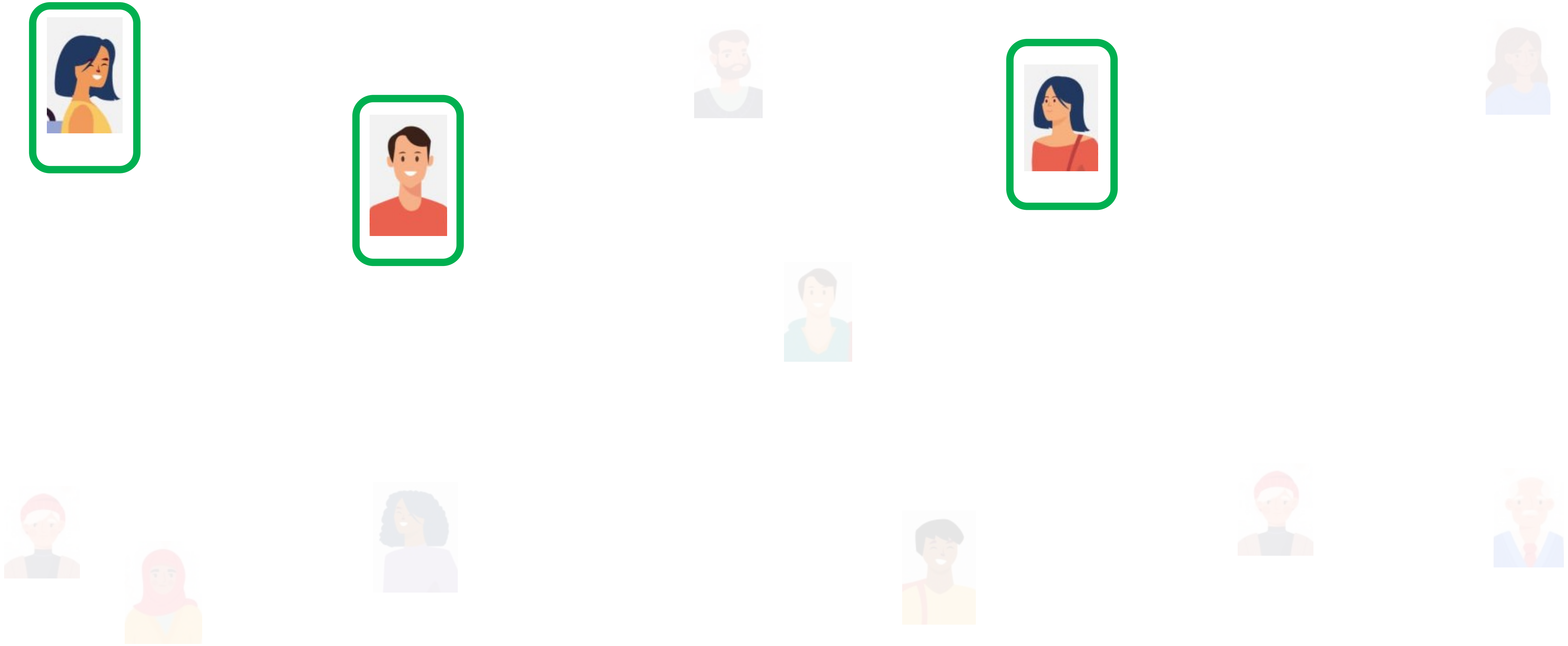
Fair greedy capture: an example ($n = 12, k = 3$)



Fair greedy capture: an example ($n = 12, k = 3$)



Fair greedy capture: an example ($n = 12, k = 3$)



Fair greedy capture

- Fair greedy capture is **fair**
- Theorem: For $k = O(\varepsilon^{-3} \ln m)$, fair greedy capture has **ex ante metric distortion $1 + \varepsilon$** and **constant ex post distortion** for every set of m alternatives
- Drawback: our current bound on ex post distortion is large (127)

End of the second story

- More results:
  - For small panels, **deterministic selection algorithms** cannot have (ex post) distortion better than 5
  - In contrast, **fair** selection algorithms have ex ante distortion at most 3
  - The panel size of $k = \Omega(\varepsilon^{-2} \ln m)$ is **optimal** for ex ante distortion $1 + \varepsilon$

- Open problems
  - **Better ex post distortion bounds** for fair greedy capture?
  - Improving the **dependency of the panel size on ε**?

Thank you!