Cryptographic Techniques to Ensure Fair Randomness in Legal Processes

Henry Corrigan-Gibbs and Keith Winstein
Department of Computer Science

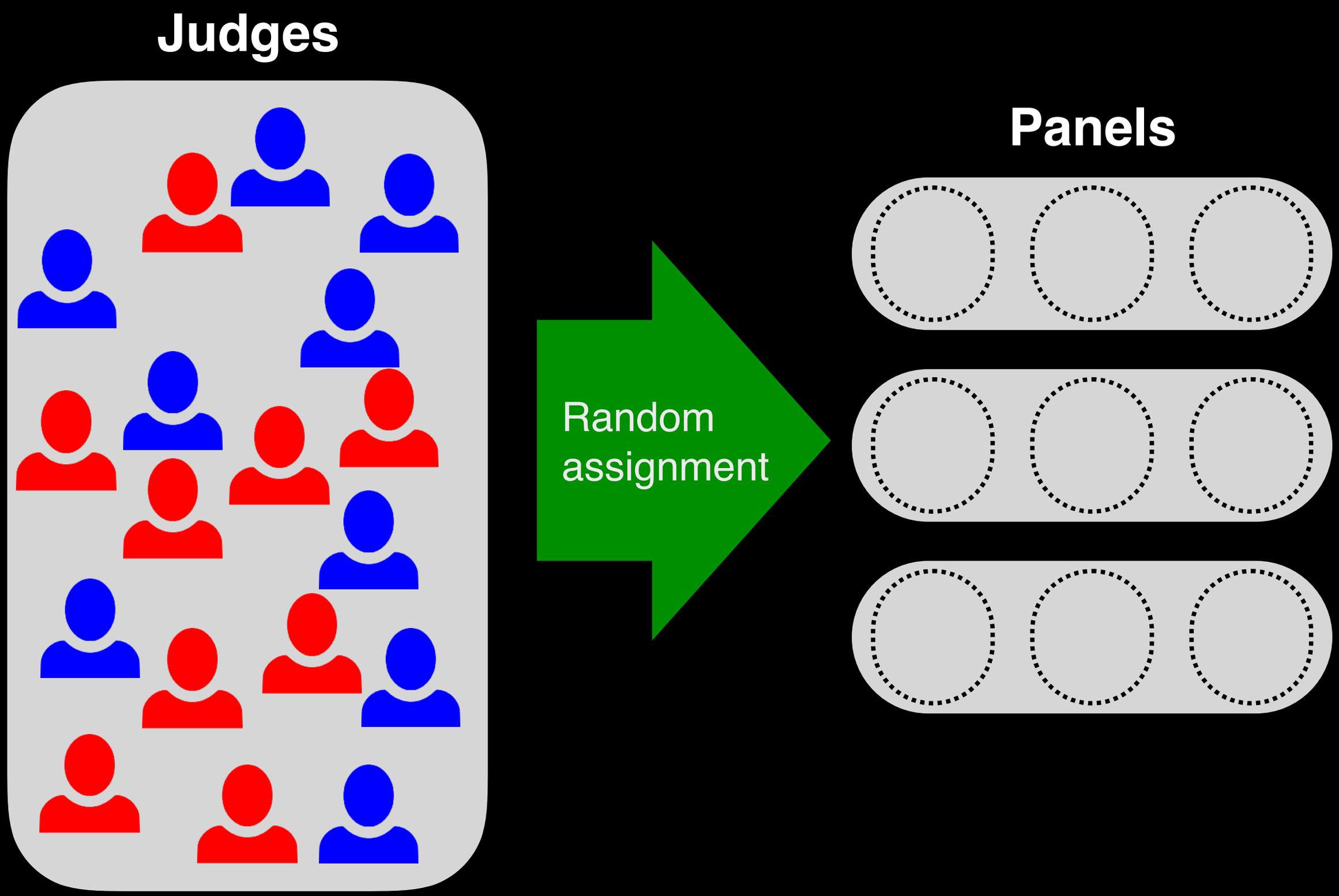
Coalition Challenges Selection of Judges in Same-Sex Marriage Case

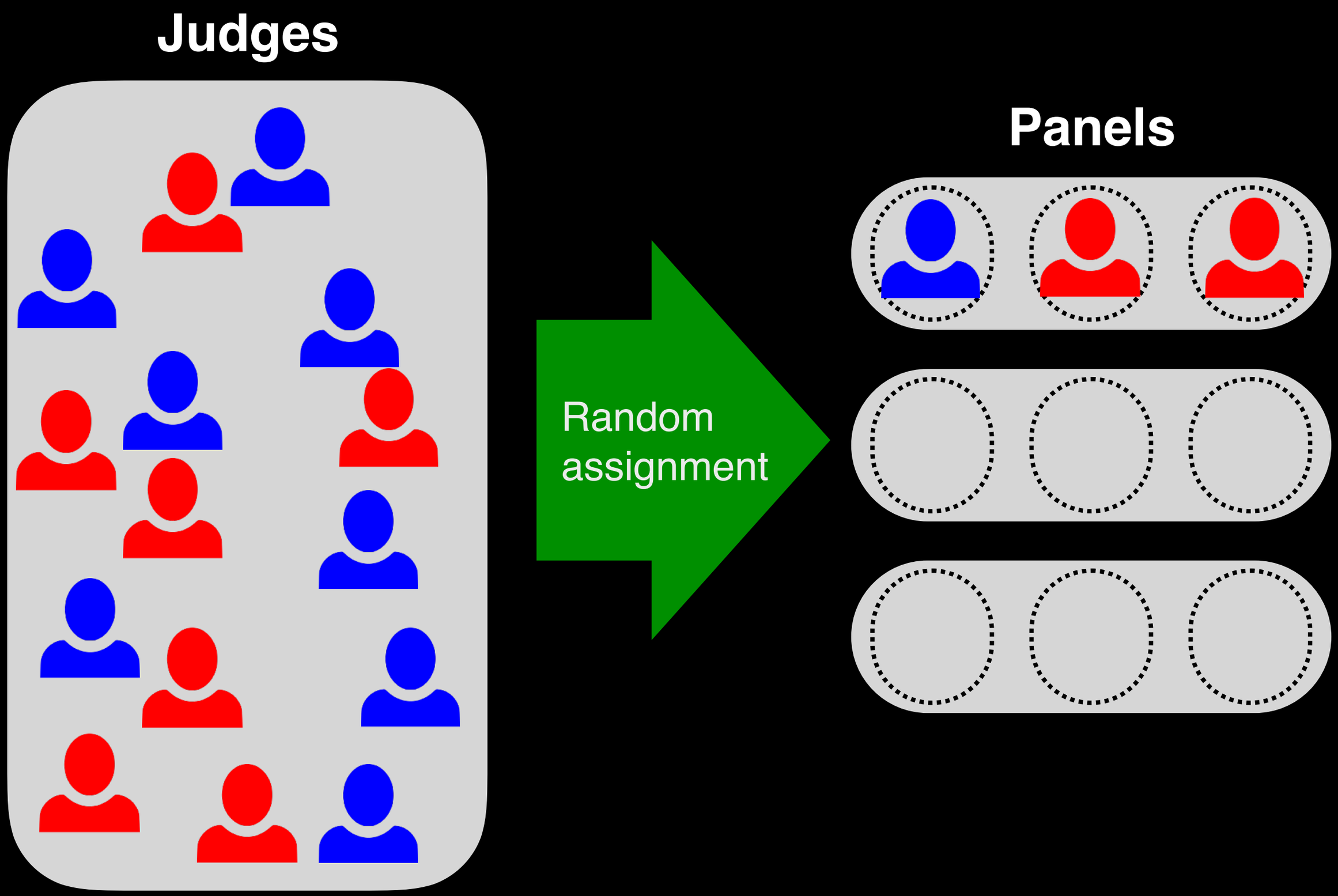
Adam Liptak

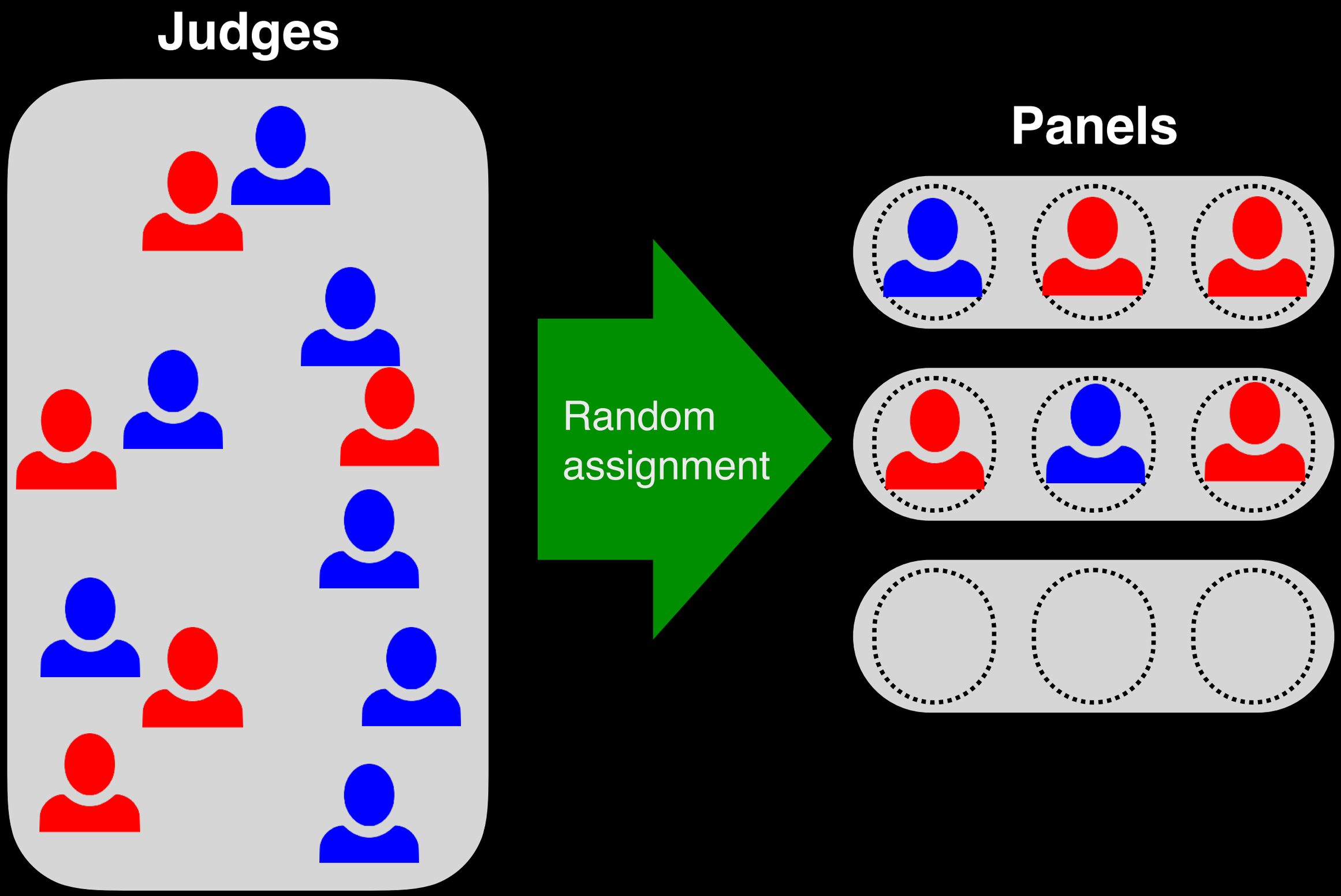
SIDEBAR NOV. 10, 2014

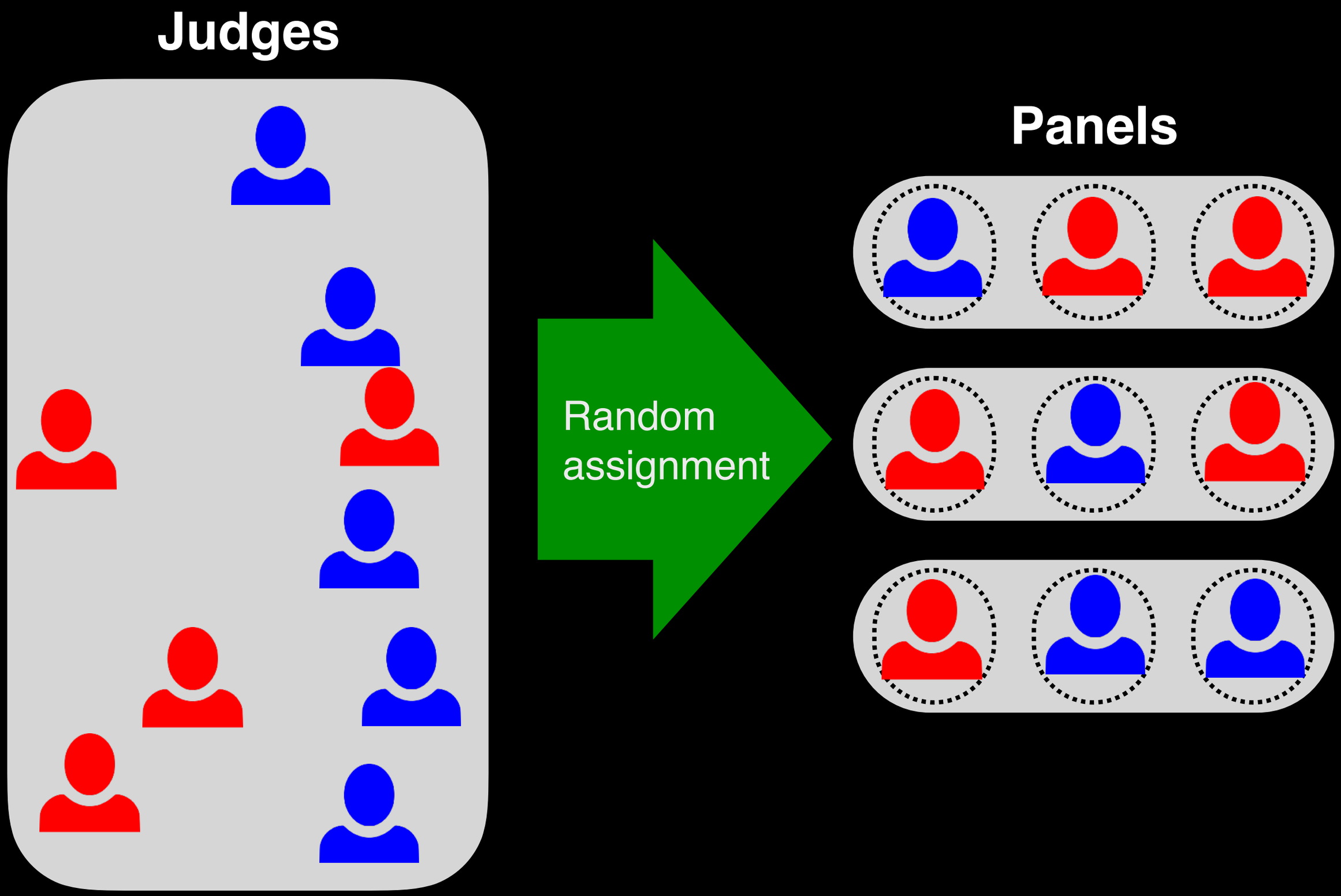
WASHINGTON — When a three-judge panel of the federal appeals court in California struck down bans last month on same-sex marriage in Idaho and Nevada, it was no surprise. The panel included two of the court's leading liberals.

A group opposing same-sex marriage said the composition of the panel was also no coincidence. In an unusual accusation in a recent court filing, the group said the two judges served on a disproportionate number of cases involving gay rights.


Judges
Random assignment
Panels


Judges
Random assignment
Panels


Judges
Random assignment
Panels


Judges
Random assignment
Panels

Coalition Challenges Selection of Judges in Same-Sex Marriage Case

Adam Liptak

SIDEBAR NOV. 10, 2014

WASHINGTON — When a three-judge panel of the federal appeals court in California struck down bans last month on same-sex marriage in Idaho and Nevada, it was no surprise. The panel included two of the court's leading liberals.

A group opposing same-sex marriage said the composition of the panel was also no coincidence. In an unusual accusation in a recent court filing, the group said the two judges served on a disproportionate number of cases involving gay rights.

Coalition Challenges Selection of Judges in Same-Sex Marriage Case

Adam Liptak

SIDEBAR NOV. 10, 2014

WASHINGTON — When a three-judge panel of the federal appeals court in California struck down bans last month on same-sex marriage in Idaho and Nevada, it was no surprise. The panel included two of the court's leading liberals.

A group opposing same-sex marriage said the composition of the panel was also no coincidence. In an unusual accusation in a recent court filing, the group said the two judges served on a disproportionate number of cases involving gay rights.

Coalition Challenges Selection of Judges in Same-Sex Marriage Case

Adam Liptak

SIDEBAR NOV. 10, 2014

WASHINGTON — When a three-judge panel of the federal appeals court in California struck down bans last month on same-sex marriage in Idaho and Nevada, it was no surprise. The panel included two of the court's leading liberals.

A group opposing same-sex marriage said the composition of the panel was also no coincidence. In an unusual accusation in a recent court filing, the group said the two judges served on a disproportionate number of cases involving gay rights.

At the same time, a new study by two law professors supports the idea that many federal appeals courts are not selecting their panels perfectly randomly.

"We found strong evidence in the majority of circuits that they're not using a random assignment process to form their panels," said Marin K. Levy, a law professor at Duke and one of the study's authors.

"We found strong evidence in the majority of circuits that they're not using a random assignment process to form their panels," said Marin K. Levy, a law professor at Duke and one of the study's authors.

Adam S. Chilton, a law professor at the University of Chicago and the study's other author, said there was only a two percent probability that the real panels were generated by pure chance.

"We found strong evidence in the majority of circuits that they're not using a random assignment process to form their panels," said Marin K. Levy, a law professor at Duke and one of the study's authors.

Adam S. Chilton, a law professor at the University of Chicago and the study's other author, said there was only a two percent probability that the real panels were generated by pure chance.

"If any of the 12 circuits are using a nonrandom process," he said, "it's most likely to be the Ninth Circuit."

Draft Document

Challenging the Randomness of Panel Assignment in the Federal Courts of Appeals

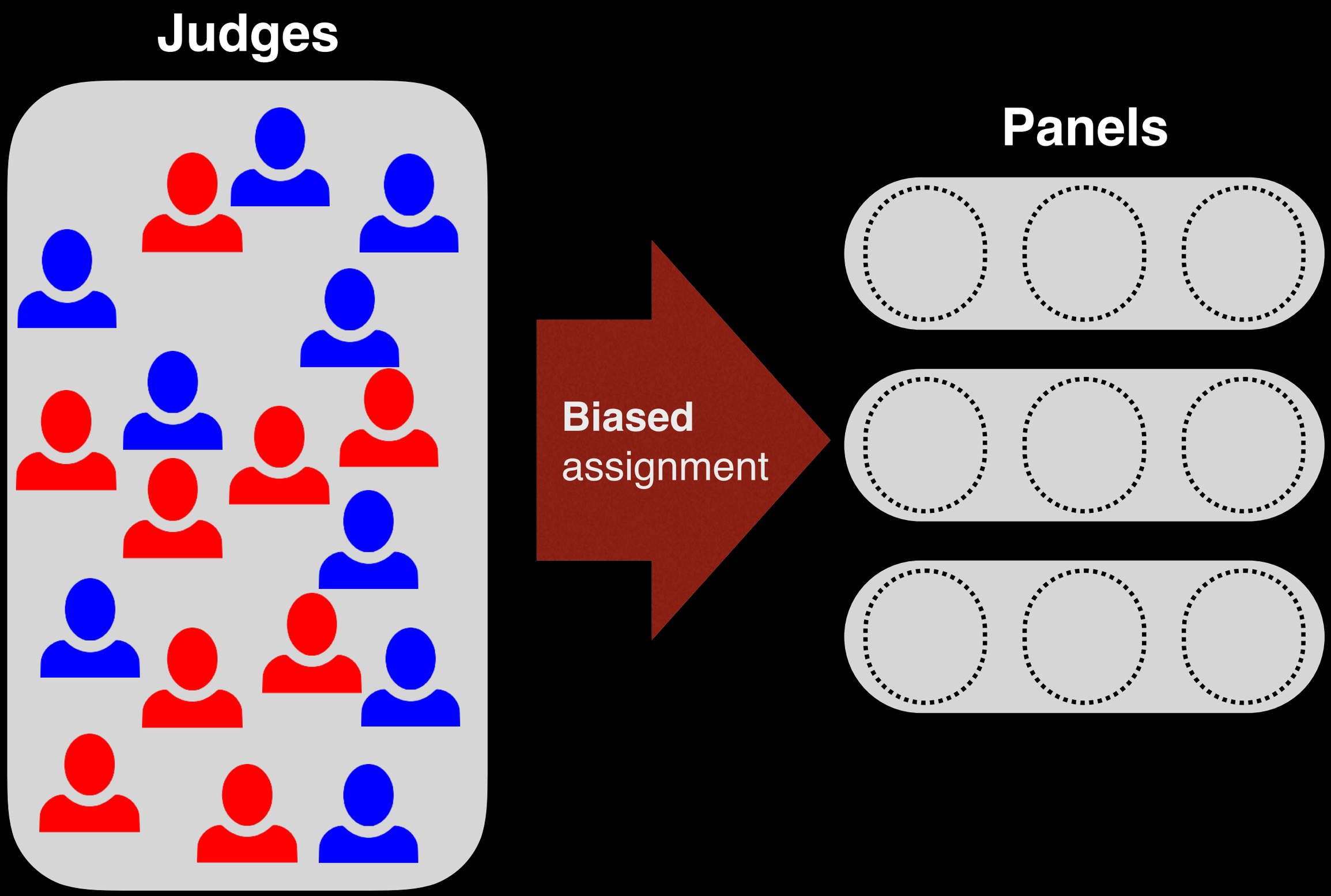
Adam S. Chilton[*] *& Marin K. Levy*[**]

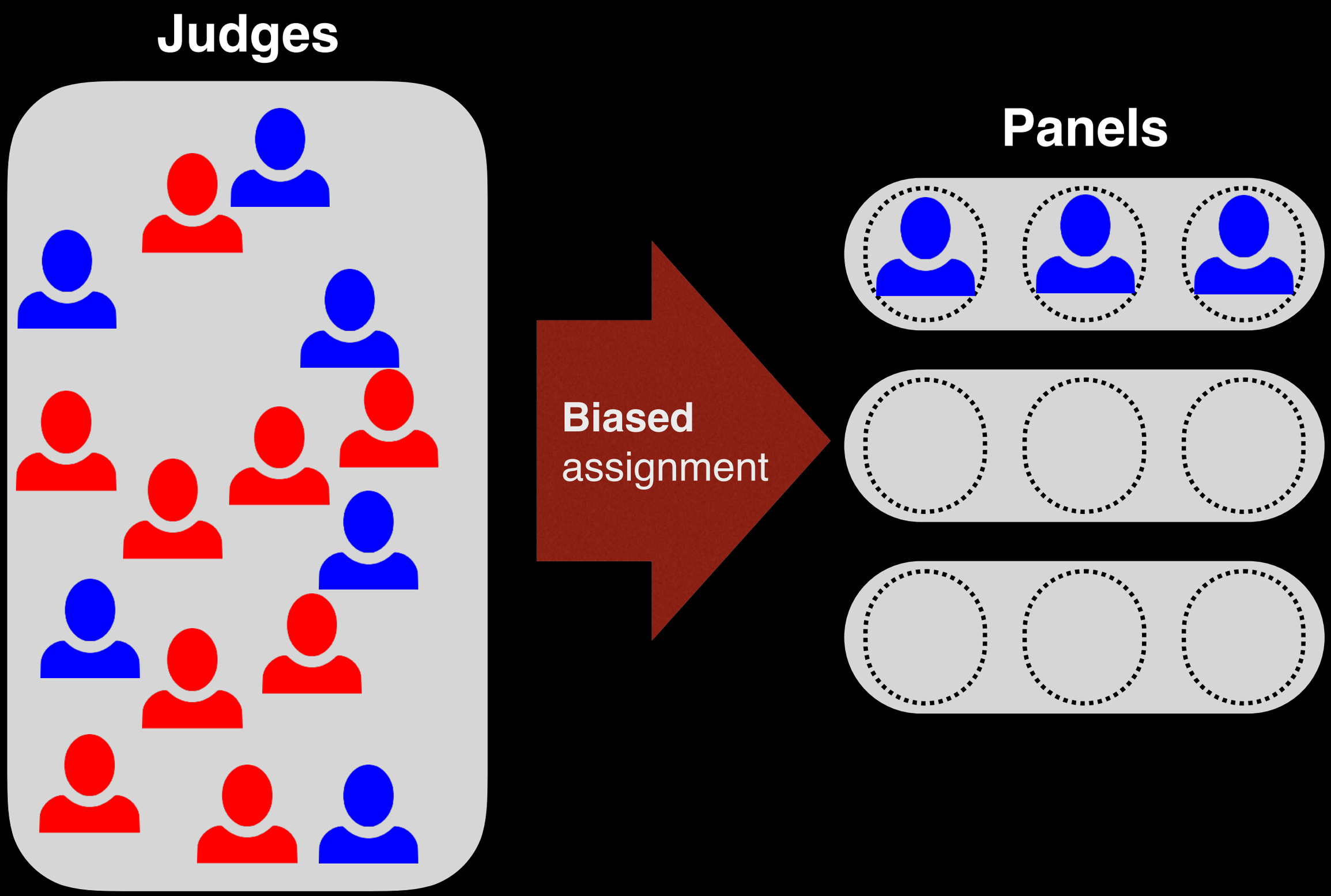
Abstract

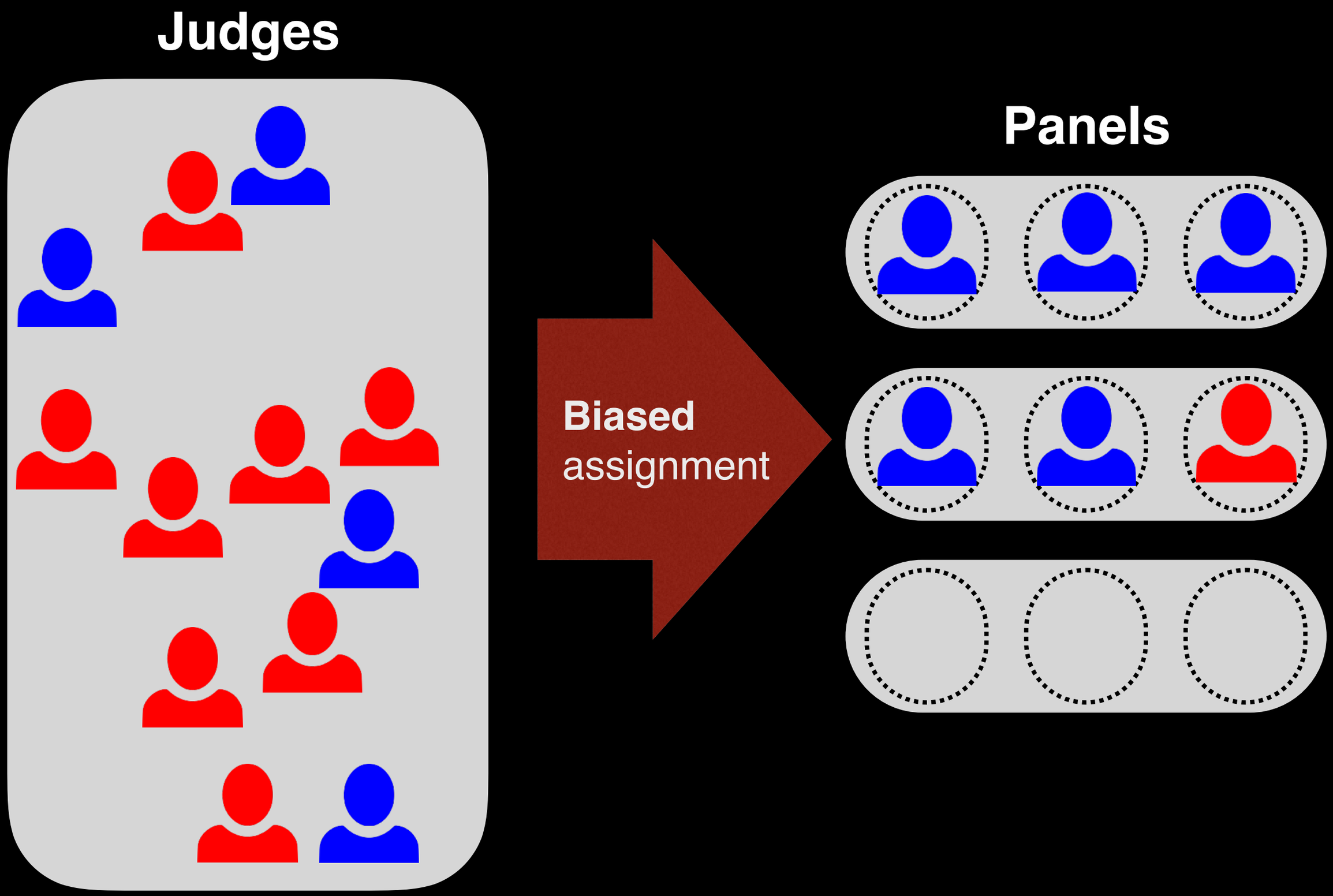
A fundamental academic assumption about the federal courts of appeals is that the three-judge panels that hear cases have been randomly configured. Scores of scholarly articles have noted this "fact," and it has been relied on heavily by empirical researchers. Even though there are practical reasons to doubt that judges would always be randomly assigned to panels, this assumption has never been tested. This Article fill this void by doing so.

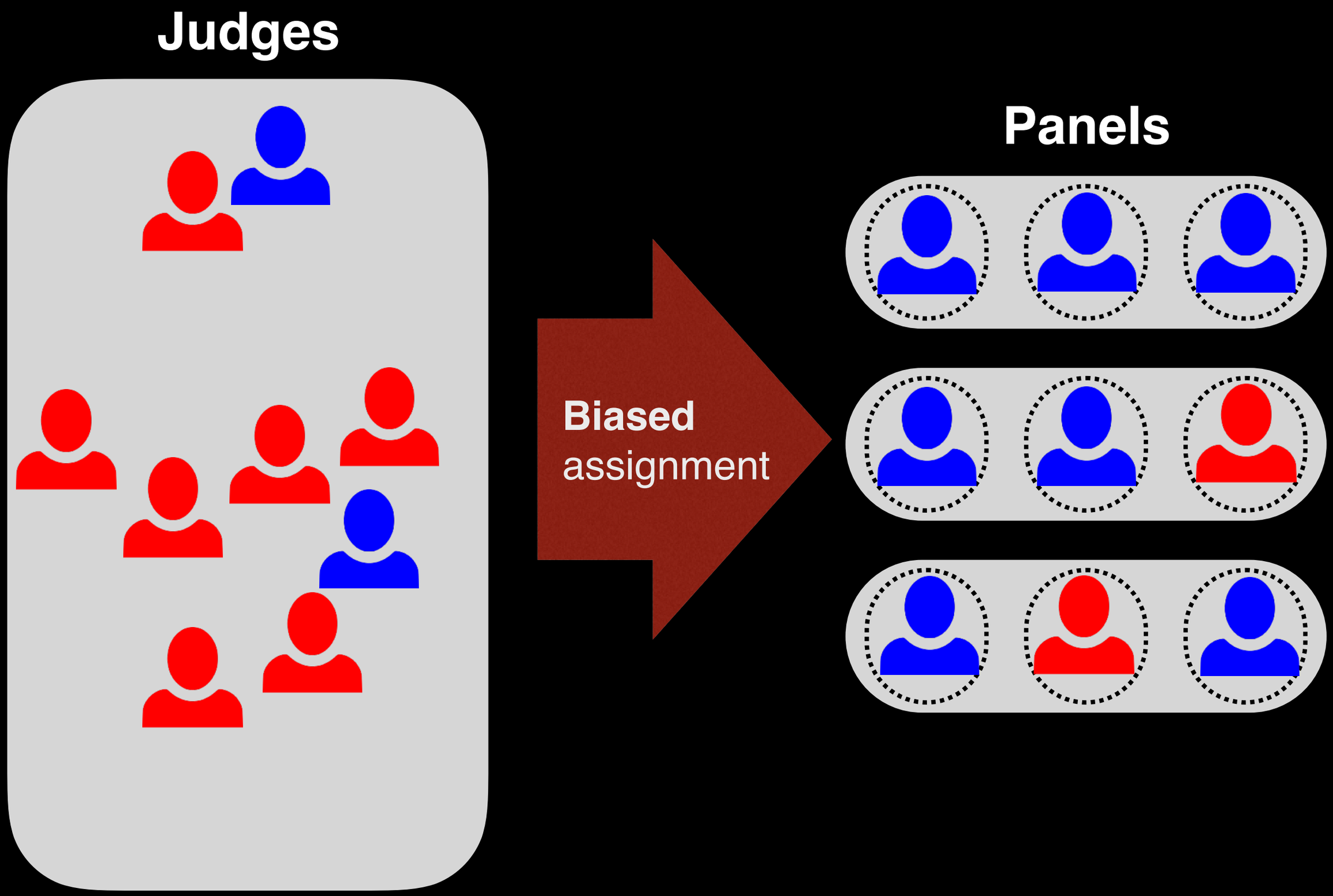
To determine whether the circuit courts utilize random assignment, we have created what we believe to be the largest dataset of panel assignments of those courts constructed to date. Using this dataset, we tested whether panel assignments are, in fact, random by comparing the actual assignments to truly random panels generated by code that we have created to simulate the panel generation process. Our results show evidence of non-randomness in the majority of the federal courts of appeals. Moreover, our findings specifically show a lack of randomness along several key dimensions: the distribution of conservative judges, female judges, and minority judges.

To be sure, the analysis here is descriptive, not explanatory or normative. We do not ourselves mean to suggest that "perfect randomness" is a desirable goal. We are simply testing an existing


Judges
Biased
assignment
Panels


Judges
Biased
assignment
Panels


Judges
Biased
assignment
Panels


Judges
Biased
assignment
Panels

Beyond panel assignments…

Beyond panel assignments…

Panel assignments are just one example of an important (but opaque) computerized government process:

Beyond panel assignments…

Panel assignments are just one example of an important (but opaque) computerized government process:

- visa lotteries, tax audits, airport screening, school lotteries, etc.

Beyond panel assignments…

Panel assignments are just one example of an important (but opaque) computerized government process:

visa lotteries, tax audits, airport screening, school lotteries, etc.

Risk: “Black-box” algorithms and processes may exhibit subtle biases and faults.

Beyond panel assignments…

Panel assignments are just one example of an important (but opaque) computerized government process:

visa lotteries, tax audits, airport screening, school lotteries, etc.

Risk: “Black-box” algorithms and processes may exhibit subtle biases and faults.

Hope: Cryptographic tools can make almost any algorithm open and accountable.

Plan

Plan

PART I: Have the federal Courts of Appeals been using biased randomness in their panel assignments?

Plan

PART I: Have the federal Courts of Appeals been using biased randomness in their panel assignments?

PART II: How can we ensure that the government faithfully executes a random process?

Plan

PART I: Have the federal Courts of Appeals been using biased randomness in their panel assignments?

The study has two major design flaws.
These flaws **invalidate the study's claims**.

PART II: How can we ensure that the government faithfully executes a random process?

Plan

PART I: Have the federal Courts of Appeals been using biased randomness in their panel assignments?

The study has two major design flaws.
These flaws **invalidate the study's claims**.

PART II: How can we ensure that the government faithfully executes a random process?

These cryptographic techniques also apply to:
visa lotteries, tax audits, airport screening, school lotteries, etc.

PART I

The Perils of Retrospective Analysis

PART II

Cryptographic Techniques to Ensure Fair Randomness

PART I

The Perils of Retrospective Analysis

PART II

Cryptographic Techniques to Ensure Fair Randomness

Draft Document

Challenging the Randomness of Panel Assignment in the Federal Courts of Appeals

Adam S. Chilton[*] *& Marin K. Levy*[**]

Abstract

A fundamental academic assumption about the federal courts of appeals is that the three-judge panels that hear cases have been randomly configured. Scores of scholarly articles have noted this "fact," and it has been relied on heavily by empirical researchers. Even though there are practical reasons to doubt that judges would always be randomly assigned to panels, this assumption has never been tested. This Article fill this void by doing so.

To determine whether the circuit courts utilize random assignment, we have created what we believe to be the largest dataset of panel assignments of those courts constructed to date. Using this dataset, we tested whether panel assignments are, in fact, random by comparing the actual assignments to truly random panels generated by code that we have created to simulate the panel generation process. Our results show evidence of non-randomness in the majority of the federal courts of appeals. Moreover, our findings specifically show a lack of randomness along several key dimensions: the distribution of conservative judges, female judges, and minority judges.

To be sure, the analysis here is descriptive, not explanatory or normative. We do not ourselves mean to suggest that "perfect randomness" is a desirable goal. We are simply testing an existing

Study Methods

Study Methods

1. **Prepare** a set of 144 hypotheses to test.

   e.g., H59: *The number of all-female panels in the 2nd Circuit is consistent with a random assignment.*

Study Methods

1. **Prepare** a set of 144 hypotheses to test.
   e.g., H59: *The number of all-female panels in the 2nd Circuit is consistent with a random assignment.*

2. **Collect** a data set of the actual panel assignments in 12 U.S. Courts of Appeals.

Study Methods

1. **Prepare** a set of 144 hypotheses to test.
   e.g., H59: *The number of all-female panels in the 2nd Circuit is consistent with a random assignment.*

2. **Collect** a data set of the actual panel assignments in 12 U.S. Courts of Appeals.

3. **Test** how likely the actual panel assignments would be under a truly random (ideal) assignment.

Appendix B:
Simulated Distribution of Female Judges

	Actual Panels	Simulated Mean	90% Confidence Interval		Significance
D.C. Circuit					
0 Women	158	166	149	184	
1 Women	272	261	242	279	
2 Women	77	79	66	93	
3 Women	3	4	1	7	
1st Circuit					
0 Women	131	137	123	150	
1 Women	148	140	126	154	
2 Women	18	21	14	29	
3 Women	1	0	0	1	
2nd Circuit					
0 Women	558	570	540	600	
1 Women	608	594	564	624	
2 Women	155	160	141	180	
3 Women	11	11	6	16	
3rd Circuit					
0 Women	311	296	275	317	
1 Women	263	297	275	318	***
2 Women	93	72	59	86	**
3 Women	2	4	1	8	
4th Circuit					
0 Women	263	263	243	283	
1 Women	266	268	248	288	
2 Women	67	62	50	75	
3 Women	1	3	1	6	*
5th Circuit					
0 Women	271	278	256	301	
1 Women	422	419	394	443	
2 Women	176	164	145	183	
3 Women	8	16	10	23	**
6th Circuit					
0 Women	391	392	366	419	
1 Women	566	563	534	591	
2 Women	219	226	204	248	
3 Women	30	25	17	33	

Problem 1: Multiple Hypothesis Testing

Problem 1: Multiple Hypothesis Testing

- When testing 144 hypotheses at the $p < 0.1$ level, we expect 14 “false discoveries” *by random chance alone*

Problem 1: Multiple Hypothesis Testing

- When testing 144 hypotheses at the $p < 0.1$ level, we expect 14 “false discoveries” *by random chance alone*

- Even if <u>all courts</u> are using good randomness, C-L will falsely accuse eight circuits of bad behavior on average.

Why can multiple hypothesis testing be problematic?

Why can multiple hypothesis testing be problematic?

By analogy… Say that every day, every state lottery picks a number between 1 and 50.

Why can multiple hypothesis testing be problematic?

By analogy… Say that every day, every state lottery picks a number between 1 and 50.

Why can multiple hypothesis testing be problematic?

By analogy… Say that every day, every state lottery picks a number between 1 and 50.

Alabama	2	Hawaii	12	Mass.	38	N.M.	39	S.D.	11
Alaska	32	Idaho	13	Michigan	12	New York	27	Tenn.	6
Arizona	17	Illinois	13	Minn.	9	N.C.	42	Texas	32
Arkansas	43	Indiana	46	Miss.	15	N.D.	12	Utah	29
California	6	Iowa	18	Missouri	50	Ohio	17	Vermont	2
Colorado	4	Kansas	39	Montana	8	Okla.	47	Virginia	12
Conn.	12	Kentucky	17	Nebraska	27	Oregon	23	Wash.	8
Delaware	43	Louisiana	1	Nevada	31	Penn.	18	W. Va.	30
Florida	32	Maine	34	N.H.	47	R.I.	5	Wisc.	9
Georgia	48	Maryland	11	N.J.	7	S.C.	2	Wyoming	7

Alabama	2	Hawaii	12	Mass.	38	N.M.	39	S.D.	11
Alaska	32	Idaho	13	Michigan	12	New York	27	Tenn.	6
Arizona	17	Illinois	13	Minn.	9	N.C.	42	Texas	32
Arkansas	43	Indiana	46	Miss.	15	N.D.	12	Utah	29
California	6	Iowa	18	Missouri	50	Ohio	17	Vermont	2
Colorado	4	Kansas	39	Montana	8	Okla.	47	Virginia	12
Conn.	12	Kentucky	17	Nebraska	27	Oregon	23	Wash.	8
Delaware	43	Louisiana	1	Nevada	31	Penn.	18	W. Va.	30
Florida	32	Maine	34	N.H.	47	R.I.	5	Wisc.	9
Georgia	48	Maryland	11	N.J.	7	S.C.	2	Wyoming	7

Alabama	2	Hawaii	12	Mass.	38	N.M.	39	S.D.	11
Alaska	32	Idaho	13	Michigan	12	New York	27	Tenn.	6
Arizona	17	Illinois	13	Minn.	9	N.C.	42	Texas	32
Arkansas	43	Indiana	46	Miss.	15	N.D.	12	Utah	29
California	6	Iowa	18	Missouri	50	Ohio	17	Vermont	2
Colorado	4	Kansas	39	Montana	8	Okla.	47	Virginia	12
Conn.	12	Kentucky	17	Nebraska	27	Oregon	23	Wash.	8
Delaware	43	Louisiana	1	Nevada	31	Penn.	18	W. Va.	30
Florida	32	Maine	34	N.H.	47	R.I.	5	Wisc.	9
Georgia	48	Maryland	11	N.J.	7	S.C.	2	Wyoming	7

Scientist Discovers Evidence of Fraud in Iowa and Pennsylvania State Lotteries

Alabama	2	Hawaii	12	Mass.	38	N.M.	39	S.D.	11
Alaska	32	Idaho	13	Michigan	12	New York	27	Tenn.	6
Arizona	17	Illinois	13	Minn.	9	N.C.	42	Texas	32
Arkansas	43	Indiana	46	Miss.	15	N.D.	12	Utah	29
California	6	Iowa	18	Missouri	50	Ohio	17	Vermont	2
Colorado	4	Kansas	39	Montana	8	Okla.	47	Virginia	12
Conn.	12	Kentucky	17	Nebraska	27	Oregon	23	Wash.	8
Delaware	43	Louisiana	1	Nevada	31	Penn.	18	W. Va.	30
Florida	32	Maine	34	N.H.	47	R.I.	5	Wisc.	9
Georgia	48	Maryland	11	N.J.	7	S.C.	2	Wyoming	7

Alabama	2	Hawaii	12	Mass.	38	N.M.	39	S.D.	11
Alaska	32	Idaho	13	Michigan	12	New York	27	Tenn.	6
Arizona	17	Illinois	13	Minn.	9	N.C.	42	Texas	32
Arkansas	43	Indiana	46	Miss.	15	N.D.	12	Utah	29
California	6	Iowa	18	Missouri	50	Ohio	17	Vermont	2
Colorado	4	Kansas	39	Montana	8	Okla.	47	Virginia	12
Conn.	12	Kentucky	17	Nebraska	27	Oregon	23	Wash.	8
Delaware	43	Louisiana	1	Nevada	31	Penn.	18	W. Va.	30
Florida	32	Maine	34	N.H.	47	R.I.	5	Wisc.	9
Georgia	48	Maryland	11	N.J.	7	S.C.	2	Wyoming	7

Alabama	2	Hawaii	12	Mass.	38	N.M.	39	S.D.	11
Alaska	32	Idaho	13	Michigan	12	New York	27	Tenn.	6
Arizona	17	Illinois	13	Minn.	9	N.C.	42	Texas	32
Arkansas	43	Indiana	46	Miss.	15	N.D.	12	Utah	29
California	6	Iowa	18	Missouri	50	Ohio	17	Vermont	2
Colorado	4	Kansas	39	Montana	8	Okla.	47	Virginia	12
Conn.	12	Kentucky	17	Nebraska	27	Oregon	23	Wash.	8
Delaware	43	Louisiana	1	Nevada	31	Penn.	18	W. Va.	30
Florida	32	Maine	34	N.H.	47	R.I.	5	Wisc.	9
Georgia	48	Maryland	11	N.J.	7	S.C.	2	Wyoming	7

Alabama	2	Hawaii	12	Mass.	38	N.M.	39	S.D.	11
Alaska	32	Idaho	13	Michigan	12	New York	27	Tenn.	6
Arizona	17	Illinois	13	Minn.	9	N.C.	42	Texas	32
Arkansas	43	Indiana	46	Miss.	15	N.D.	12	Utah	29
California	6	Iowa	18	Missouri	50	Ohio	17	Vermont	2
Colorado	4	Kansas	39	Montana	8	Okla.	47	Virginia	12
Conn.	12	Kentucky	17	Nebraska	27	Oregon	23	Wash.	8
Delaware	43	Louisiana	1	Nevada	31	Penn.	18	W. Va.	30
Florida	32	Maine	34	N.H.	47	R.I.	5	Wisc.	9
Georgia	48	Maryland	11	N.J.	7	S.C.	2	Wyoming	7

Wait a minute…

Even if the lottery *weren't rigged*, we would expect to see a few **18**s just by random chance.

Problem 1: Multiple Hypothesis Testing

Problem 1: Multiple Hypothesis Testing

The Courts of Appeals study falls into the same trap:

Problem 1: Multiple Hypothesis Testing

The Courts of Appeals study falls into the same trap:

The authors run a large number of statistical tests (144) and find 21 significant results.

Problem 1: Multiple Hypothesis Testing

The Courts of Appeals study falls into the same trap:

The authors run a large number of statistical tests (144) and find 21 significant results.

- The probability that we would see these results due to random chance is > 5%.

Problem 1: Multiple Hypothesis Testing

The Courts of Appeals study falls into the same trap:

The authors run a large number of statistical tests (144) and find 21 significant results.

- The probability that we would see these results due to random chance is > 5%.

"[T]here is evidence that the ideological balance of panels is non-random in five circuits: the D.C. Circuit, the Second Circuit, the Fourth Circuit, the Eight Circuit, and the Ninth Circuit."

Problem 1: Multiple Hypothesis Testing

The Courts of Appeals study falls into the same trap:

The authors run a large number of statistical tests (144) and find 21 significant results.

- The probability that we would see these results due to random chance is > 5%.

"[T]here is evidence that the ideological balance of panels is non-random in five circuits: the D.C. Circuit, the Second Circuit, the Fourth Circuit, the Eight Circuit, and the Ninth Circuit."

- As in the lotto example: we expect false discoveries due to chance alone. We expect <u>more than four</u> false discoveries, the study finds five.

Problem 1: Multiple Hypothesis Testing

The Courts of Appeals study falls into the same trap:

The authors run a large number of statistical tests (144) and find 21 significant results.

- The probability that we would see these results due to random chance is > 5%.

"[T]here is evidence that the ideological balance of panels is non-random in five circuits: the D.C. Circuit, the Second Circuit, the Fourth Circuit, the Eight Circuit, and the Ninth Circuit."

- As in the lotto example: we expect false discoveries due to chance alone. We expect <u>more than four</u> false discoveries, the study finds five.

"[T]here was evidence of non-randomness in **seven** of the twelve region circuits"

Problem 1: Multiple Hypothesis Testing

The Courts of Appeals study falls into the same trap:

The authors run a large number of statistical tests (144) and find 21 significant results.

- The probability that we would see these results due to random chance is > 5%.

"[T]here is evidence that the ideological balance of panels is non-random in five circuits: the D.C. Circuit, the Second Circuit, the Fourth Circuit, the Eight Circuit, and the Ninth Circuit."

- As in the lotto example: we expect false discoveries due to chance alone. We expect <u>more than four</u> false discoveries, the study finds five.

"[T]here was evidence of non-randomness in **seven** of the twelve region circuits"

- By chance alone, we would expect to find such statistical deviations in <u>more than eight</u> circuits.

Problem 1: Multiple Hypothesis Testing

The Courts of Appeals study falls into the same trap:

The authors run a large number of statistical tests (144) and find 21 significant results.

- Th
  du

Random chance explains all of the results in the study.

"[T]here dom in five circ e Eight Circuit,

- As ce al ds five.

"[T]here was evidence of non-randomness in **seven** of the twelve region circuits"

- By chance alone, we would expect to find such statistical deviations in more than eight circuits.

Problem 2: “Cherry-Picking” Hypotheses

Problem 2: “Cherry-Picking” Hypotheses

Nov 10, 2014 *New York Times* article appears.

Problem 2: “Cherry-Picking” Hypotheses

Nov 10, 2014	*New York Times* article appears.
Nov 12	We explain our concerns about the study design to the authors via email.

Problem 2: “Cherry-Picking” Hypotheses

Nov 10, 2014	*New York Times* article appears.
Nov 12	We explain our concerns about the study design to the authors via email.
Nov 17	We post a note to SSRN explaining the problems with the study’s methods.

Problem 2: “Cherry-Picking” Hypotheses

Nov 10, 2014	*New York Times* article appears.
Nov 12	We explain our concerns about the study design to the authors via email.
Nov 17	We post a note to SSRN explaining the problems with the study’s methods.
Jan 3, 2015	The authors send us a revised draft for comment.

Problem 2: “Cherry-Picking” Hypotheses

Nov 10, 2014	*New York Times* article appears.
Nov 12	We explain our concerns about the study design to the authors via email.
Nov 17	We post a note to SSRN explaining the problems with the study’s methods.
Jan 3, 2015	The authors send us a revised draft for comment.

- In the revising the paper, the authors altered their statistical tests and **discarded two thirds of their hypotheses**.

Problem 2: “Cherry-Picking” Hypotheses

Nov 10, 2014	*New York Times* article appears.
Nov 12	We explain our concerns about the study design to the authors via email.
Nov 17	We post a note to SSRN explaining the problems with the study’s methods.
Jan 3, 2015	The authors send us a revised draft for comment.

- In the revising the paper, the authors altered their statistical tests and **discarded two thirds of their hypotheses**.
- They discarded the hypotheses that the data set didn’t support (i.e., with the weakest *p*-values).

Problem 2: “Cherry-Picking” Hypotheses

Nov 10, 2014	*New York Times* article appears.
Nov 12	We explain our concerns about the study design to the authors via email.
Nov 17	We post a note to SSRN explaining the problems with the study’s methods.
Jan 3, 2015	The authors send us a revised draft for comment.

- In the revising the paper, the authors altered their statistical tests and **discarded two thirds of their hypotheses**.
- They discarded the hypotheses that the data set didn’t support (i.e., with the weakest *p*-values).

Problem: Editing the hypotheses after running the statistical tests can lead to spurious findings (“*p*-hacking”).

Problem 2: “Cherry-Picking” Hypotheses

Example: Say you run a medical trial comparing Aspirin vs New Drug X.

Problem 2: “Cherry-Picking” Hypotheses

Example: Say you run a medical trial comparing Aspirin vs New Drug X.

Problem 2: “Cherry-Picking” Hypotheses

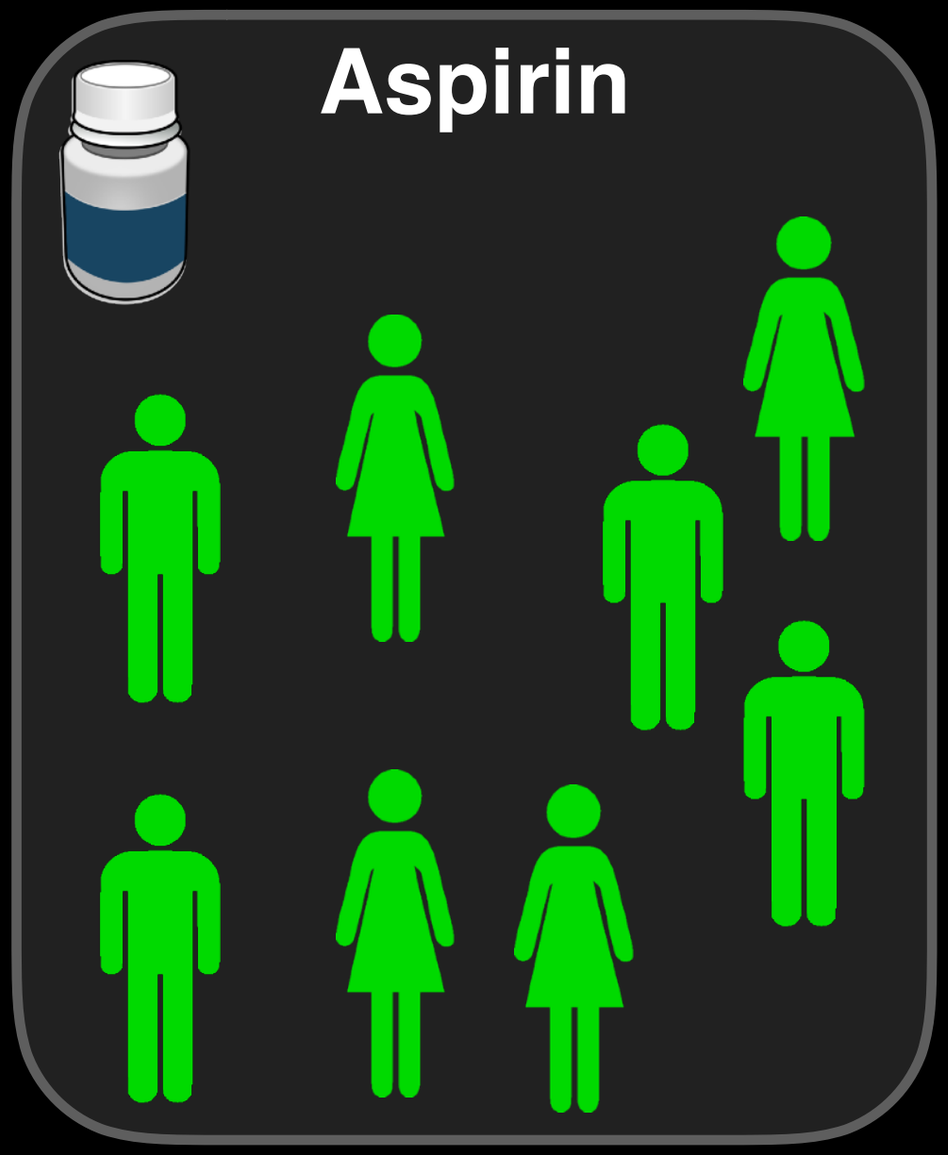
Example: Say you run a medical trial comparing Aspirin vs New Drug X.

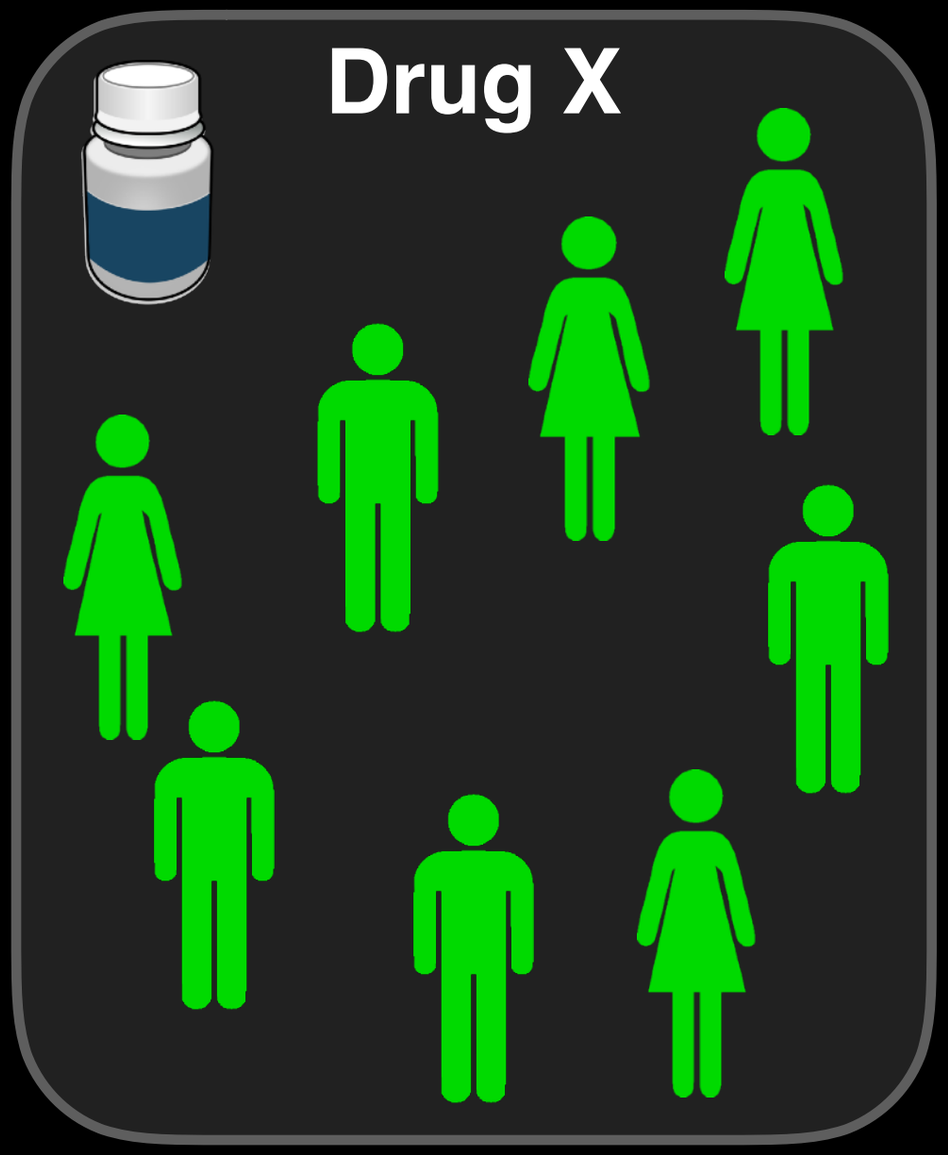
Aspirin

Drug X

Problem 2: “Cherry-Picking” Hypotheses

Example: Say you run a medical trial comparing Aspirin vs New Drug X.



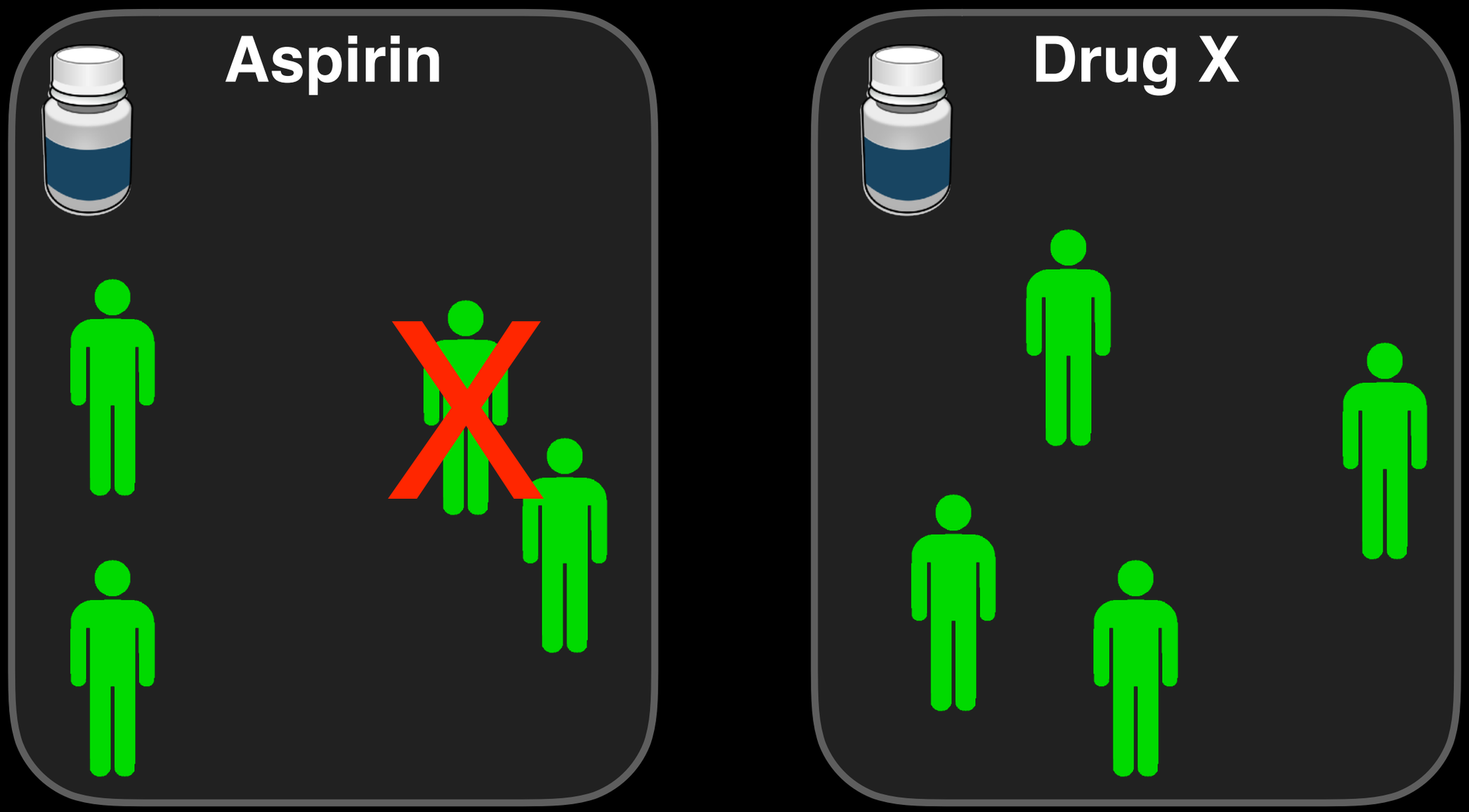


Problem 2: “Cherry-Picking” Hypotheses

Example: Say you run a medical trial comparing Aspirin vs New Drug X.

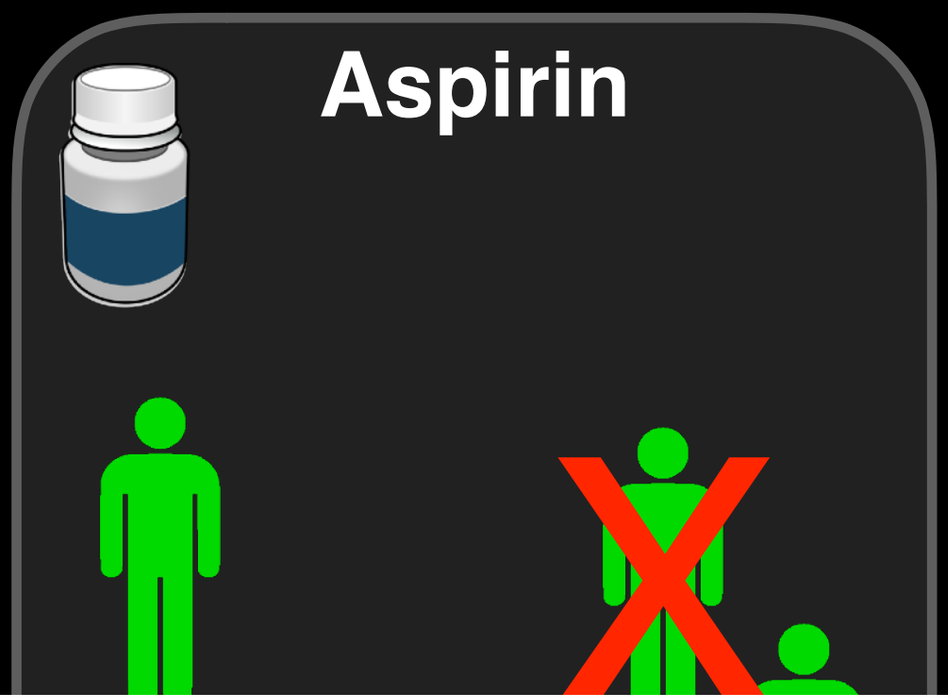
Problem 2: “Cherry-Picking” Hypotheses

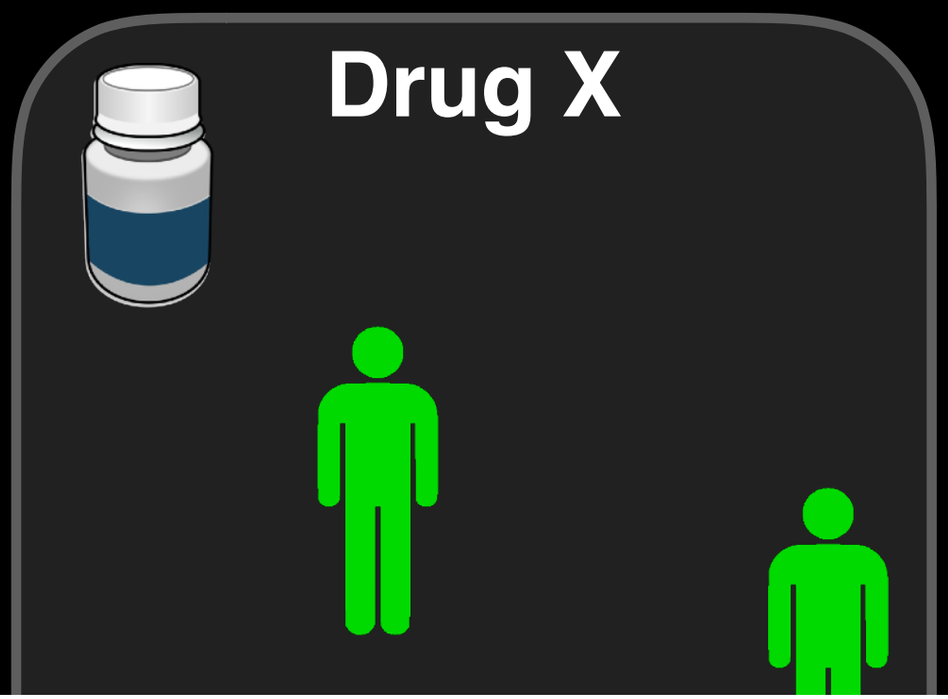
Example: Say you run a medical trial comparing Aspirin vs New Drug X.



Problem 2: “Cherry-Picking” Hypotheses

Example: Say you run a medical trial comparing Aspirin vs New Drug X.





Clinical Trial Finds New Drug X Safer Than Aspirin Among Men

Outcome

Outcome

- In response to the authors' request, we reviewed an updated draft.

Outcome

- In response to the authors' request, we reviewed an updated draft.
- We replied that omitting the weaker results from the final paper is not legitimate (this would amount to *p*-hacking).

Outcome

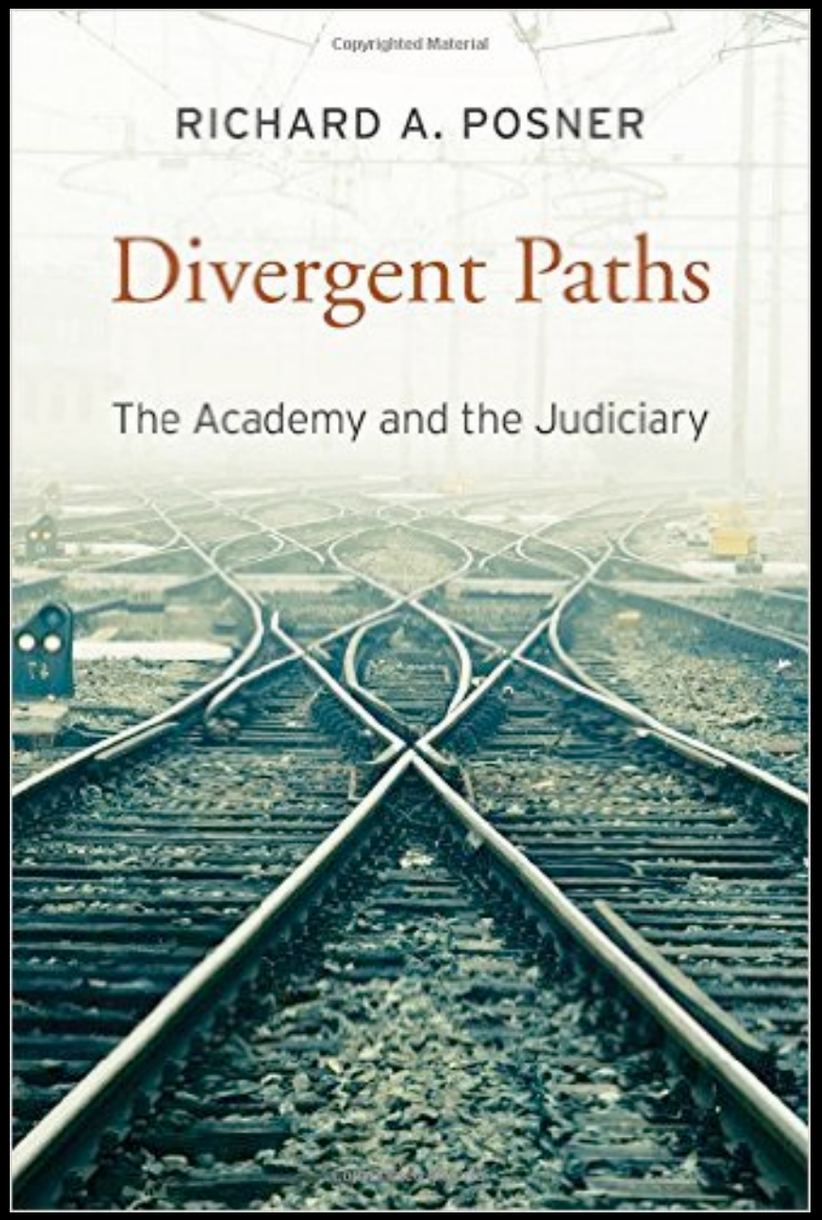
- In response to the authors' request, we reviewed an updated draft.
- We replied that omitting the weaker results from the final paper is not legitimate (this would amount to *p*-hacking).

Result: Adam Chilton & Marin Levy. *Challenging the Randomness of Panel Assignments in the Federal Courts of Appeals*, 101 CORNELL L. REV. 1 (2015).

Outcome

- In response to the authors' request, we reviewed an updated draft.
- We replied that omitting the weaker results from the final paper is not legitimate (this would amount to *p*-hacking).

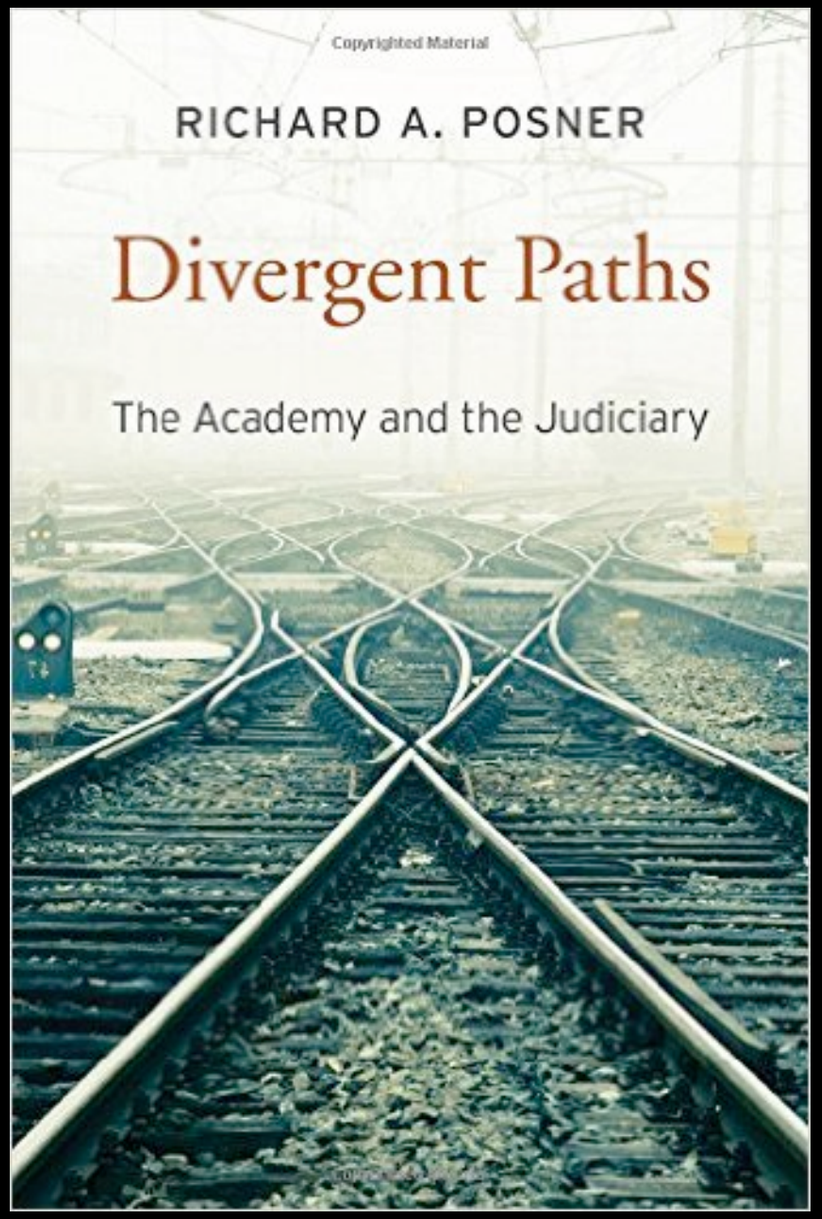
Result: Adam Chilton & Marin Levy. *Challenging the Randomness of Panel Assignments in the Federal Courts of Appeals*, 101 CORNELL L. REV. 1 (2015).



Outcome

"A recent article, however, employing **sophisticated statistical methods**, identifies four circuits—D.C., Second, Eighth, and Ninth—in which assignments are not random, but indeed appear to produce an ideological slant."

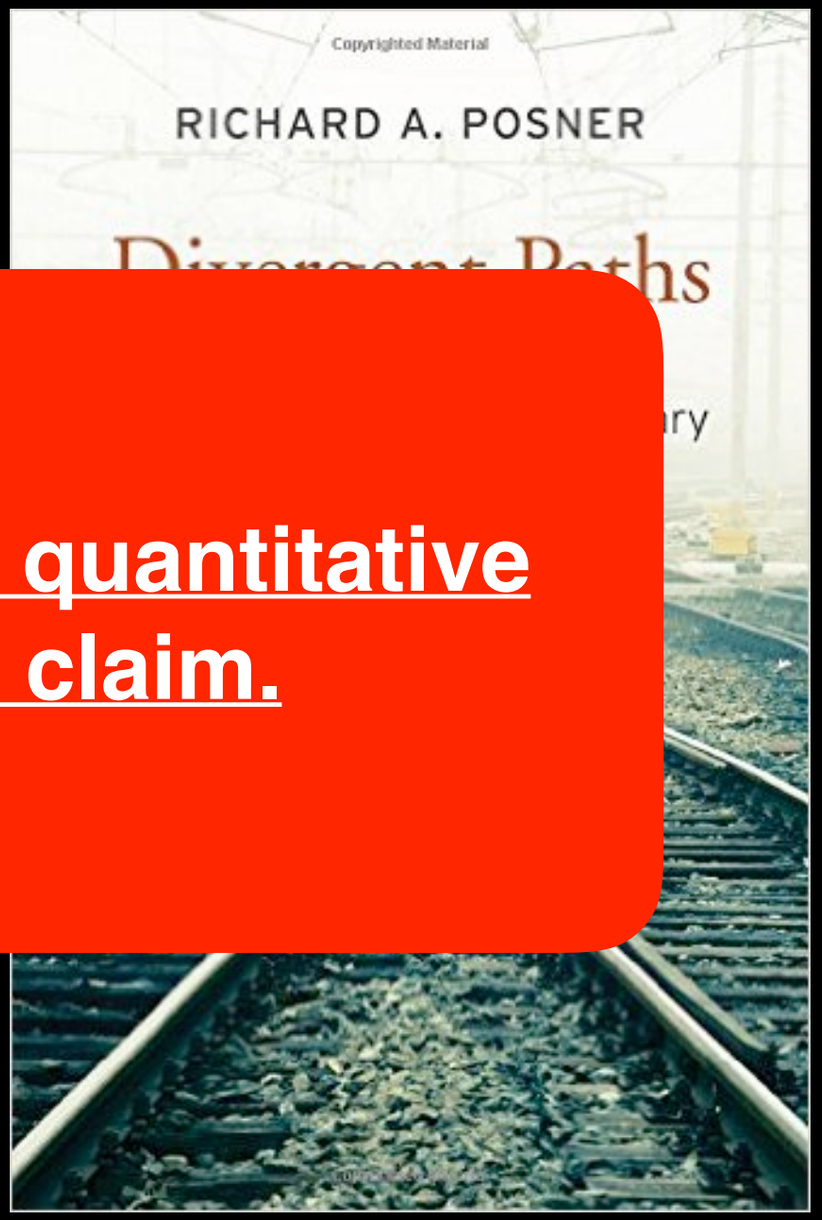
Result: Adam Chilton & Mark Levy. *Challenging the Randomness of Panel Assignments in the Federal Courts of Appeals*, 101 CORNELL L. REV. 1 (2015).



Outcome

"A recent article, however, employing **sophisticated statis**
ident
Seco
in wh
not r
appe
ideol

Levy. *Challenging the Randomness of Panel Assignments in the Federal Courts of Appeals*, 101 CORNELL L. REV. 1 (2015).



<u>The study provides no quantitative evidence for this claim.</u>

Recommendations

Recommendations

For lawyers: Be wary of statistical claims; it is easy to generate spurious “statistically significant” ($p < 0.05$) findings:

Recommendations

For lawyers: Be wary of statistical claims; it is easy to generate spurious “statistically significant” ($p < 0.05$) findings:

Bad Science

Recommendations

For lawyers: Be wary of statistical claims; it is easy to generate spurious “statistically significant” ($p < 0.05$) findings:

Bad Science

- Run 20 independent studies.

Recommendations

For lawyers: Be wary of statistical claims; it is easy to generate spurious "statistically significant" ($p < 0.05$) findings:

Bad Science

- Run 20 independent studies.
  - → On average, one will yield a "significant" result.

Recommendations

For lawyers: Be wary of statistical claims; it is easy to generate spurious "statistically significant" ($p < 0.05$) findings:

Bad Science

- Run 20 independent studies.
  - → On average, one will yield a "significant" result.
- Discard the 19 non-significant results.

Recommendations

For lawyers: Be wary of statistical claims; it is easy to generate spurious "statistically significant" ($p < 0.05$) findings:

Bad Science

- Run 20 independent studies.
  - → On average, one will yield a "significant" result.
- Discard the 19 non-significant results.
- Insert the "significant" result into your petition.

Recommendations

For lawyers: Be wary of statistical claims; it is easy to generate spurious "statistically significant" ($p < 0.05$) findings:

Bad Science

- Run 20 independent studies.
  → On average, one will yield a "significant" result.
- Discard the 19 non-significant results.
- Insert the "significant" result into your petition.

For law review editors: Consider consulting external reviewers about papers with a heavy computational or statistical angle.

PART I

The Perils of Retrospective Analysis

PART II

Cryptographic Techniques to Ensure Fair Randomness

PART I

The Perils of Retrospective Analysis

PART II

Cryptographic Techniques to Ensure Fair Randomness

Randomness

The Chilton-Levy study raises a provocative question:

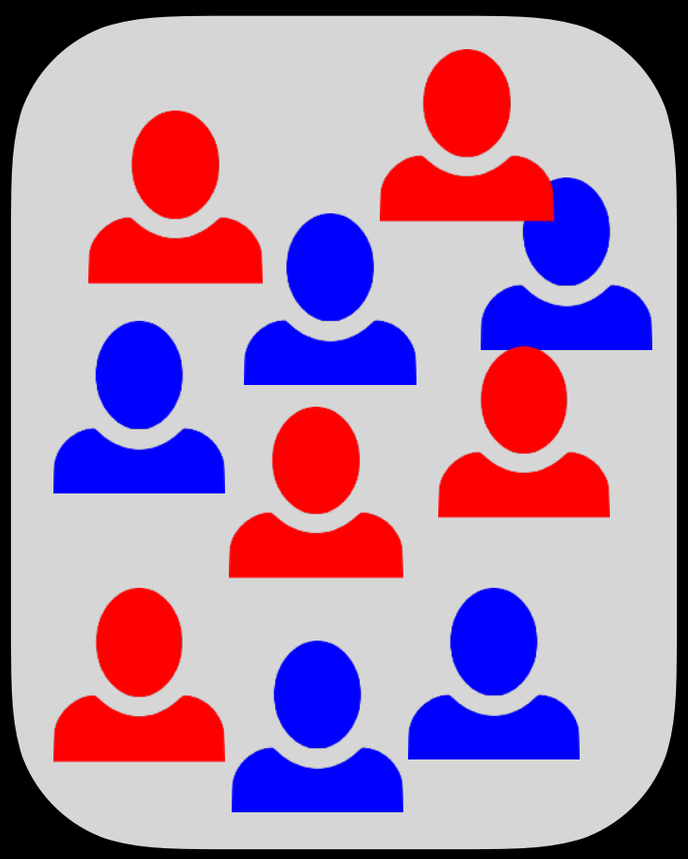
How can the public be sure that the assignments are random?

Randomness

The Chilton-Levy study raises a provocative question:

How can the public be sure that the assignments are random?

Judges

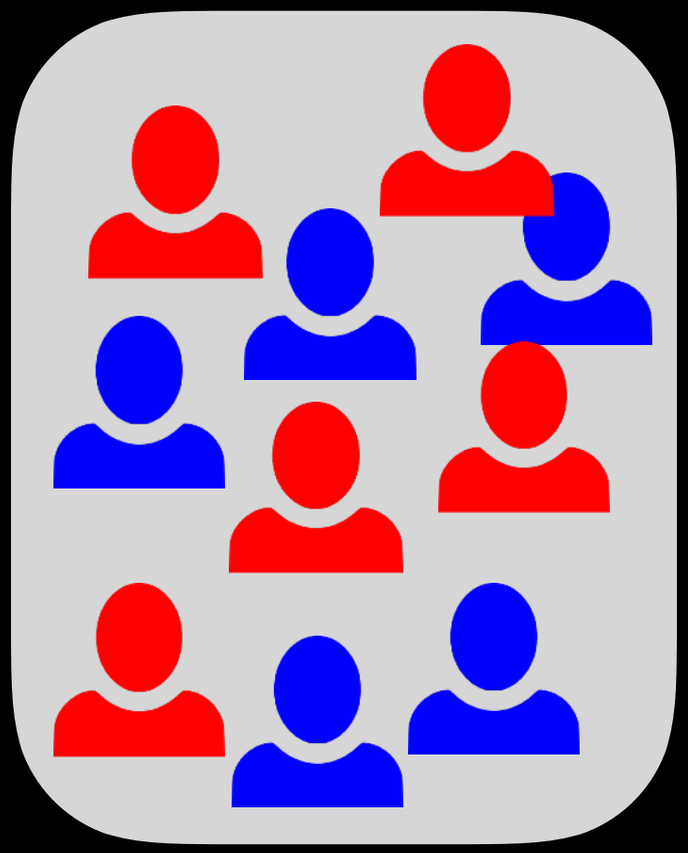


Randomness

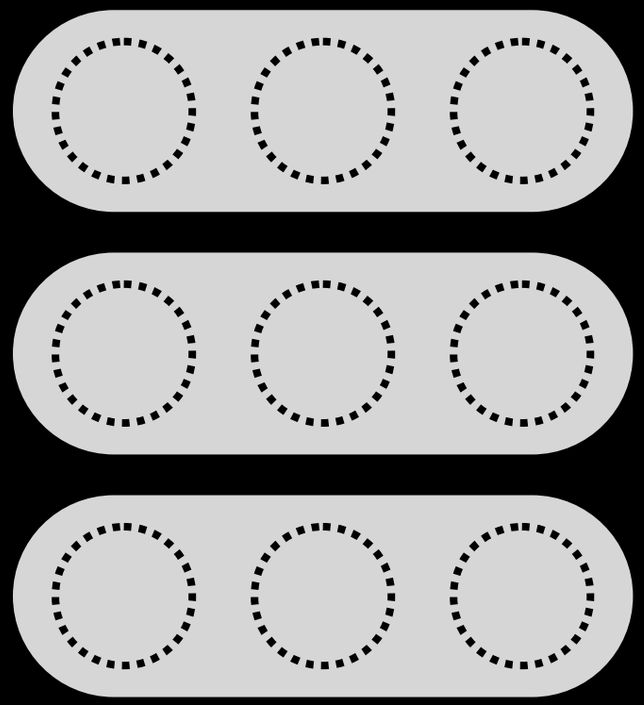
The Chilton-Levy study raises a provocative question:

How can the public be sure that the assignments are random?

Judges



Panels

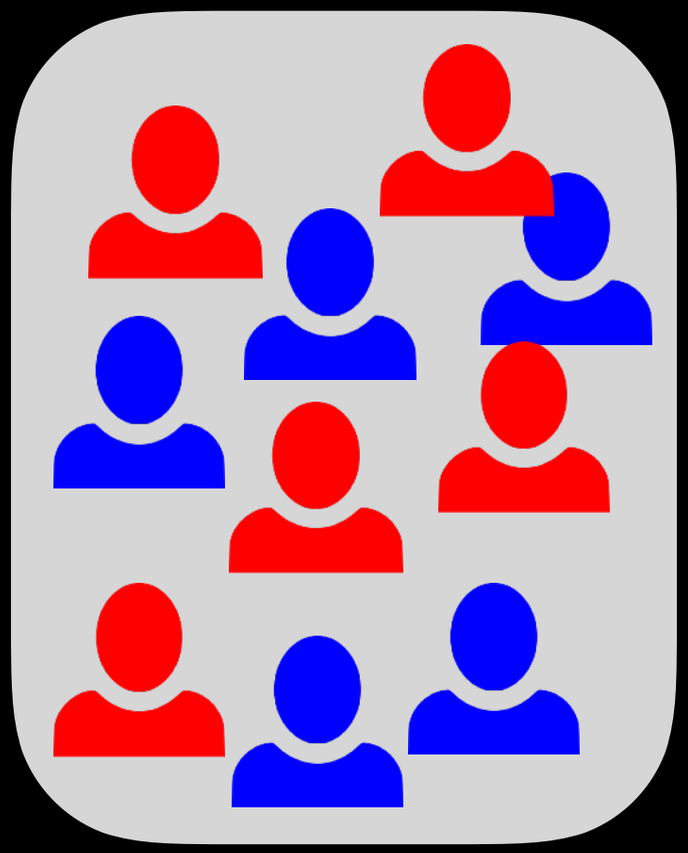


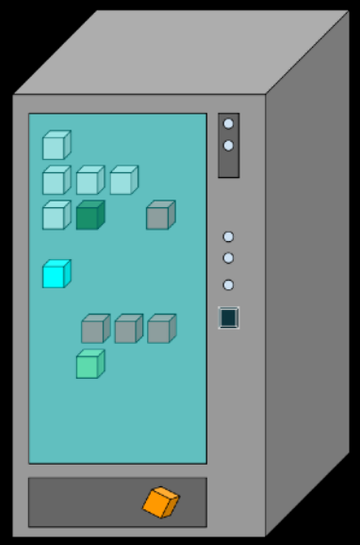
Randomness

The Chilton-Levy study raises a provocative question:

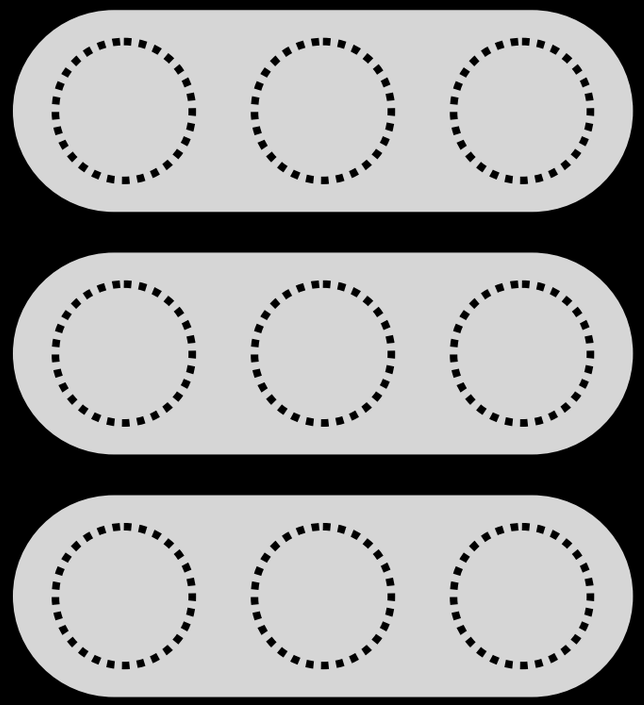
How can the public be sure that the assignments are random?

Judges





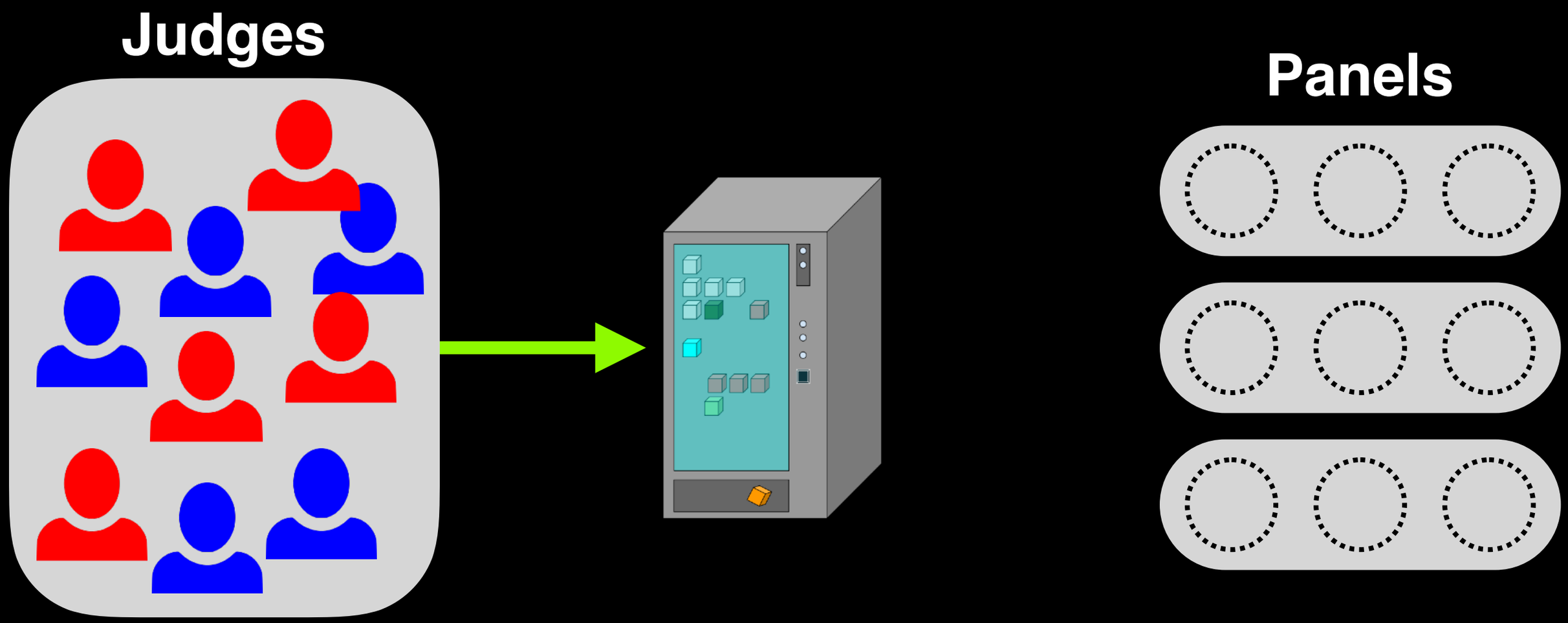
Panels



Randomness

The Chilton-Levy study raises a provocative question:

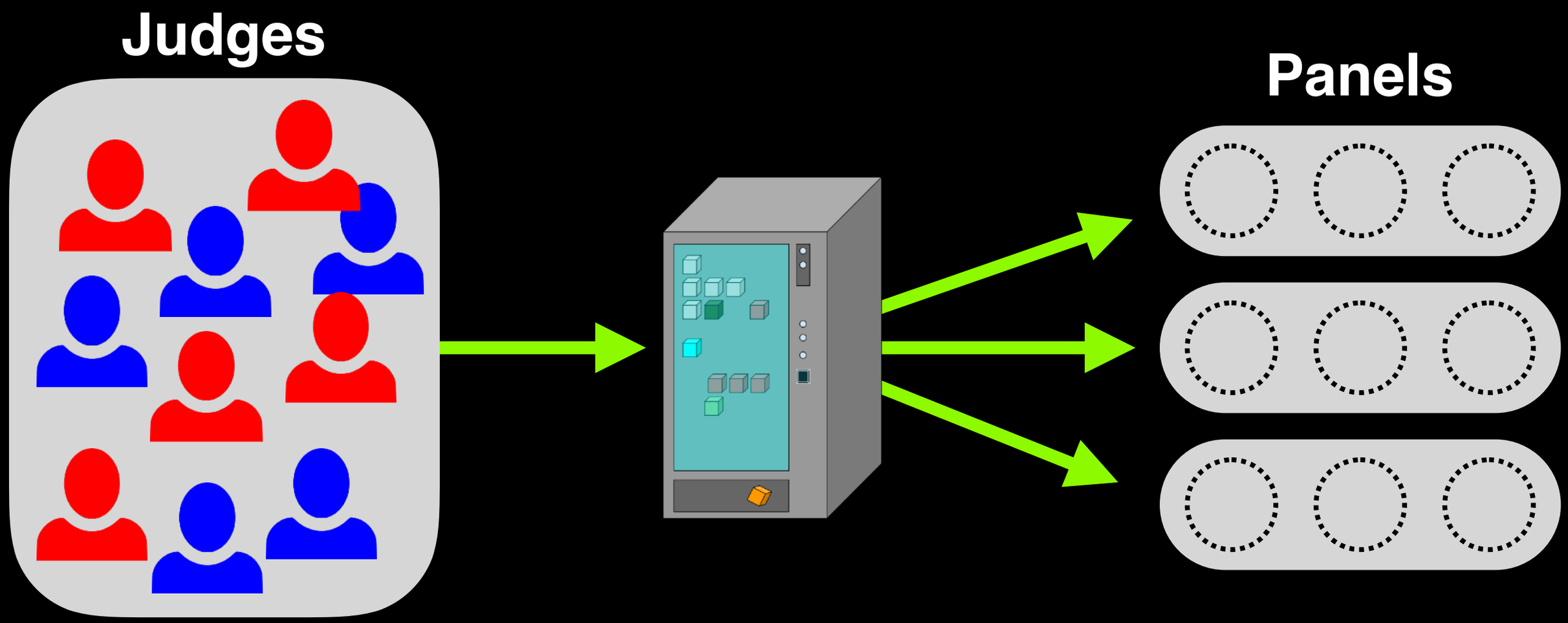
How can the public be sure that the assignments are random?



Randomness

The Chilton-Levy study raises a provocative question:

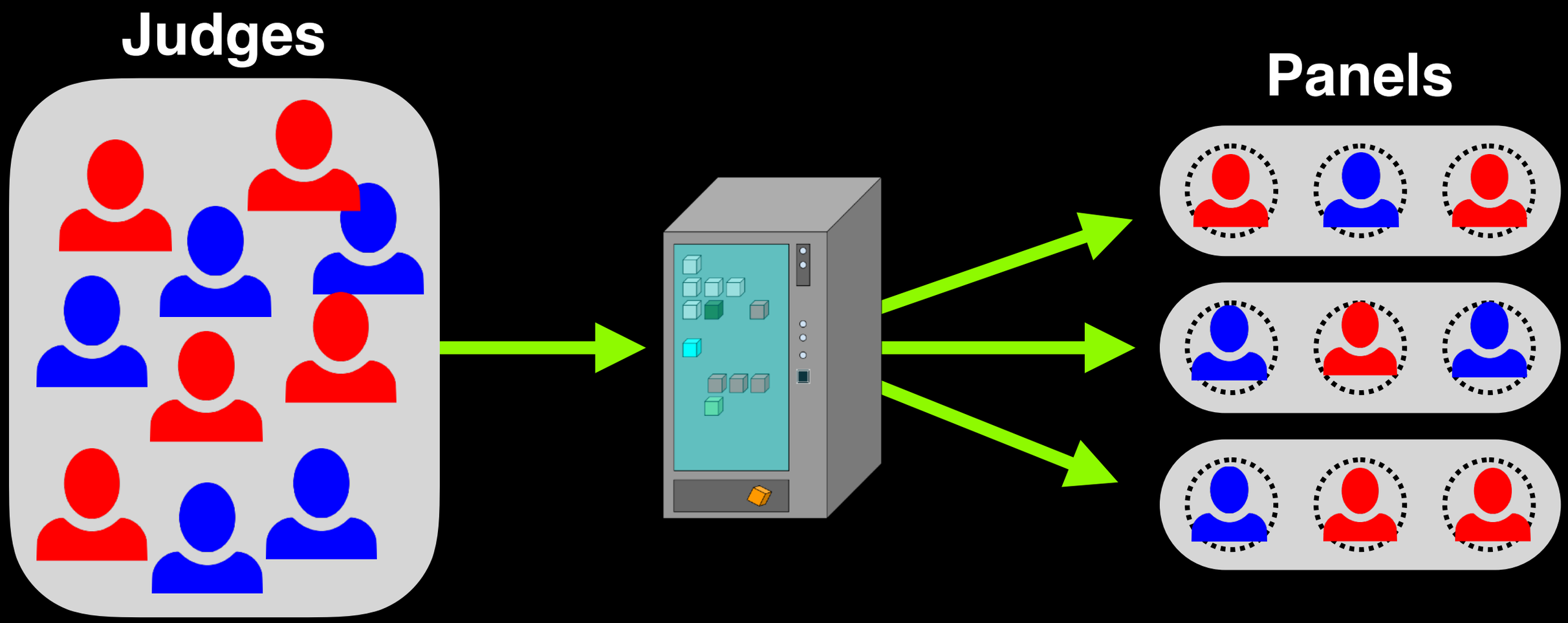
How can the public be sure that the assignments are random?



Randomness

The Chilton-Levy study raises a provocative question:

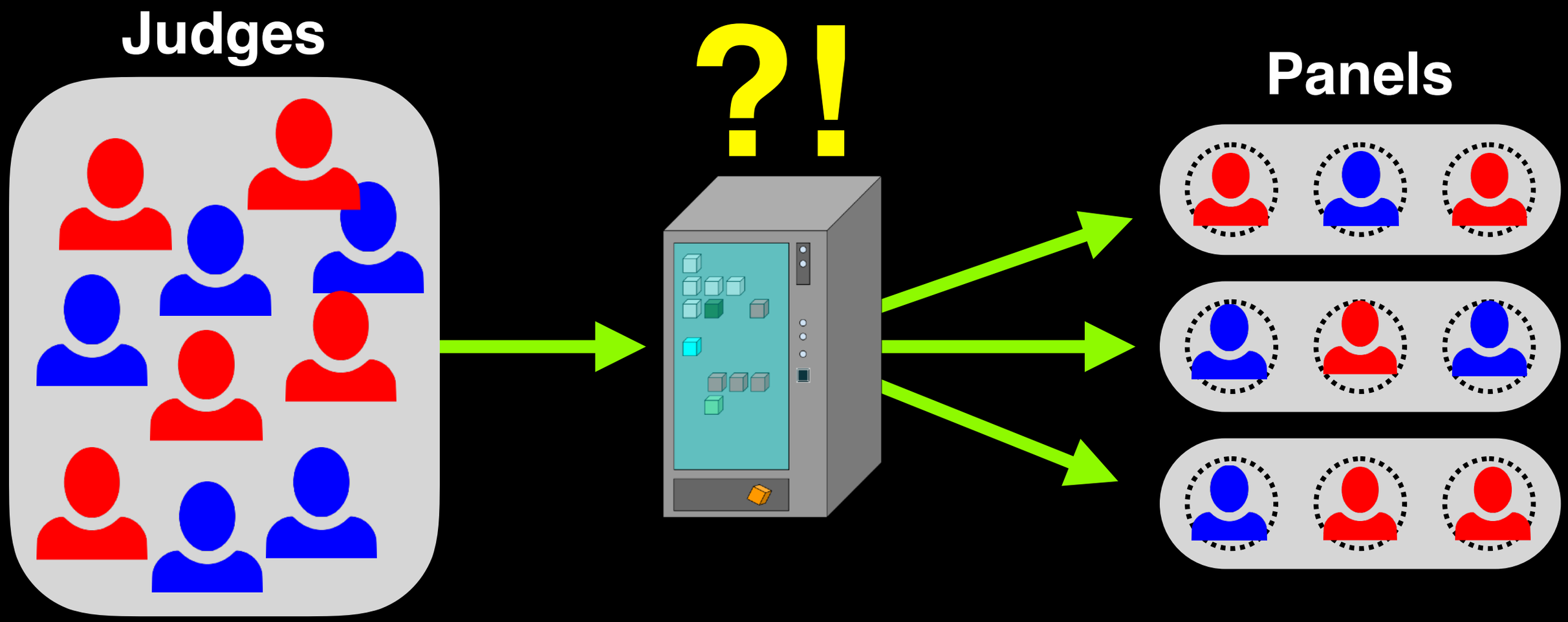
How can the public be sure that the assignments are random?



Randomness

The Chilton-Levy study raises a provocative question:

How can the public be sure that the assignments are random?



Other Applications

Other Applications

Verifiable randomness sources have a host of other applications:
[Bonneau et al. 2015] [Kroll et al., 2017]

Other Applications

Verifiable randomness sources have a host of other applications:
[Bonneau et al. 2015] [Kroll et al., 2017]

- Visa lottery

Other Applications

Verifiable randomness sources have a host of other applications:
[Bonneau et al. 2015] [Kroll et al., 2017]

- Visa lottery
- Actual lottery

Other Applications

Verifiable randomness sources have a host of other applications:
[Bonneau et al. 2015] [Kroll et al., 2017]

- Visa lottery
- Actual lottery
- IRS audits

Other Applications

Verifiable randomness sources have a host of other applications:
[Bonneau et al. 2015] [Kroll et al., 2017]

- Visa lottery
- Actual lottery
- IRS audits
- Random screening at airports (customs, security, etc.)

Other Applications

Verifiable randomness sources have a host of other applications:
[Bonneau et al. 2015] [Kroll et al., 2017]

- Visa lottery
- Actual lottery
- IRS audits
- Random screening at airports (customs, security, etc.)
- School lotteries

Other Applications

Verifiable randomness sources have a host of other applications:
[Bonneau et al. 2015] [Kroll et al., 2017]

- Visa lottery
- Actual lottery
- IRS audits
- Random screening at airports (customs, security, etc.)
- School lotteries
- …

Three Cryptographic Ideas

1. Coin-Flipping Protocols
2. Fair Exchange
3. Randomness Beacons

Three Cryptographic Ideas

1. **Coin-Flipping Protocols**
2. Fair Exchange
3. Randomness Beacons

Tool 1: Coin-Flipping Protocol

Tool 1: Coin-Flipping Protocol

Claim: If we have a means to produce a single unpredictable number *N*, that is enough to select a random panel.

Tool 1: Coin-Flipping Protocol

Claim: If we have a means to produce a single unpredictable number *N*, that is enough to select a random panel.

→ Make a list of all possible panels, use the *N*-th panel.
(We can think of *N* as between 1 and 100.)

Tool 1: Coin-Flipping Protocol

Claim: If we have a means to produce a single unpredictable number *N*, that is enough to select a random panel.

→ Make a list of all possible panels, use the *N*-th panel.
(We can think of *N* as between 1 and 100.)

Principle: The best (only?) way to be sure that a “random” number is really random is to pick it yourself. [Blum 1983]

Tool 1: Coin-Flipping Protocol

Petitioner 

Respondent 

Tool 1: Coin-Flipping Protocol

“Cryptographic commitment”

Petitioner 

Respondent 

Tool 1: Coin-Flipping Protocol

"Cryptographic commitment"

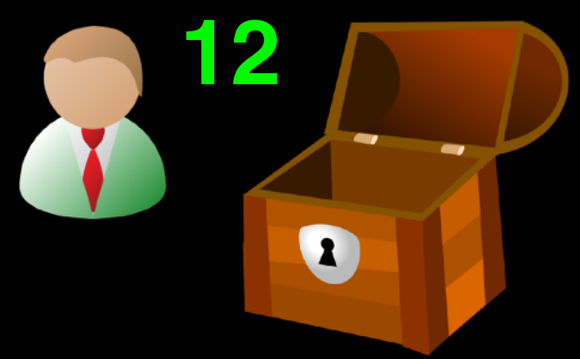
Petitioner

Respondent

Tool 1: Coin-Flipping Protocol

"Cryptographic commitment"

Petitioner



Respondent

Tool 1: Coin-Flipping Protocol

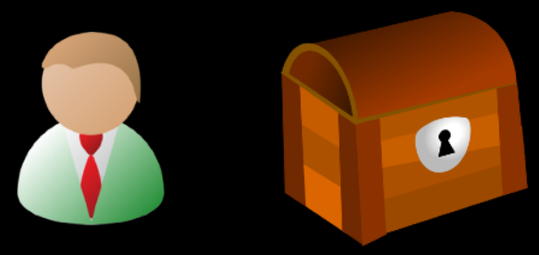
"Cryptographic commitment"

Petitioner  

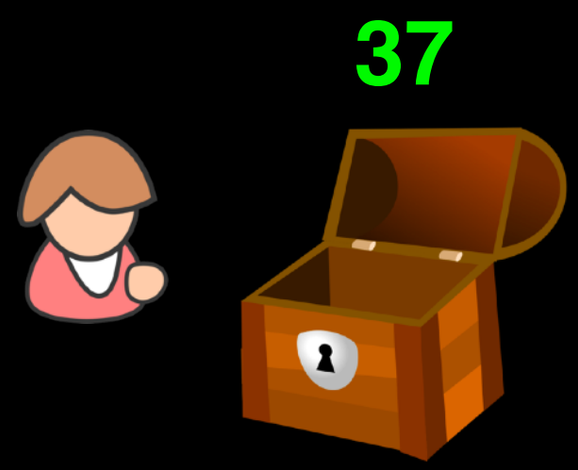
Respondent 

Tool 1: Coin-Flipping Protocol

Petitioner



Respondent



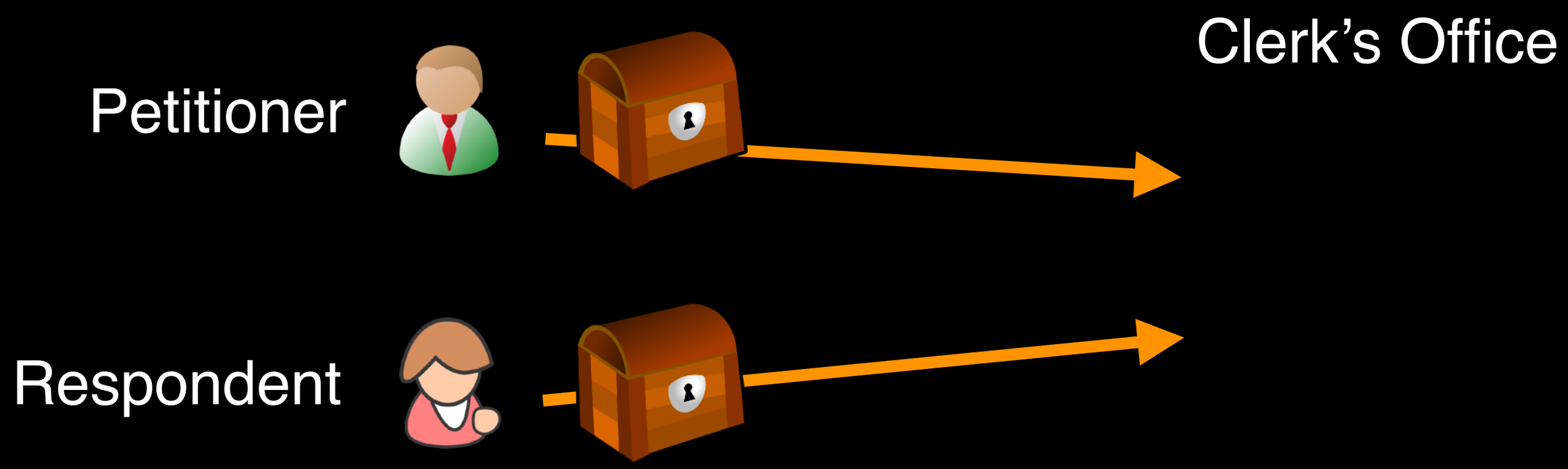
Tool 1: Coin-Flipping Protocol

Clerk's Office

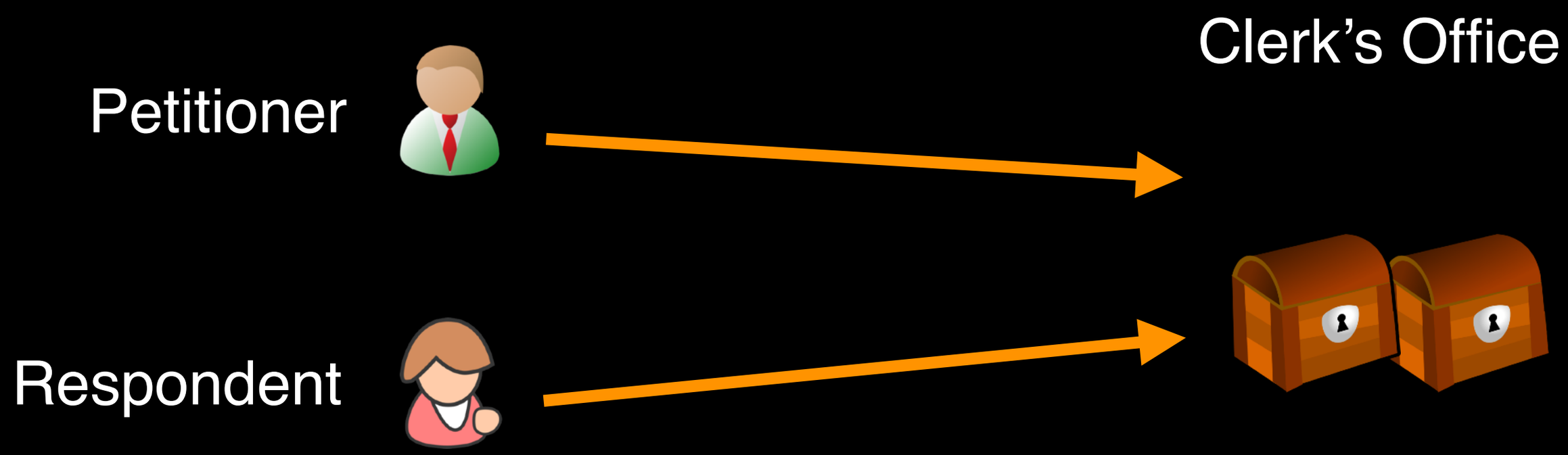
Petitioner  

Respondent  

Tool 1: Coin-Flipping Protocol



Tool 1: Coin-Flipping Protocol



Tool 1: Coin-Flipping Protocol

Clerk's Office

Petitioner

Respondent

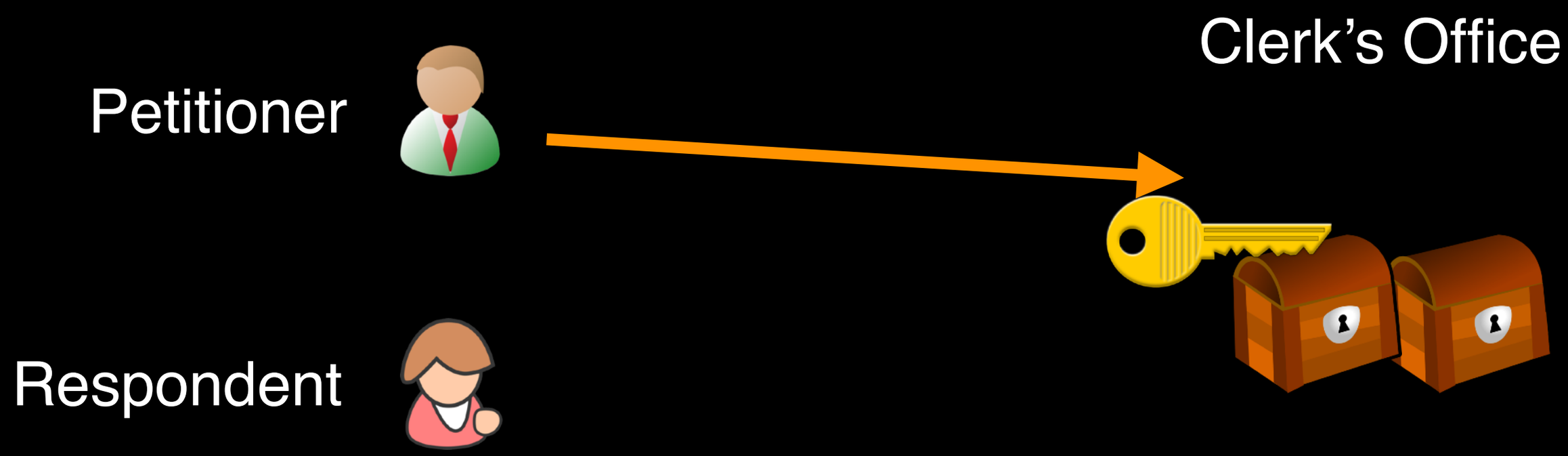
Tool 1: Coin-Flipping Protocol

Clerk's Office

Petitioner

Respondent

Tool 1: Coin-Flipping Protocol



Tool 1: Coin-Flipping Protocol

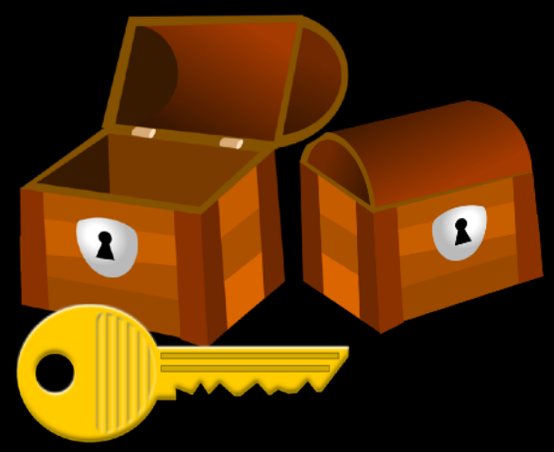
Clerk's Office

Petitioner

Respondent

Tool 1: Coin-Flipping Protocol

Clerk's Office



Petitioner

Respondent

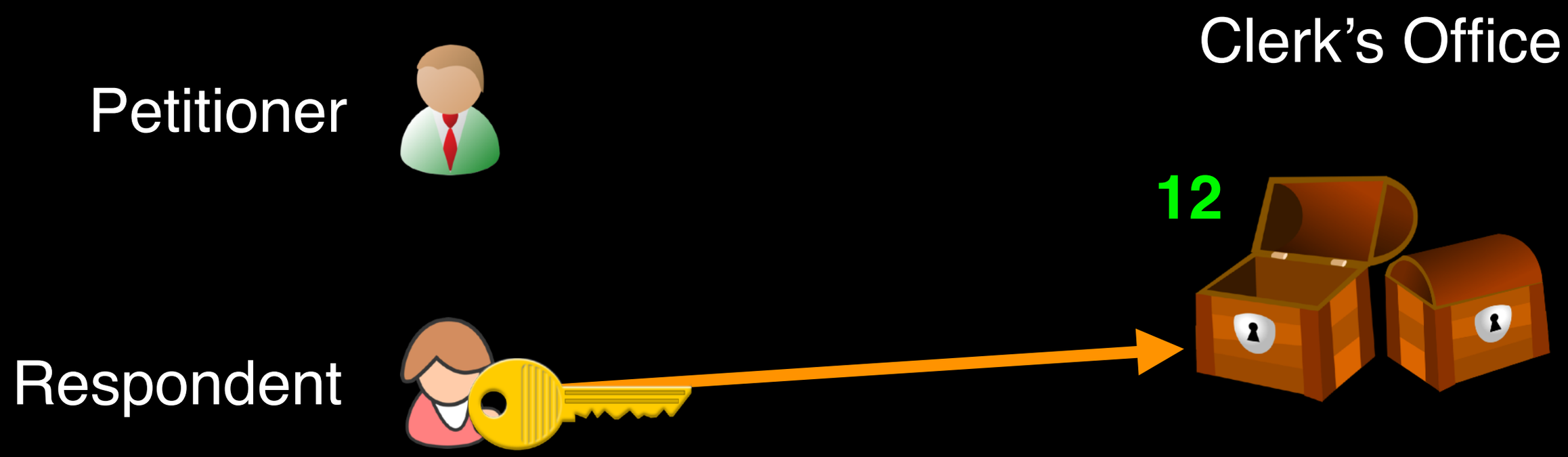
Tool 1: Coin-Flipping Protocol

Clerk's Office

Petitioner

Respondent

Tool 1: Coin-Flipping Protocol

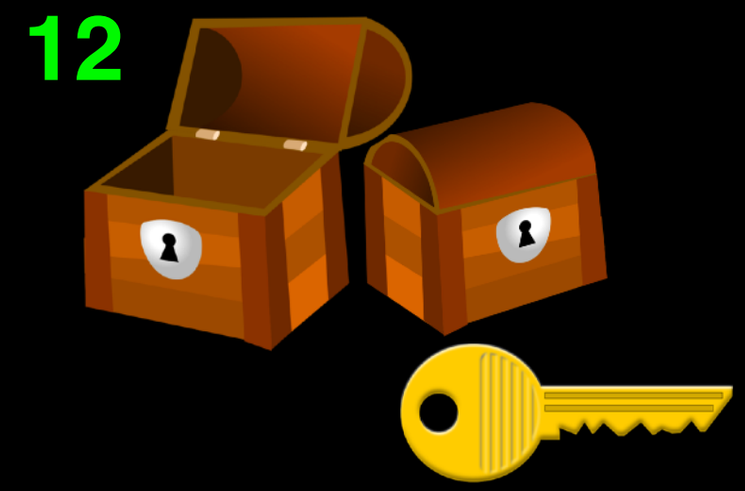


Tool 1: Coin-Flipping Protocol

Tool 1: Coin-Flipping Protocol

Clerk's Office

Petitioner



Respondent

Tool 1: Coin-Flipping Protocol

Clerk's Office

Petitioner

Respondent

Tool 1: Coin-Flipping Protocol

Clerk's Office

Petitioner

Respondent

Tool 1: Coin-Flipping Protocol

Clerk's Office

Petitioner

Respondent

Tool 1: Coin-Flipping Protocol

Petitioner

Respondent

Clerk's Office

Use **12** + **37** = **49** as the random number.

Tool 1: Coin-Flipping Protocol

Clerk's Office

Petitioner

Respondent

Use **12** + **37** = **49** as the random number.

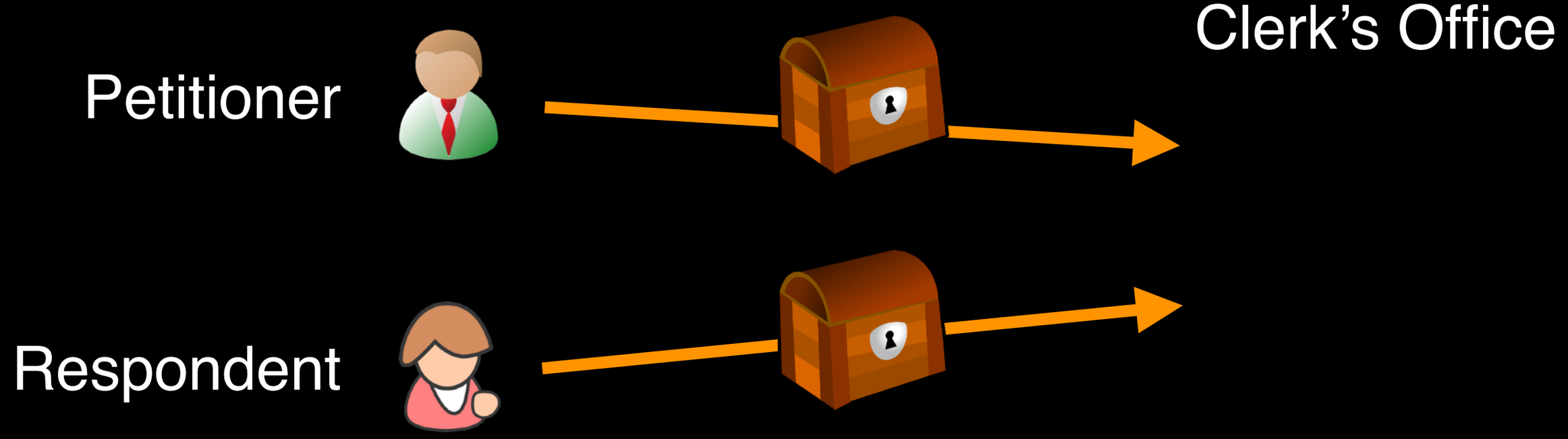
If the sum is over 100, use only the last two digits

Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…



Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

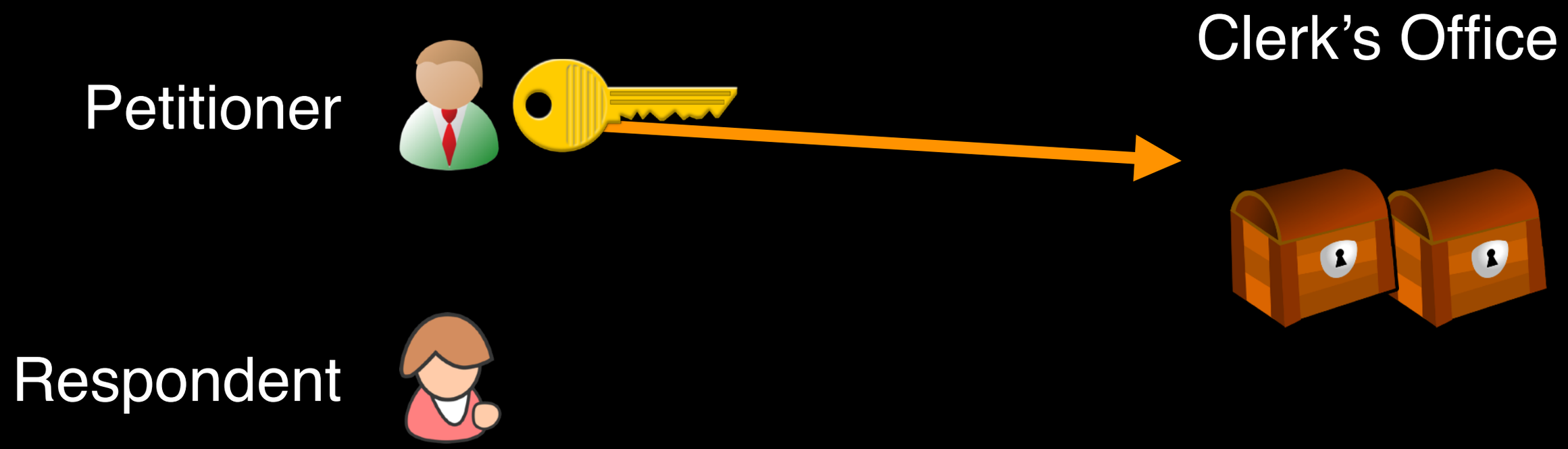
Clerk's Office

Petitioner

Respondent

Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…



Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

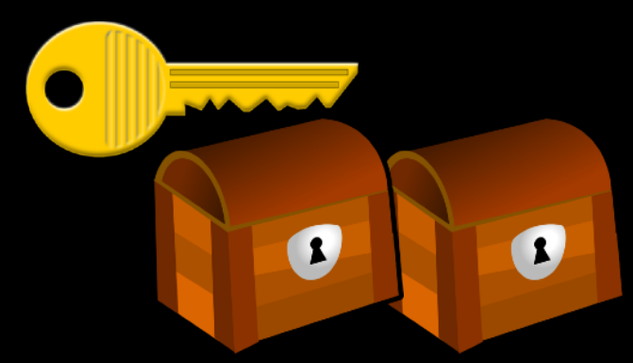
Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Petitioner

Respondent

Clerk's Office



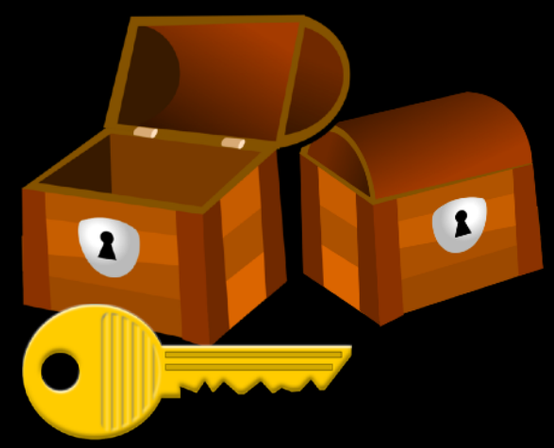
Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Petitioner

Respondent

Clerk's Office



Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Petitioner

Respondent

Clerk's Office

Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Petitioner

Respondent

My number is 37, so we will get panel 49.

Clerk's Office

Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Petitioner

Respondent

I don’t like panel 49.

Clerk’s Office

Are we done?

Problem: The respondent sees the output N before the petitioner does. This creates a loophole…

Petitioner

Respondent

I lost my key!

Let's start over.

Clerk's Office

Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Petitioner

Respondent

The Russians hacked my laptop!

Let's start over.

Clerk's Office

Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Petitioner

Respondent

Clerk's Office

Are we done?

Problem: The respondent sees the output *N* before the petitioner does. This creates a loophole…

Petitioner

Petitioner may choose the panel.

Clerk's Office

Respondent

Three Cryptographic Ideas

1. **Coin-Flipping Protocols**
2. Fair Exchange
3. Randomness Beacons

Three Cryptographic Ideas

1. Coin-Flipping Protocols
2. **Fair Exchange**
3. Randomness Beacons

Tool 2: Fair Exchange

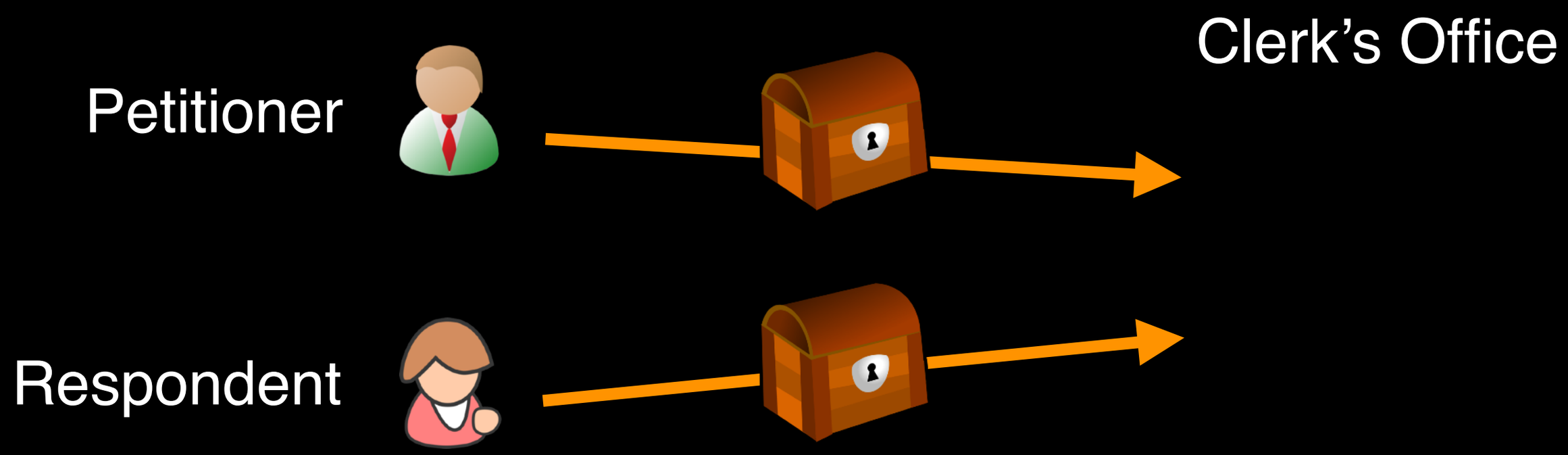
[Blum 1981], [Damgård 1993]

Clerk's Office

Petitioner 

Respondent 

Tool 2: Fair Exchange



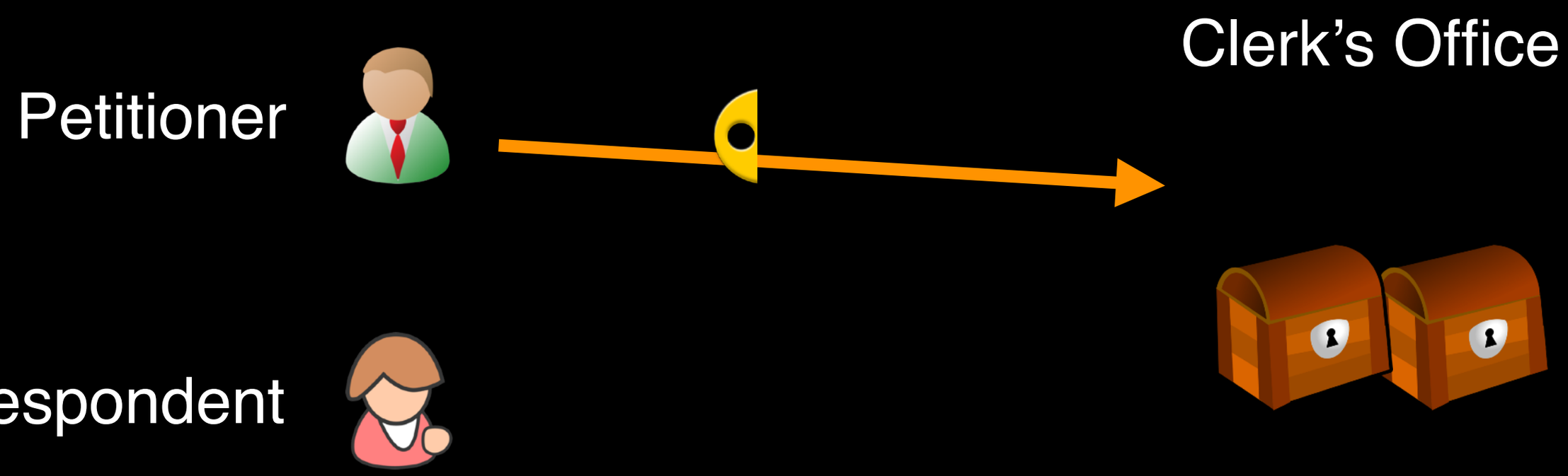
Tool 2: Fair Exchange

Clerk's Office

Petitioner

Respondent

Tool 2: Fair Exchange

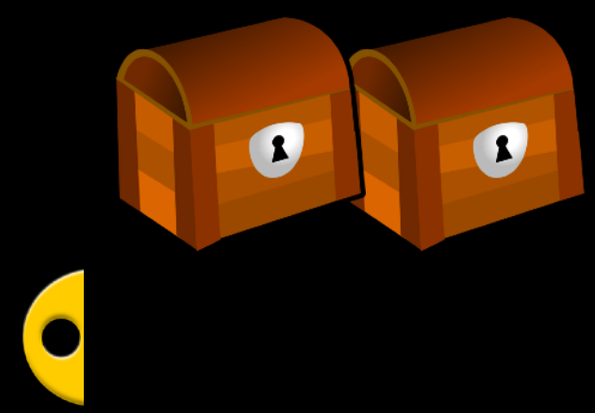


Tool 2: Fair Exchange

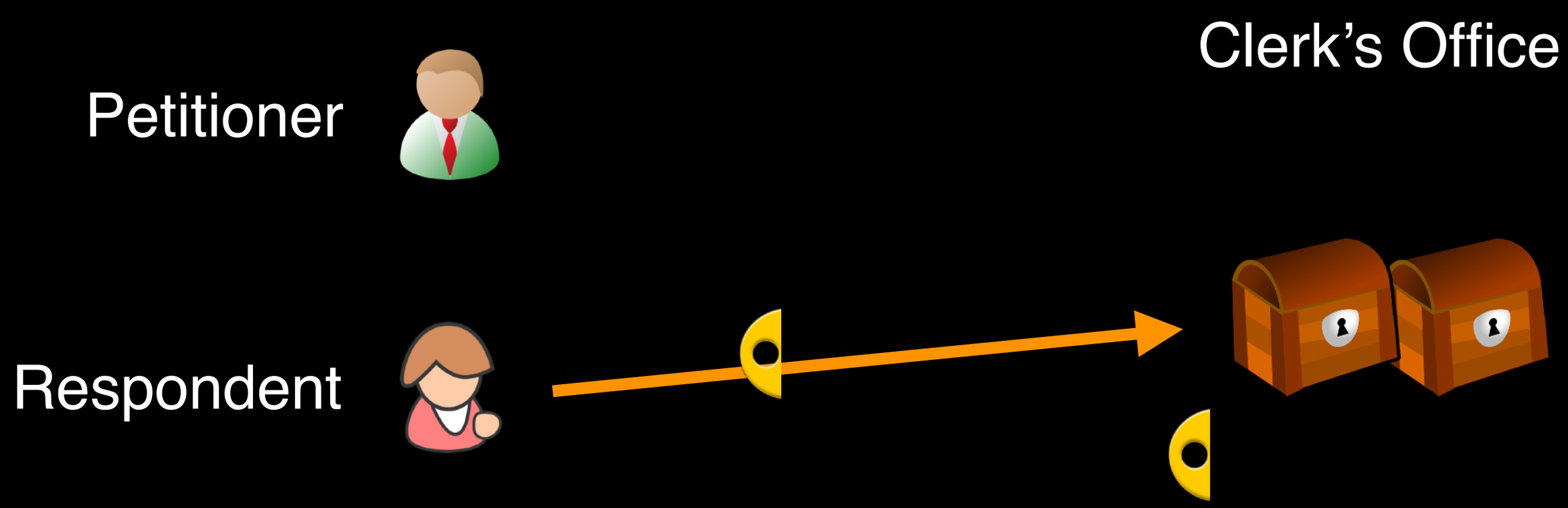
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

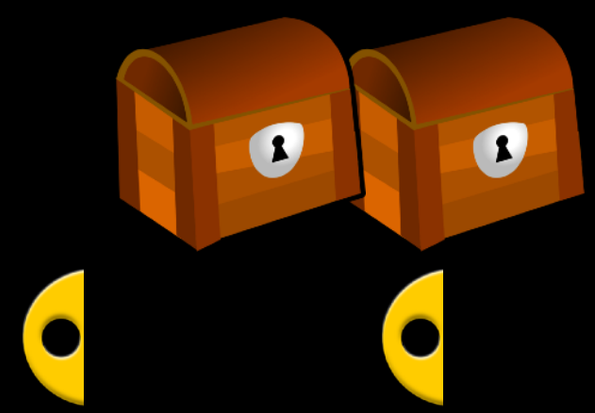


Tool 2: Fair Exchange

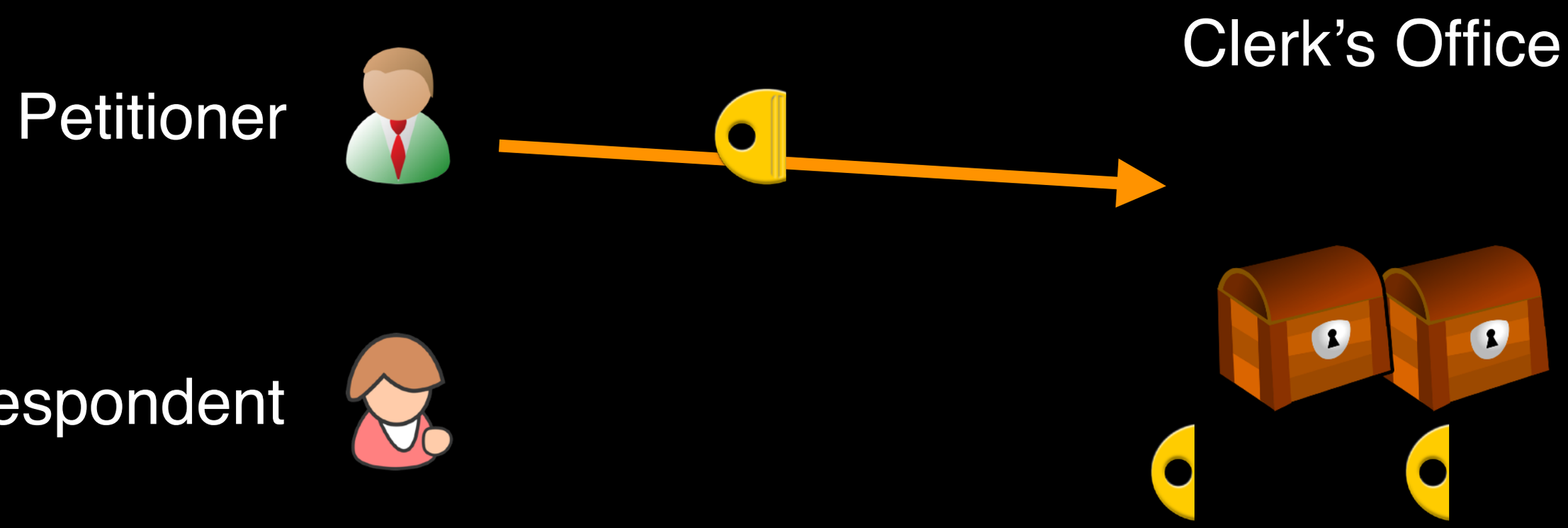
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

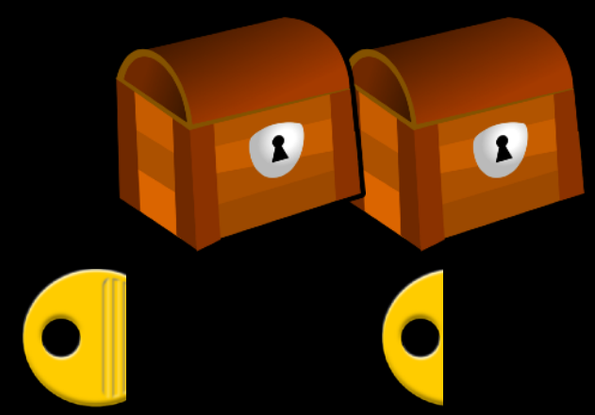


Tool 2: Fair Exchange

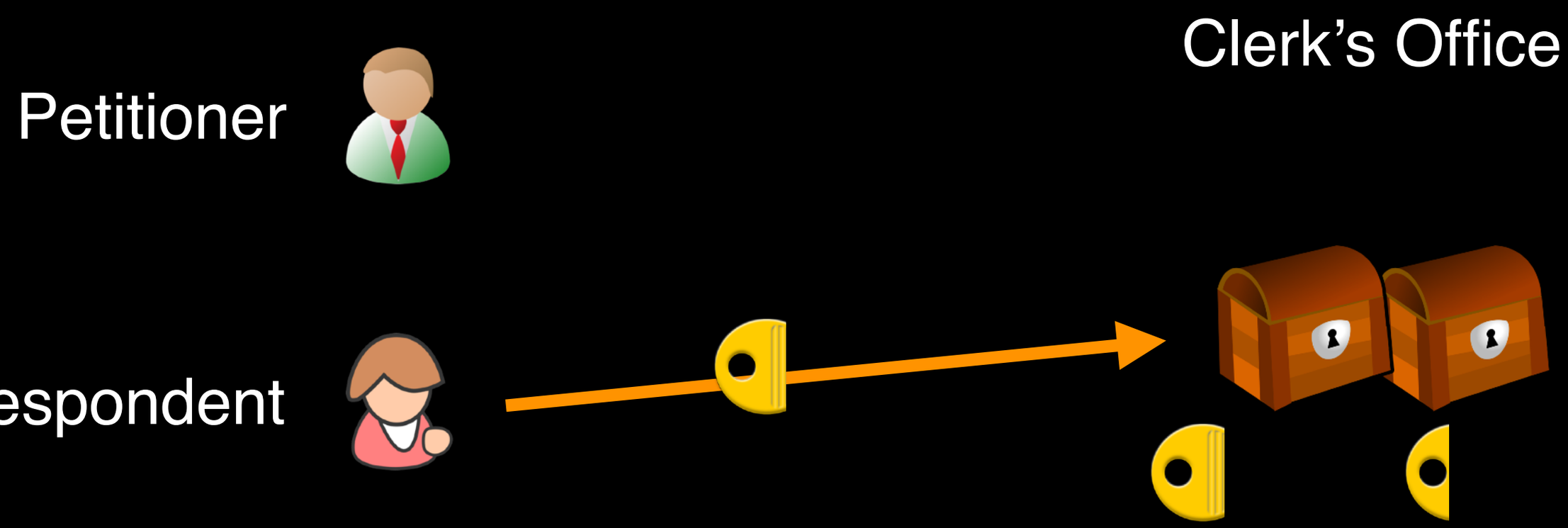
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

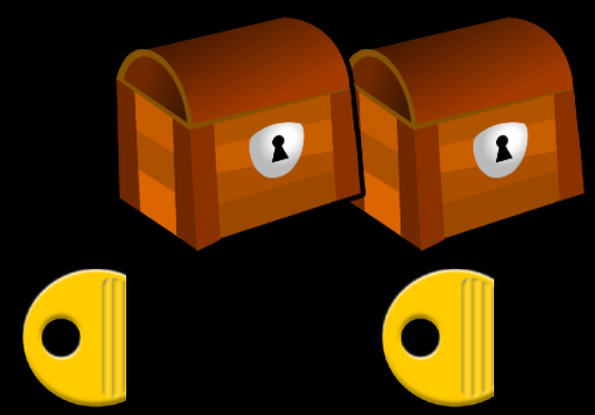


Tool 2: Fair Exchange

Clerk's Office

Petitioner 

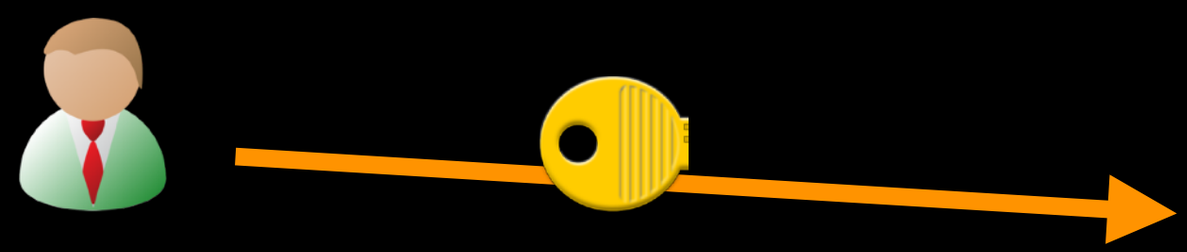
Respondent 



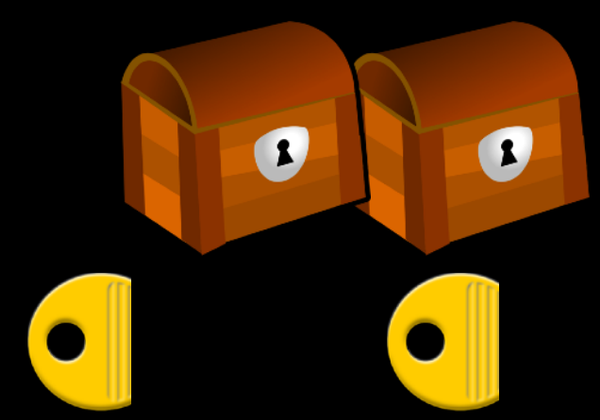
Tool 2: Fair Exchange

Clerk's Office

Petitioner



Respondent

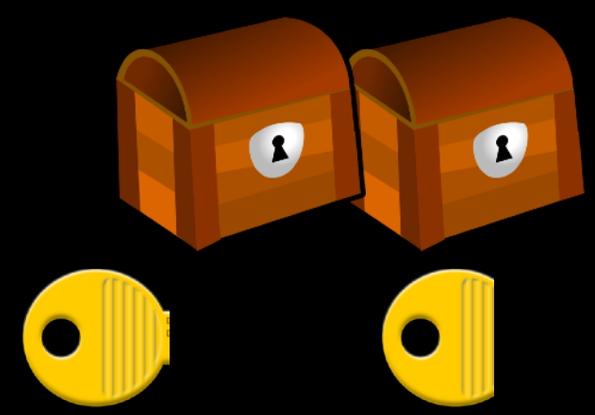


Tool 2: Fair Exchange

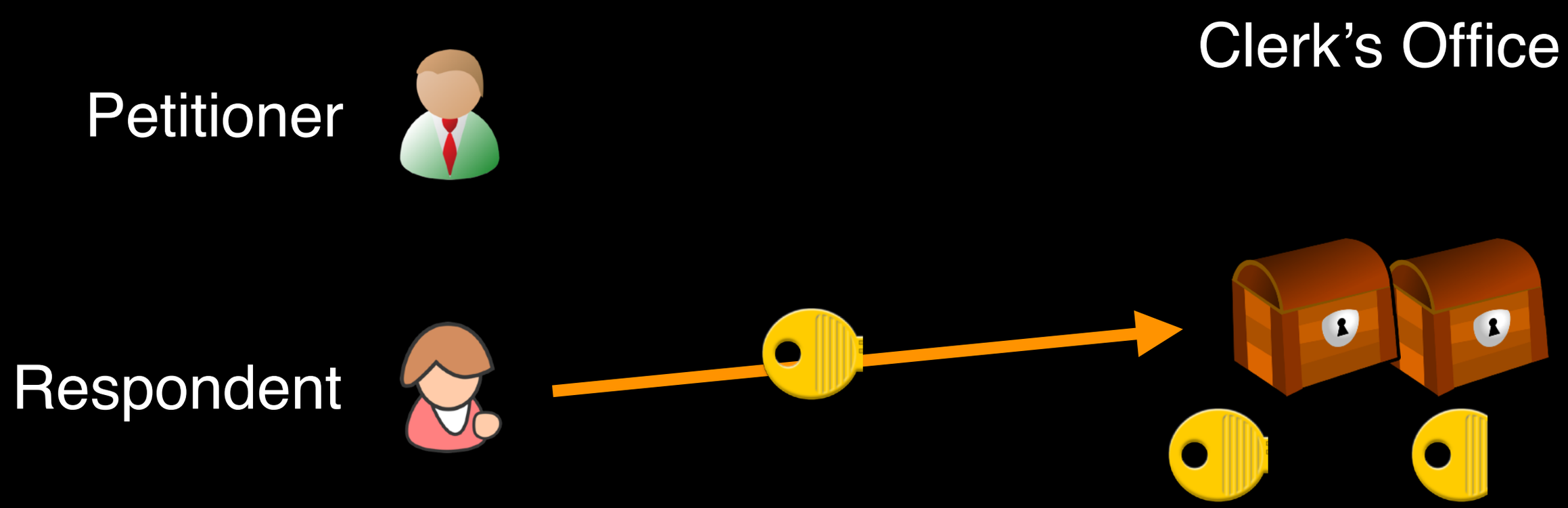
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

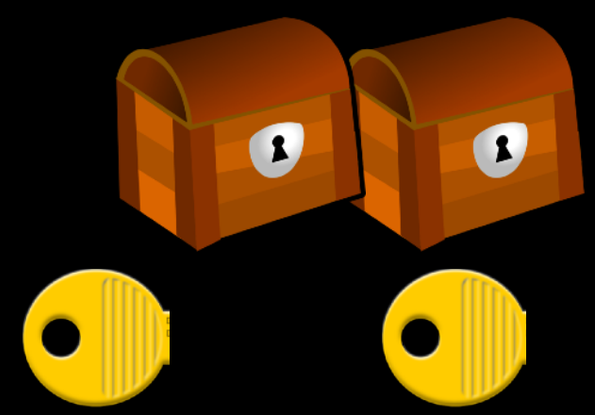


Tool 2: Fair Exchange

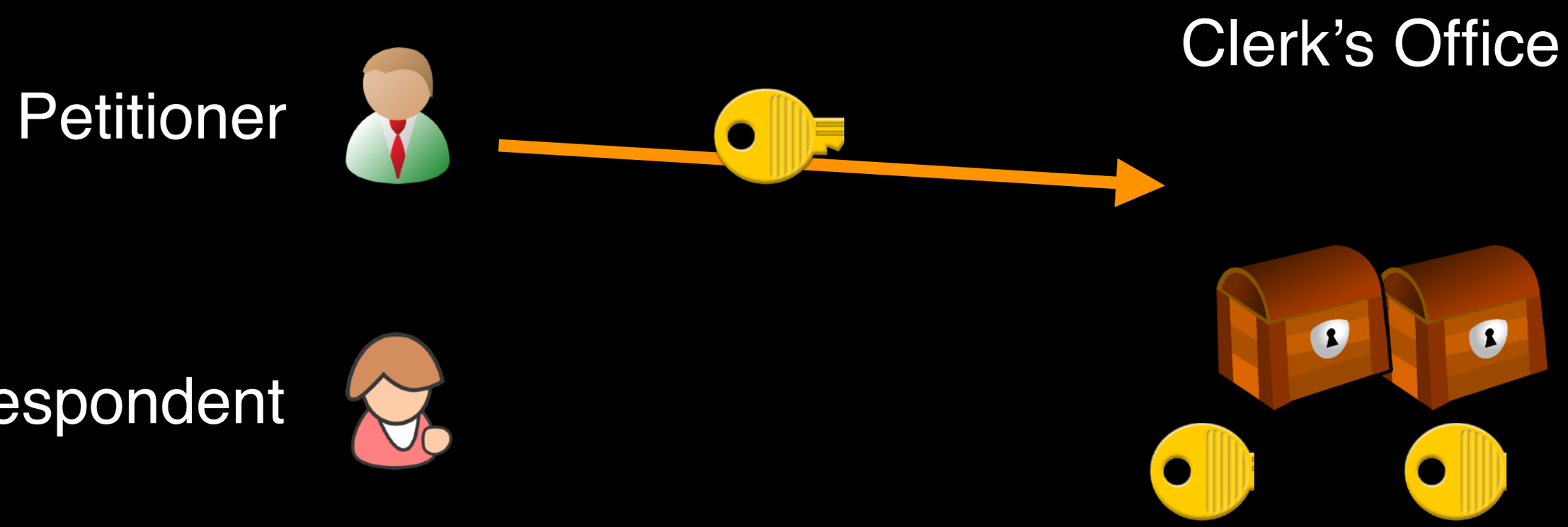
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

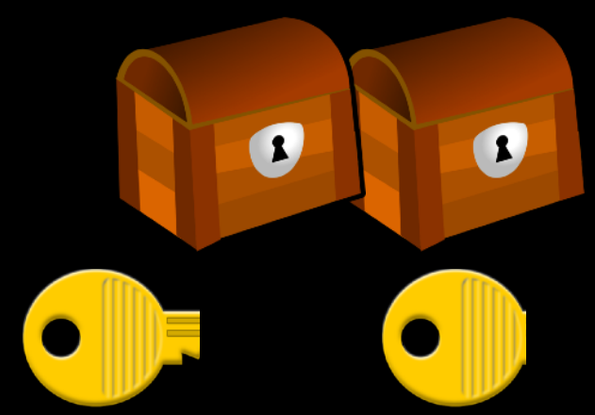


Tool 2: Fair Exchange

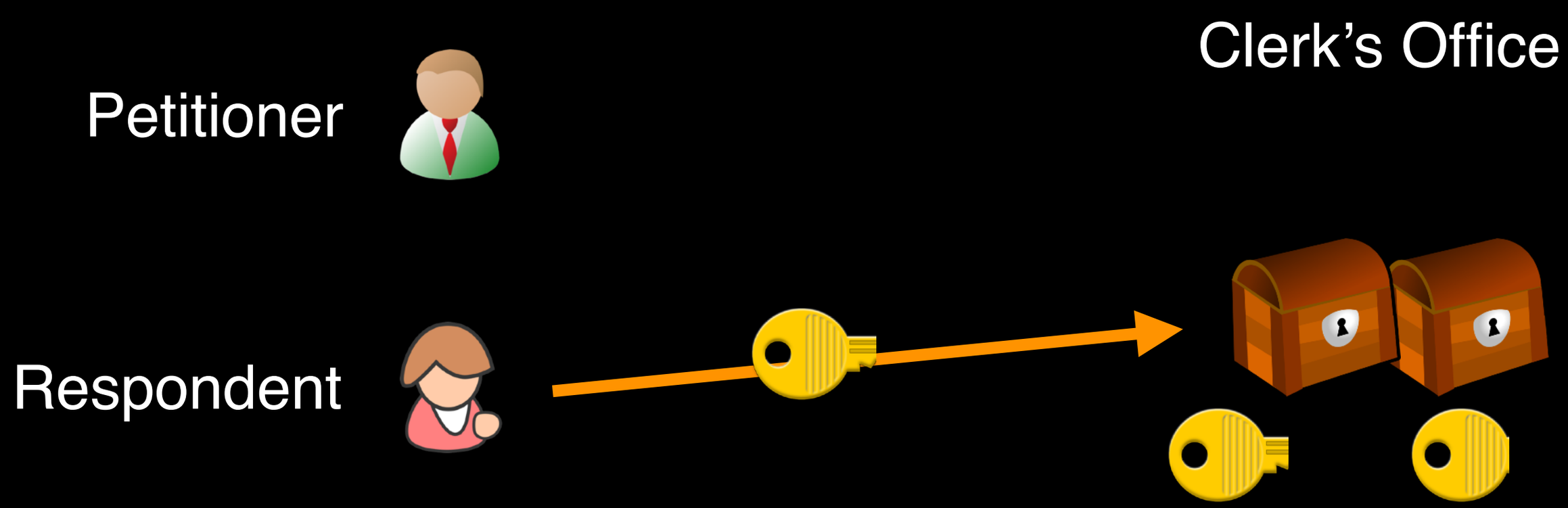
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

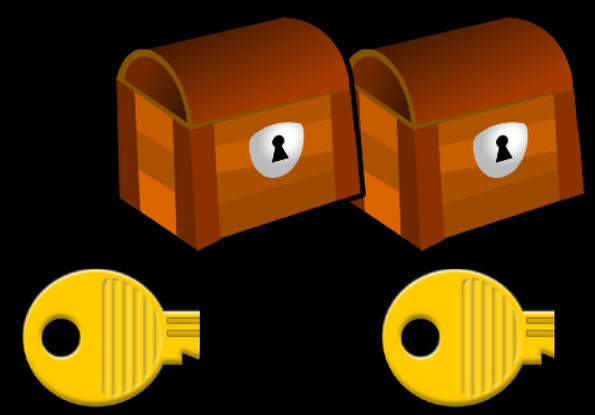


Tool 2: Fair Exchange

Clerk's Office

Petitioner 

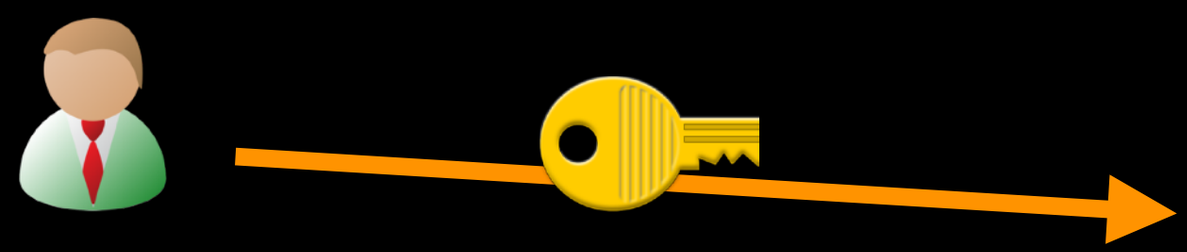
Respondent 



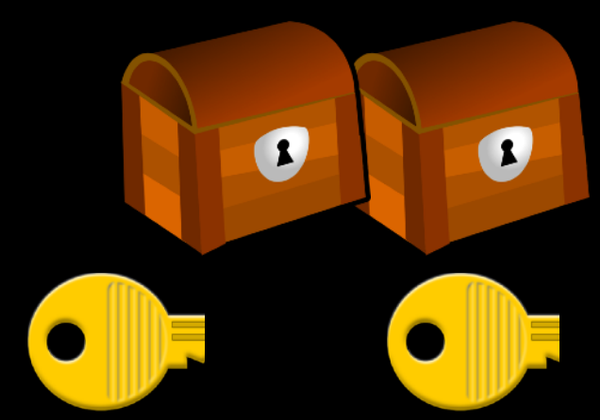
Tool 2: Fair Exchange

Clerk's Office

Petitioner



Respondent

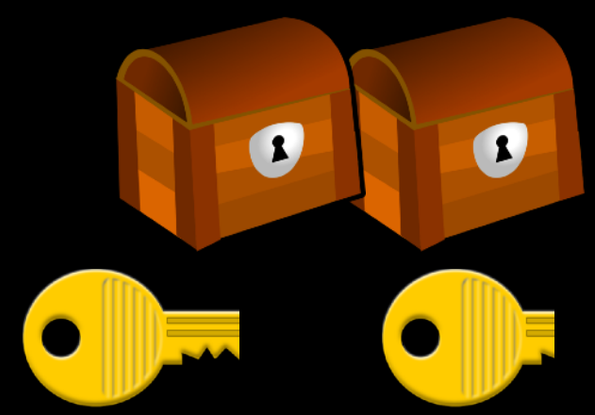


Tool 2: Fair Exchange

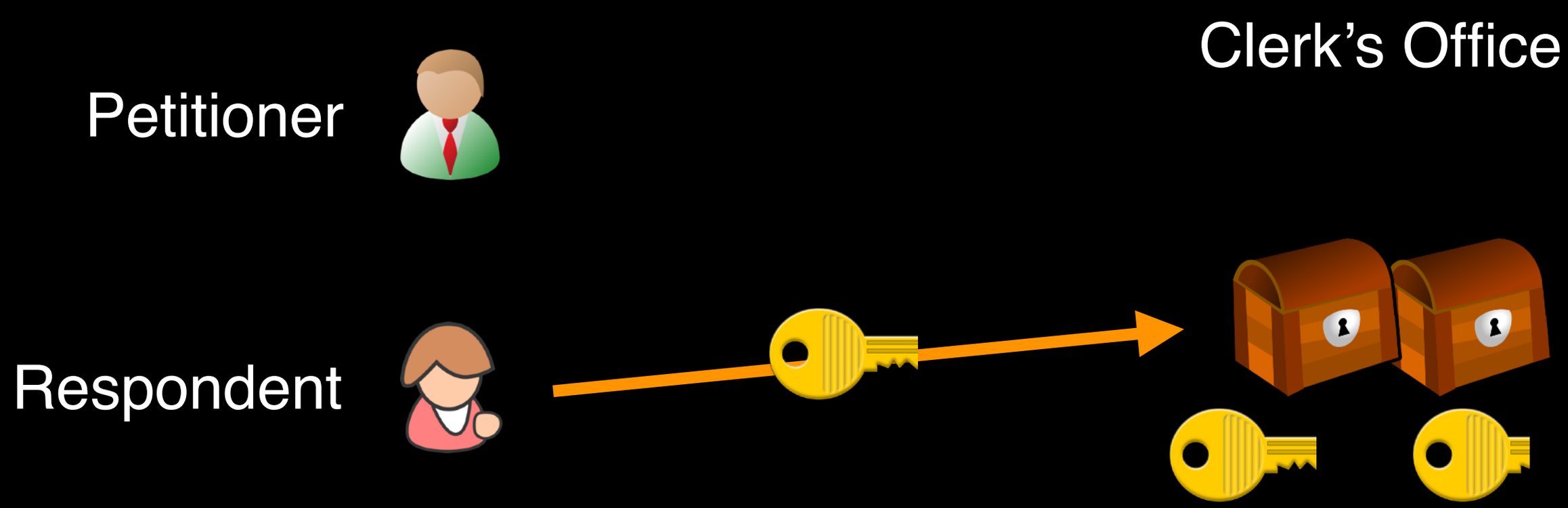
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

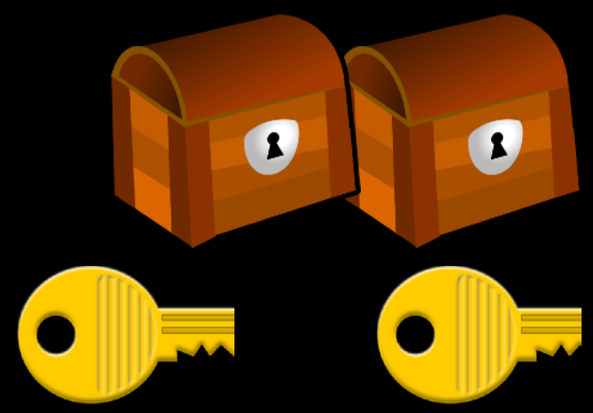


Tool 2: Fair Exchange

Clerk's Office

Petitioner 

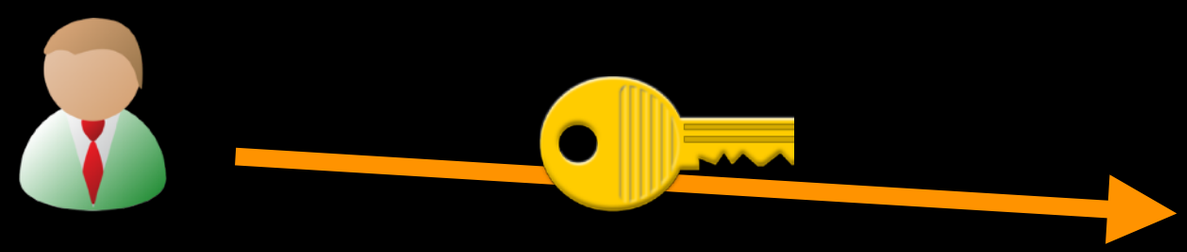
Respondent 



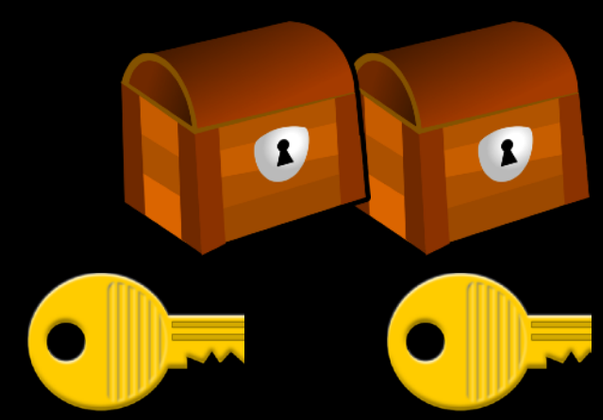
Tool 2: Fair Exchange

Clerk's Office

Petitioner



Respondent

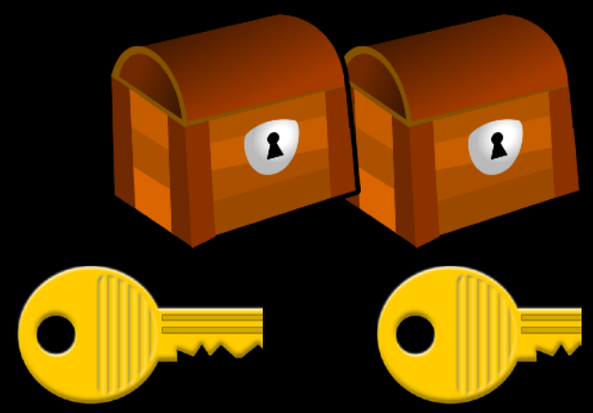


Tool 2: Fair Exchange

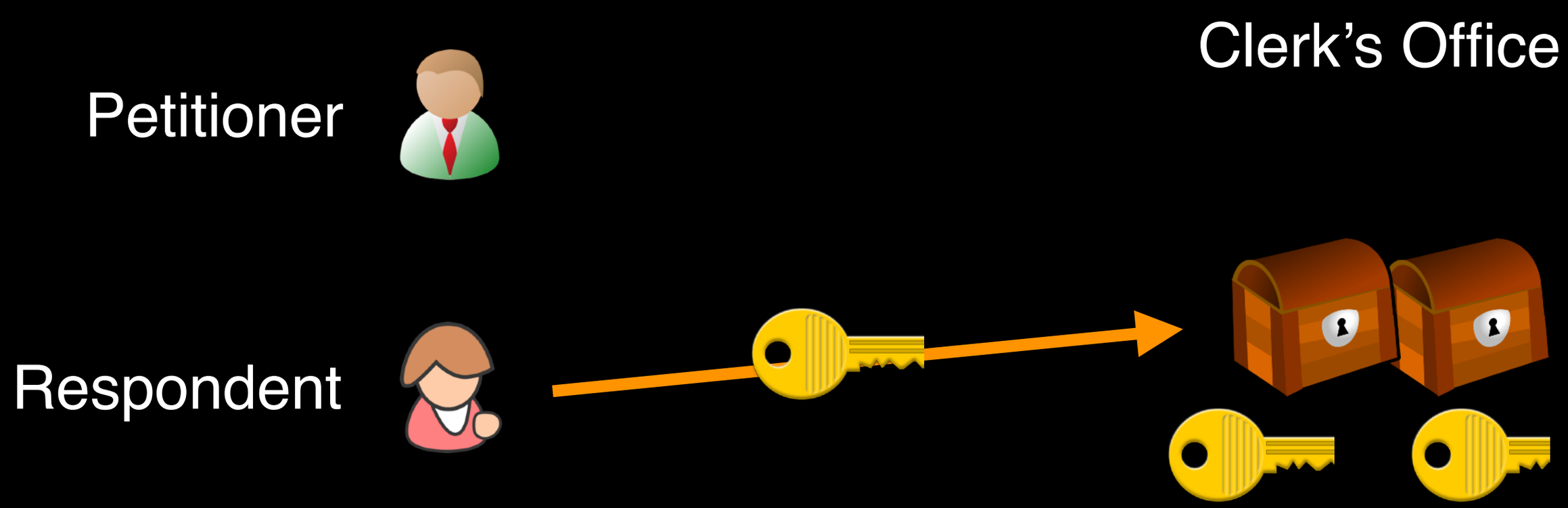
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

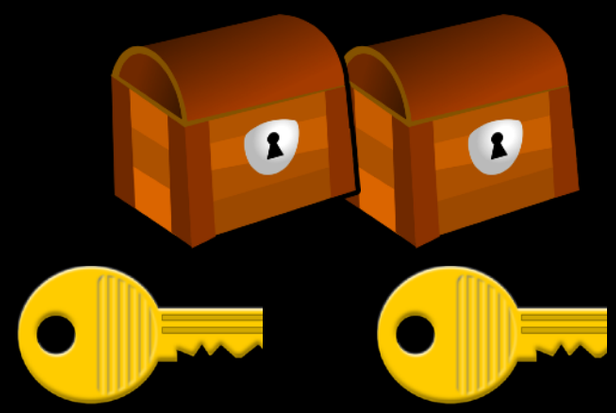


Tool 2: Fair Exchange

Clerk's Office

Petitioner 

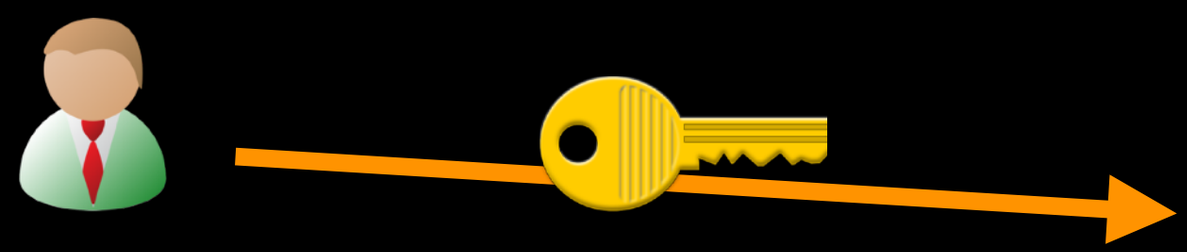
Respondent 



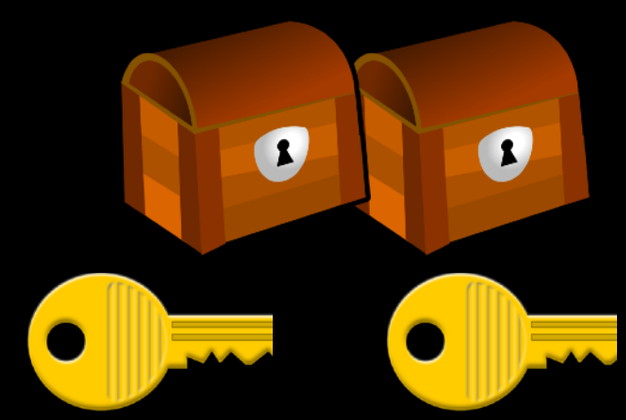
Tool 2: Fair Exchange

Clerk's Office

Petitioner



Respondent

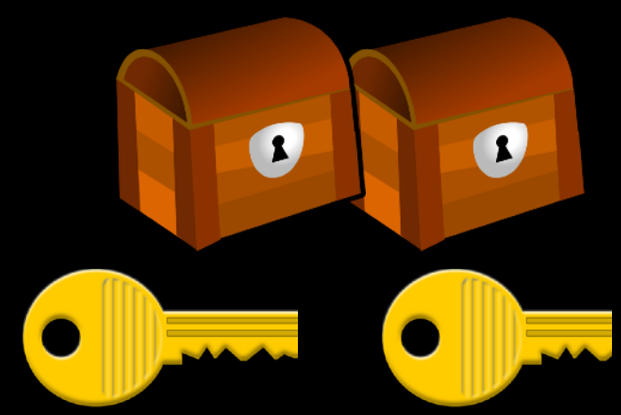


Tool 2: Fair Exchange

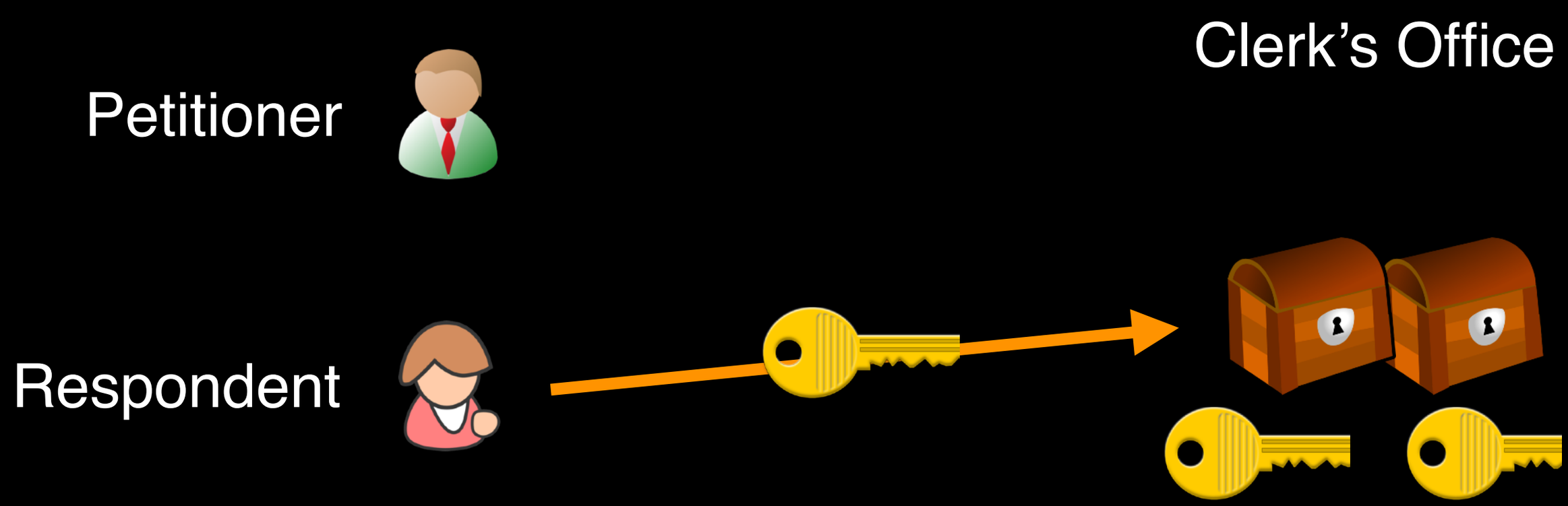
Clerk's Office

Petitioner 

Respondent 



Tool 2: Fair Exchange

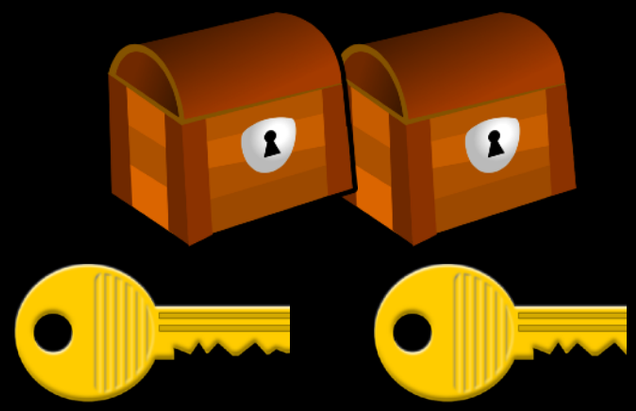


Tool 2: Fair Exchange

Clerk's Office

Petitioner 

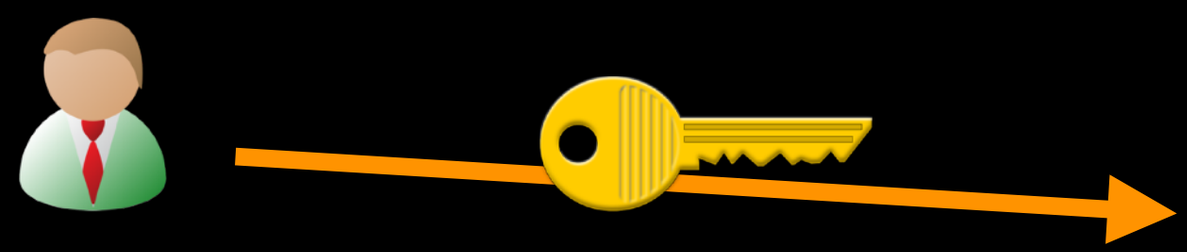
Respondent 



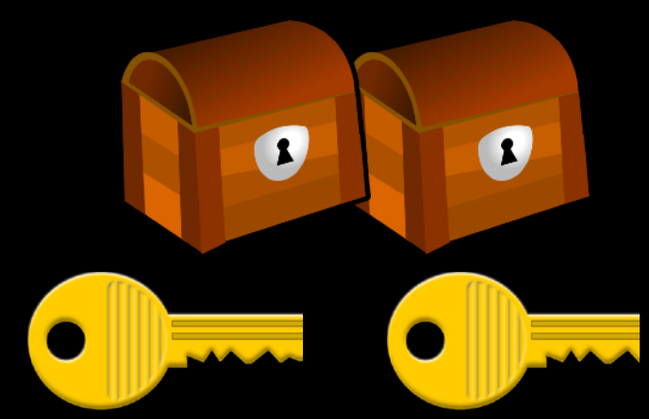
Tool 2: Fair Exchange

Clerk's Office

Petitioner



Respondent

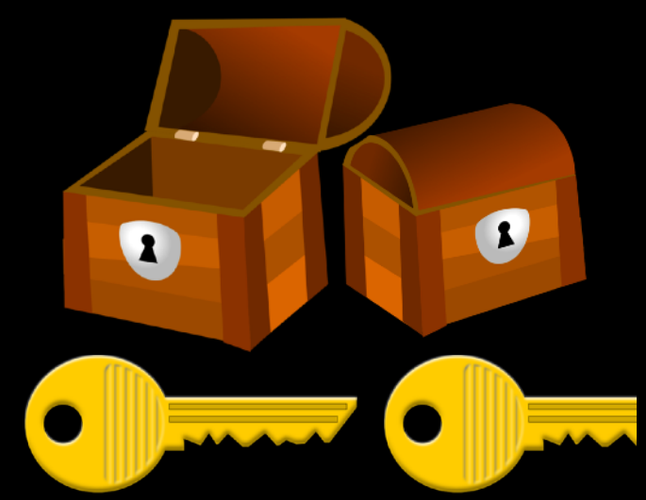


Tool 2: Fair Exchange

Clerk's Office

Petitioner

Respondent

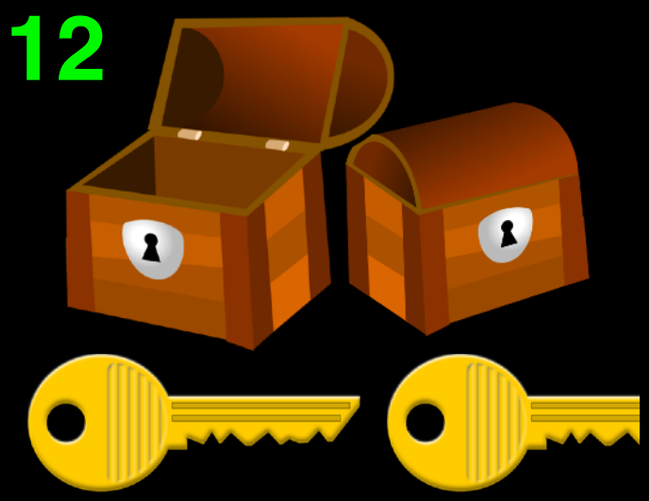


Tool 2: Fair Exchange

Clerk's Office

Petitioner

12



Respondent

Tool 2: Fair Exchange

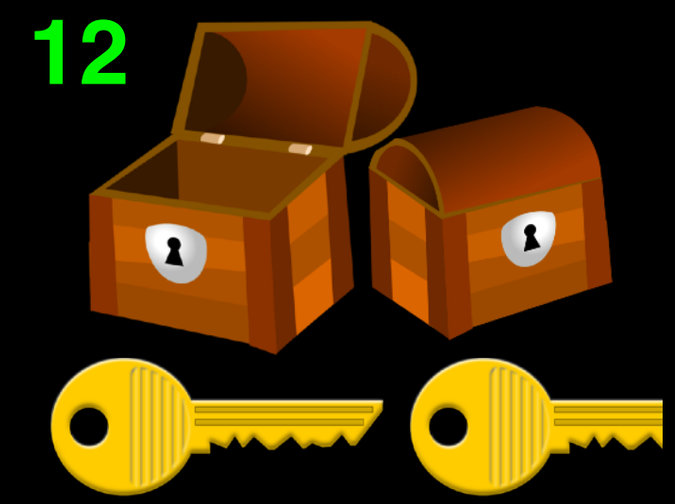
Clerk's Office

Petitioner

I lost my key!

Let's start over.

Respondent



Tool 2: Fair Exchange

Petitioner 

Respondent 

12

Tool 2: Fair Exchange

Petitioner 

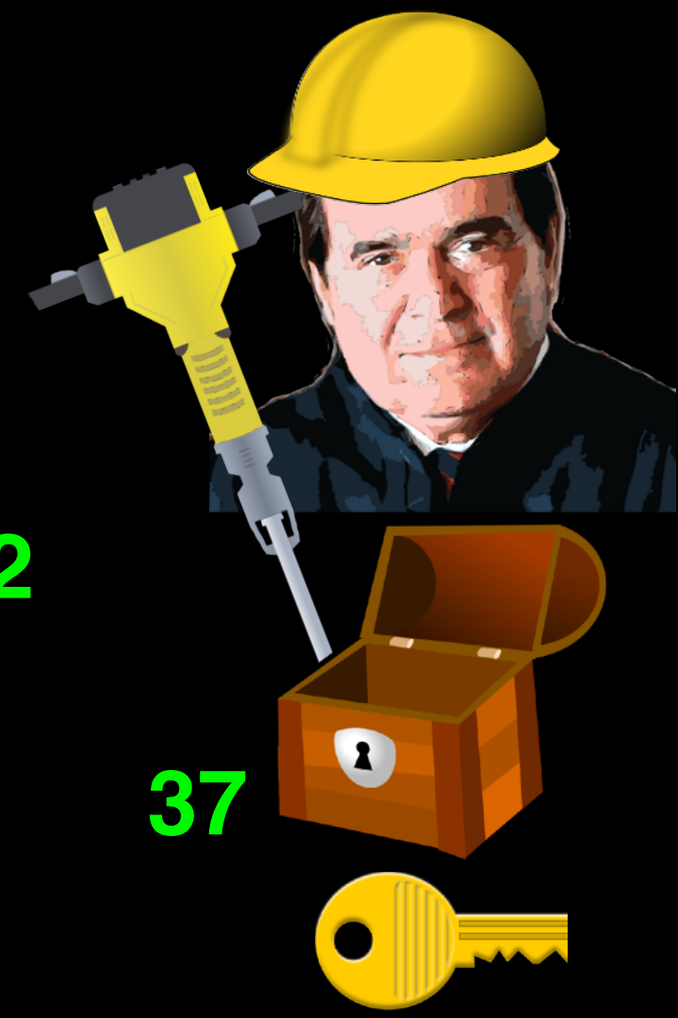
Respondent 

12

Tool 2: Fair Exchange

Petitioner 

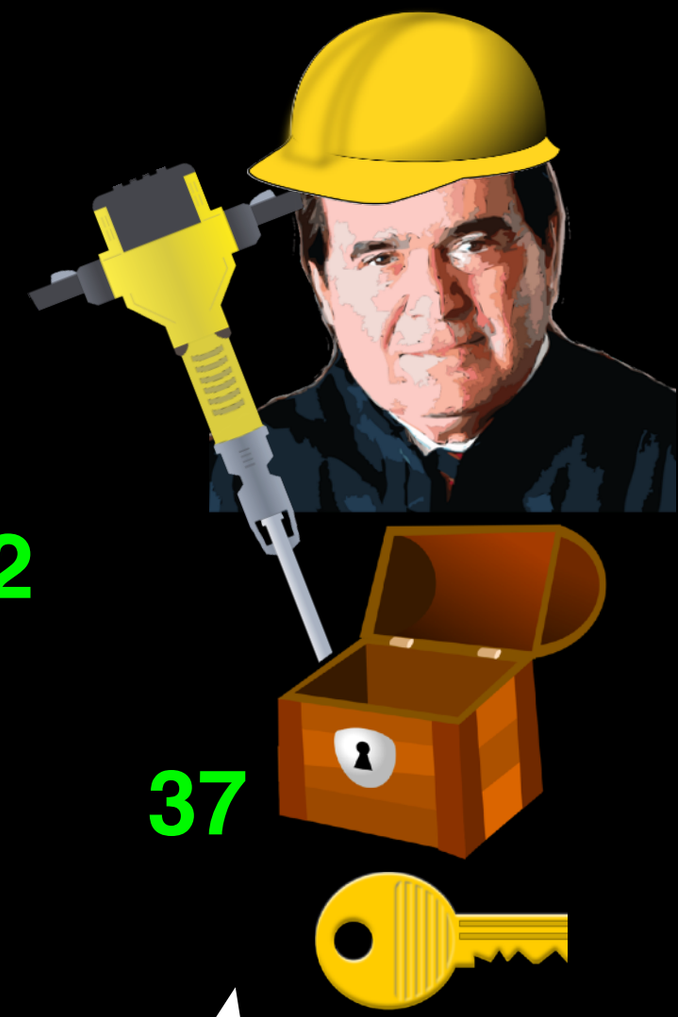
Respondent 



Tool 2: Fair Exchange

Petitioner 

Respondent 



You get panel number
12 + 37 = **49**.

Three Cryptographic Ideas

1. Coin-Flipping Protocols
2. **Fair Exchange**
3. Randomness Beacons

Three Cryptographic Ideas

1. Coin-Flipping Protocols
2. Fair Exchange
3. **Randomness Beacons**

Tool 3: Randomness Beacon

[Rabin 1983]

Tool 3: Randomness Beacon

[Rabin 1983]

- These first two solutions (coin-flipping and fair exchange) require **interaction** between the parties.

Tool 3: Randomness Beacon

[Rabin 1983]

- These first two solutions (coin-flipping and fair exchange) require **interaction** between the parties.
- An alternative approach is to derive panel assignments from a *public unpredictable randomness source.*

Tool 3: Randomness Beacon

[Rabin 1983]

- These first two solutions (coin-flipping and fair exchange) require **interaction** between the parties.
- An alternative approach is to derive panel assignments from a *public unpredictable randomness source.*
  - Tomorrow's closing prices of the NYSE and NASDAQ

Tool 3: Randomness Beacon

[Rabin 1983]

- These first two solutions (coin-flipping and fair exchange) require **interaction** between the parties.
- An alternative approach is to derive panel assignments from a *public unpredictable randomness source.*
  - Tomorrow's closing prices of the NYSE and NASDAQ
  - Front page of next week's *New York Times*

Tool 3: Randomness Beacon

[Rabin 1983]

- These first two solutions (coin-flipping and fair exchange) require **interaction** between the parties.
- An alternative approach is to derive panel assignments from a *public unpredictable randomness source.*
  - Tomorrow's closing prices of the NYSE and NASDAQ
  - Front page of next week's *New York Times*
  - Blocks in the Bitcoin network [Bonneau et al. 2015]

Tool 3: Randomness Beacon

[Rabin 1983]

- These first two solutions (coin-flipping and fair exchange) require **interaction** between the parties.

- An alternative approach is to derive panel assignments from a *public unpredictable randomness source.*
  - Tomorrow's closing prices of the NYSE and NASDAQ
  - Front page of next week's *New York Times*
  - Blocks in the Bitcoin network [Bonneau et al. 2015]

- Using a public randomness source allows <u>anyone</u> to verify the correct use of randomness.

Three Cryptographic Ideas

1. Coin-Flipping Protocols
2. Fair Exchange
3. Randomness Beacons

Are these schemes practical?

Are these schemes practical?

Maybe…

Are these schemes practical?

Maybe…

+ Many randomized government processes (panel assignments, tax audits, etc.) are computerized.

Are these schemes practical?

Maybe…

+ Many randomized government processes (panel assignments, tax audits, etc.) are computerized.
+ The cryptographic building blocks are all completely standard and are already in your PC.

Are these schemes practical?

Maybe…

+ Many randomized government processes (panel assignments, tax audits, etc.) are computerized.
+ The cryptographic building blocks are all completely standard and are already in your PC.

Maybe not…

Are these schemes practical?

Maybe…

+ Many randomized government processes (panel assignments, tax audits, etc.) are computerized.
+ The cryptographic building blocks are all completely standard and are already in your PC.

Maybe not…

– The software that runs our government is rarely open to public scrutiny.

Are these schemes practical?

Maybe…

+ Many randomized government processes (panel assignments, tax audits, etc.) are computerized.
+ The cryptographic building blocks are all completely standard and are already in your PC.

Maybe not…

– The software that runs our government is rarely open to public scrutiny.
– There is an incentive problem: codifying procedures removes flexibility.
  See: Kroll et al. *Accountable Algorithms*. 165 U. Pa. L. Rev. (2017).

Conclusions

Conclusions

The legal system, like everything else, is increasingly computerized.

Conclusions

The legal system, like everything else, is increasingly computerized.

- Risk: “Black-box” algorithms and processes may exhibit subtle biases and faults.

Conclusions

The legal system, like everything else, is increasingly computerized.

- Risk: “Black-box” algorithms and processes may exhibit subtle biases and faults.
- Hope: In theory, we can use cryptographic tools to make almost any algorithm open and accountable.

Conclusions

The legal system, like everything else, is increasingly computerized.

- Risk: “Black-box” algorithms and processes may exhibit subtle biases and faults.
- Hope: In theory, we can use cryptographic tools to make almost any algorithm open and accountable.

What will it take to turn this theory into practice?

Conclusions

The legal system, like everything else, is increasingly computerized.

- Risk: “Black-box” algorithms and processes may exhibit subtle biases and faults.
- Hope: In theory, we can use cryptographic tools to make almost any algorithm open and accountable.

What will it take to turn this theory into practice?

Help out!

Do you know anyone who could help us get access to the code used to generate panel assignments in the Courts of Appeals? Email us!

`henrycg@cs.stanford.edu`
`keithw@cs.stanford.edu`

Conclusions

The legal system, like everything else, is increasingly computerized.

- Risk: “Black-box” algorithms and processes may exhibit subtle biases and faults.
- Hope: In theory, we can use cryptographic tools to make almost any algorithm open and accountable.

What will it take to turn this theory into practice?

Help out!

Do you know anyone who could help us get access to the code used to generate panel assignments in the Courts of Appeals? Email us!

`henrycg@cs.stanford.edu`
`keithw@cs.stanford.edu`

Want more?

Cybersecurity: A Legal and Technical Perspective
LAW 4004 (Spring 2017) - Boneh and Granick