# Hiring and Bias 2

**SOC 121D: People Analytics**

**Austin van Loon**

# Some notes from the Week 2 Response Survey

- Fist off, thank you all so much for this feedback; I genuinely appreciate it
- From now on…
  - I'll be sure to structure group work more (e.g., "chunking" the group work with class check-ins and discussions)
  - I'll try to include more visual representations of examples in the slides
- I've discussed with some faculty, and I've decided I won't put recordings of the lecture online
- Hopefully, we're through the part of the class that is unexpectedly math-heavy (though there might be a *little* more math during the digital trace data week)

# Overview of the Week

- Hiring and Bias 1
  - Common Hiring Practices
  - Evidence for Discrimination in the Labor Market
  - Theories of Labor Market Discrimination
- Hiring and Bias 2
  - Can Algorithms Improve Hiring?
  - When/How Might Algorithms Re-create/Amplify/Diminish Bias?
  - Models of Human/Algorithm Hiring Models

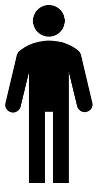
# Hiring as a Prediction Problem

- **Observations:** previous hiring decisions (e.g., your employees or secondary data)
- **Features:** all the things measured about applicants pre-hire (we discussed this last class)
    - Personality profile
    - Integrity test score
    - Emotional intelligence score
- **Outcome:** Thing you want future employees to (not) have
    - Performance (we'll talk more about next week)
    - Firing/Quitting (after X days)

| Name | Extraversion | Integrity | Quit? |
| --- | --- | --- | --- |
| Joey | 3 | -1 | Yes |
| Ross | -1 | 1 | No |
| Rachel | 2.5 | -1 | Yes |
| Monica | -0.25 | 3 | Yes |
| Chandler | -0.3 | 0 | No |
| Phoebe | 3 | 2 | No |

| Name | Extraversion | Integrity | Quit? |
| --- | --- | --- | --- |
| Gunter | 3 | -1 | ? |
| Janice | -1 | 1 | ? |
| Carol | 2.5 | -1 | ? |

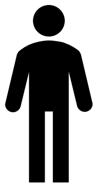
| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Joey | 3 | -1 | Yes |
| Ross | -1 | 1 | No |
| Rachel | 2.5 | -1 | Yes |
| Monica | -0.25 | 3 | Yes |
| Chandler | -0.3 | 0 | No |
| Phoebe | 3 | 2 | No |



| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Gunter | 3 | -1 | ? |
| Janice | -1 | 1 | ? |
| Carol | 2.5 | -1 | ? |

| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Joey | 3 | -1 | Yes |
| Ross | -1 | 1 | No |
| Rachel | 2.5 | -1 | Yes |
| Monica | -0.25 | 3 | Yes |
| Chandler | -0.3 | 0 | No |
| Phoebe | 3 | 2 | No |

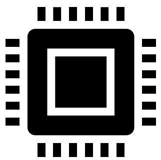
| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Gunter | 3 | -1 | 1 |
| Janice | -1 | 1 | 0 |
| Carol | 2.5 | -1 | 0 |



| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Joey | 3 | -1 | Yes |
| Ross | -1 | 1 | No |
| Rachel | 2.5 | -1 | Yes |
| Monica | -0.25 | 3 | Yes |
| Chandler | -0.3 | 0 | No |
| Phoebe | 3 | 2 | No |

| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Gunter | 3 | -1 | ? |
| Janice | -1 | 1 | ? |
| Carol | 2.5 | -1 | ? |

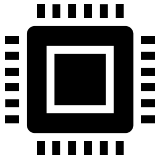
| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Joey | 3 | -1 | Yes |
| Ross | -1 | 1 | No |
| Rachel | 2.5 | -1 | Yes |
| Monica | -0.25 | 3 | Yes |
| Chandler | -0.3 | 0 | No |
| Phoebe | 3 | 2 | No |



| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Gunter | 3 | -1 | ? |
| Janice | -1 | 1 | ? |
| Carol | 2.5 | -1 | ? |

| Name | Extraversion | Integrity | Quit? |
|---|---|---|---|
| Joey | 3 | -1 | Yes |
| Ross | -1 | 1 | No |
| Rachel | 2.5 | -1 | Yes |
| Monica | -0.25 | 3 | Yes |
| Chandler | -0.3 | 0 | No |
| Phoebe | 3 | 2 | No |

| Name | Extraversion | Integrity | P(Quit?) |
|---|---|---|---|
| Gunter | 3 | -1 | 0.3 |
| Janice | -1 | 1 | 0.6 |
| Carol | 2.5 | -1 | 0.01 |



# Some Issues

- Problem of censored outcomes (those who you didn't hire)
  - One solution: only apply to those you might reasonably otherwise hire
  - Another: build in preference for "exploration"
  - Exploit random variation when possible (e.g., in leniency of different hiring managers)
- Variability in feature/outcome relationships
  - Be constantly updating your model as more data is available (build the pipeline!)
  - Build in preference for "exploration"
  - Be sure to check model performance by department, position, etc.
- Often difficult to quantify "how good" an employee is (wait till next week)
- You need A LOT of data to do all this well!
  - Some models generally require less than others (e.g., regression < NN)

# Hiring: More than Prediction?

- Sometimes hiring is meant to *change* something about your organization
  - Increase diversity
  - Promote a different organizational culture
- Sometimes hiring is meant to *signal* something to others in the field or to *express* organizational values
- You could hypothetically introduce these factors into the algorithm's cost function, but this can be difficult

# Two Basic Ways to Use Algorithms in Hiring

- Algorithms for measurement
  - Can handle varied information (e.g., social media data, voice recordings, large numbers of test responses) to predict things we want to know about candidates (e.g., personality)
  - Can help us customize and refine traditional tools for our organizational context (e.g., which situational judgement test are predictive of employee behavior?)
- Algorithms for decision-making
  - When humans are faced with a lot of variable information, they can get confused and over-weight unimportant information. This is the strength of algorithms: systematically and consistently weights each feature to optimally predict the outcome.
  - Feed measures into algorithm, let it weight information and make recommendations

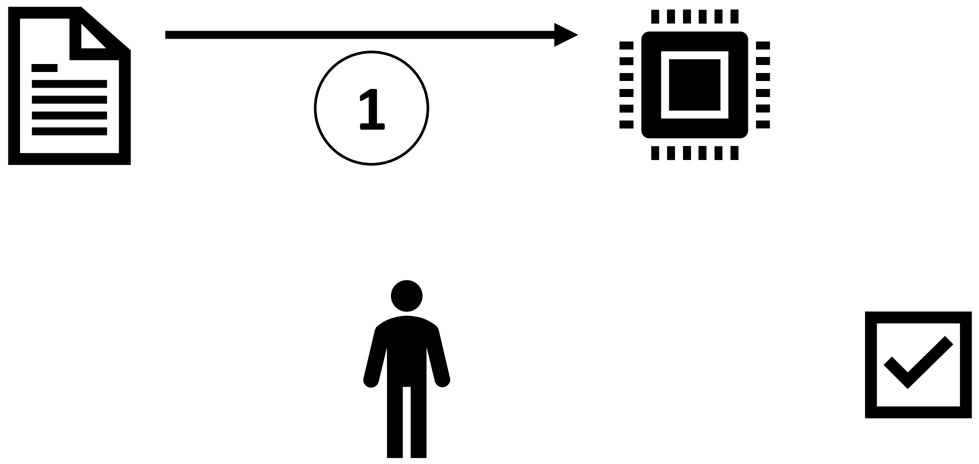
# When do algorithms re-create bias?

- **Outsider bias**: When the input data is created from biases (e.g., if women quit more often because your firm is full of misogynist men, the algorithm might predict that women are more likely to quit)
  - Explore behavior of model with/without characteristics (e.g., gender)
  - Explore the causes behind the correlations (other methodologies needed)
- **Insider bias**: When biased decisions go into making the algorithm (e.g., when features such as SES, address, and name are included as "not race" proxies)
  - Have a diverse team of people working on the algorithm
  - Be sure that the team allows for open communication and decisions are held to account
- When humans aren't looking for bias
  - So many ways bias could be introduced
  - Ultimately, while humans create bias, we're also the only ones who can see it
  - Good to have independent entities examining/auditing algorithms for bias

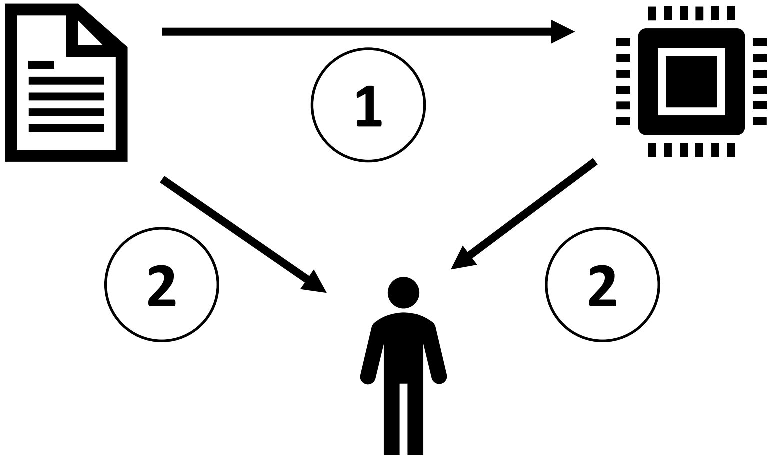
# Human/Algorithm Hiring Models

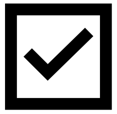
1. **Algorithms-as-Advisors model**: Algorithmic predictions are input to manager judgement
2. **Algorithms-as-Decider model**: Manager input can be features used by algorithm, who makes ultimate decision
3. **Justified deviation model**: Manager can deviate from algorithms decision if they provide public (to some “group”) account of decision

# Algorithms-as-Advisors model

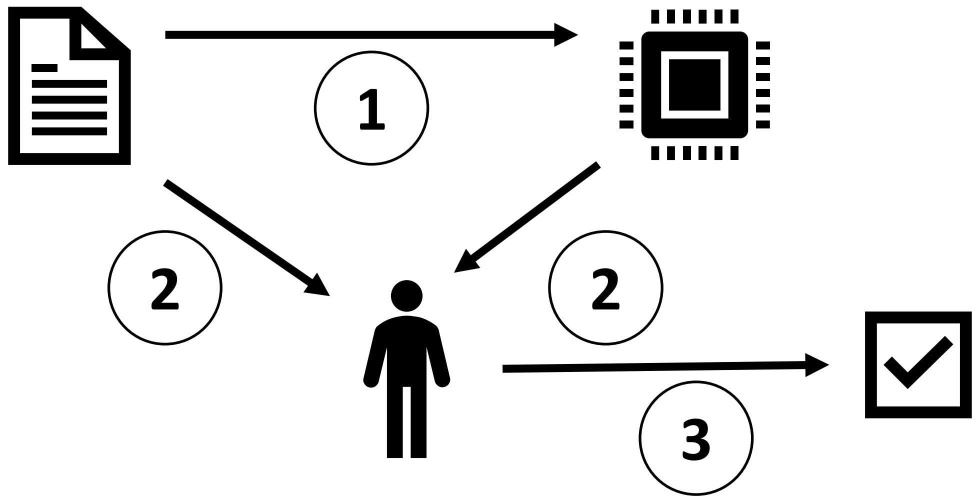


# Algorithms-as-Advisors model

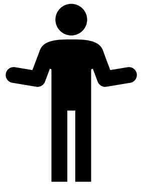


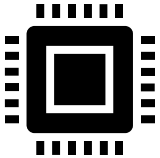



# Algorithms-as-Advisors model

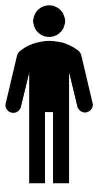


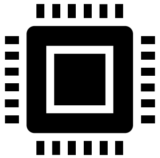
| Name | Extraversion | Integrity | Quit? |
| --- | --- | --- | --- |
| Gunter | 3 | -1 | ? |
| Janice | -1 | 1 | ? |
| Carol | 2.5 | -1 | ? |



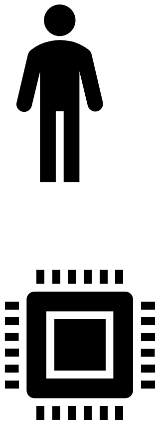


| Name | Extraversion | Integrity | P(Quit?) |
|---|---|---|---|
| Gunter | 3 | -1 | 0.3 |
| Janice | -1 | 1 | 0.6 |
| Carol | 2.5 | -1 | 0.01 |

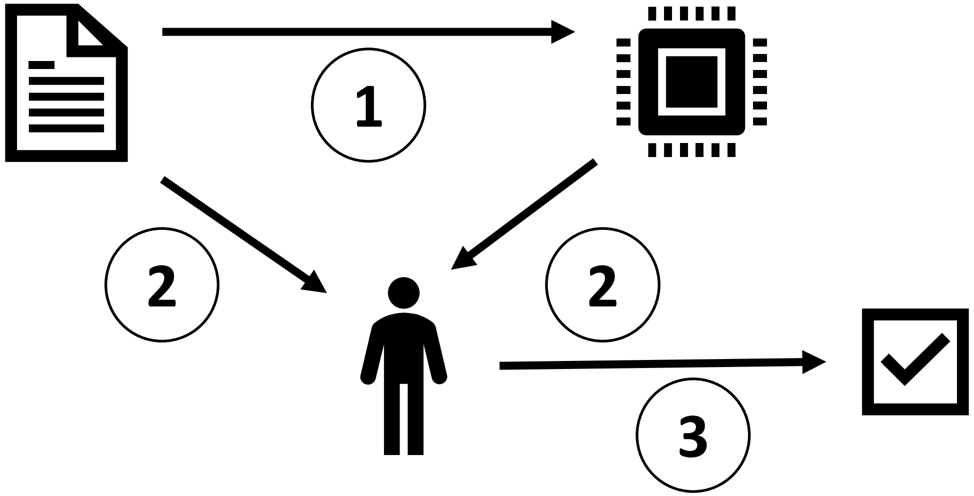




| Name | Extraversion | Integrity | | P(Quit?) |
|---|---|---|---|---|
| Gunter | 3 | -1 | | 0.3 |
| Janice | -1 | 1 | | 0.6 |
| Carol | 2.5 | -1 | | 0.01 |



# Algorithms-as-Advisors model



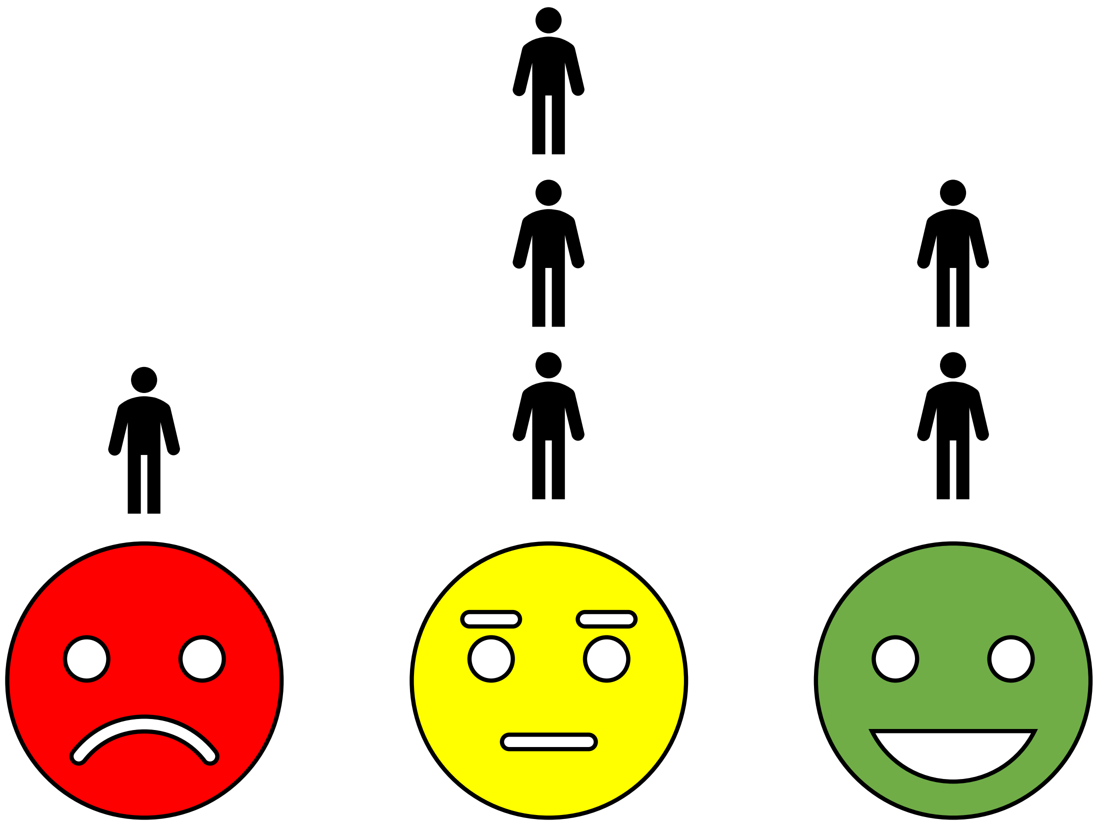

- Easiest buy-in ("We're empowering managers with AI…")
- Gives the most authority to the manager
- Downside: managers might only heed predictions when they confirm their own (potentially biased) intuitions/preferences
- Susceptible to taste-based and stereotype/status discrimination

# Discretion in Hiring*

Mitchell Hoffman, Lisa B Kahn, Danielle Li

*The Quarterly Journal of Economics*, Volume 133, Issue 2, May 2018, Pages 765–800, https://doi.org/10.1093/qje/qjx042

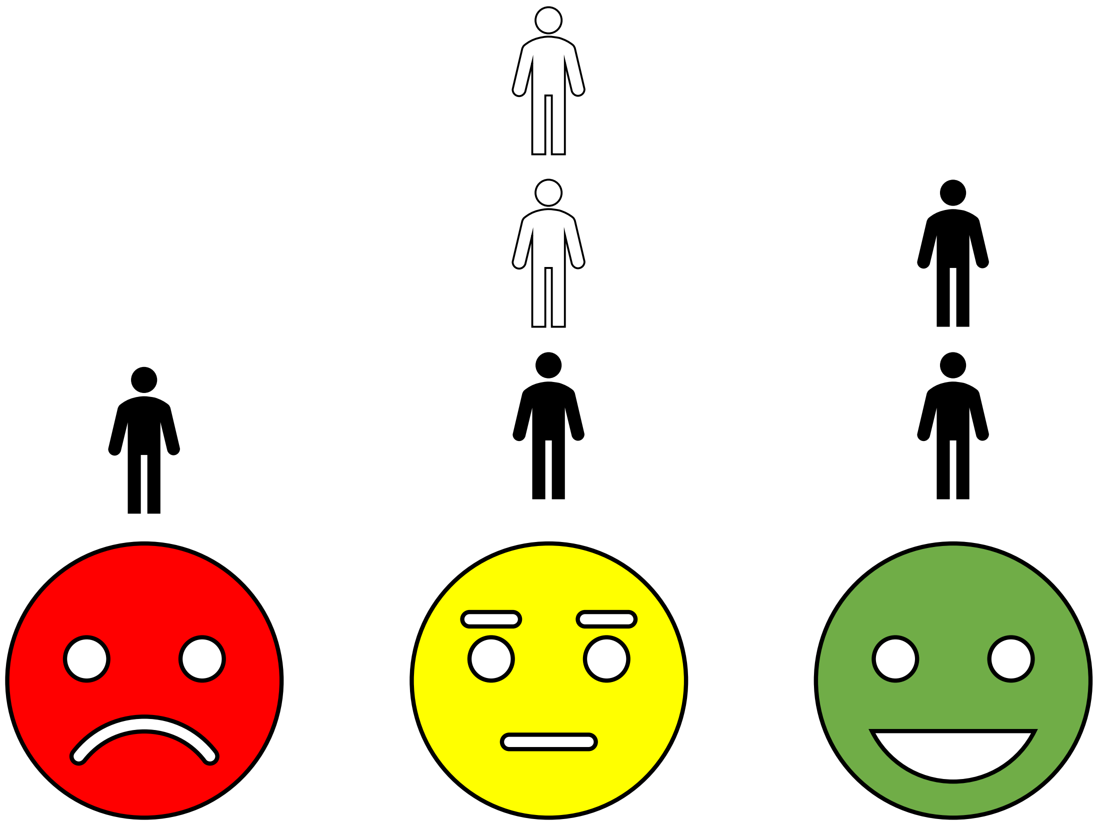
**Published:** 10 October 2017



# Discretion in Hiring*

Mitchell Hoffman, Lisa B Kahn, Danielle Li

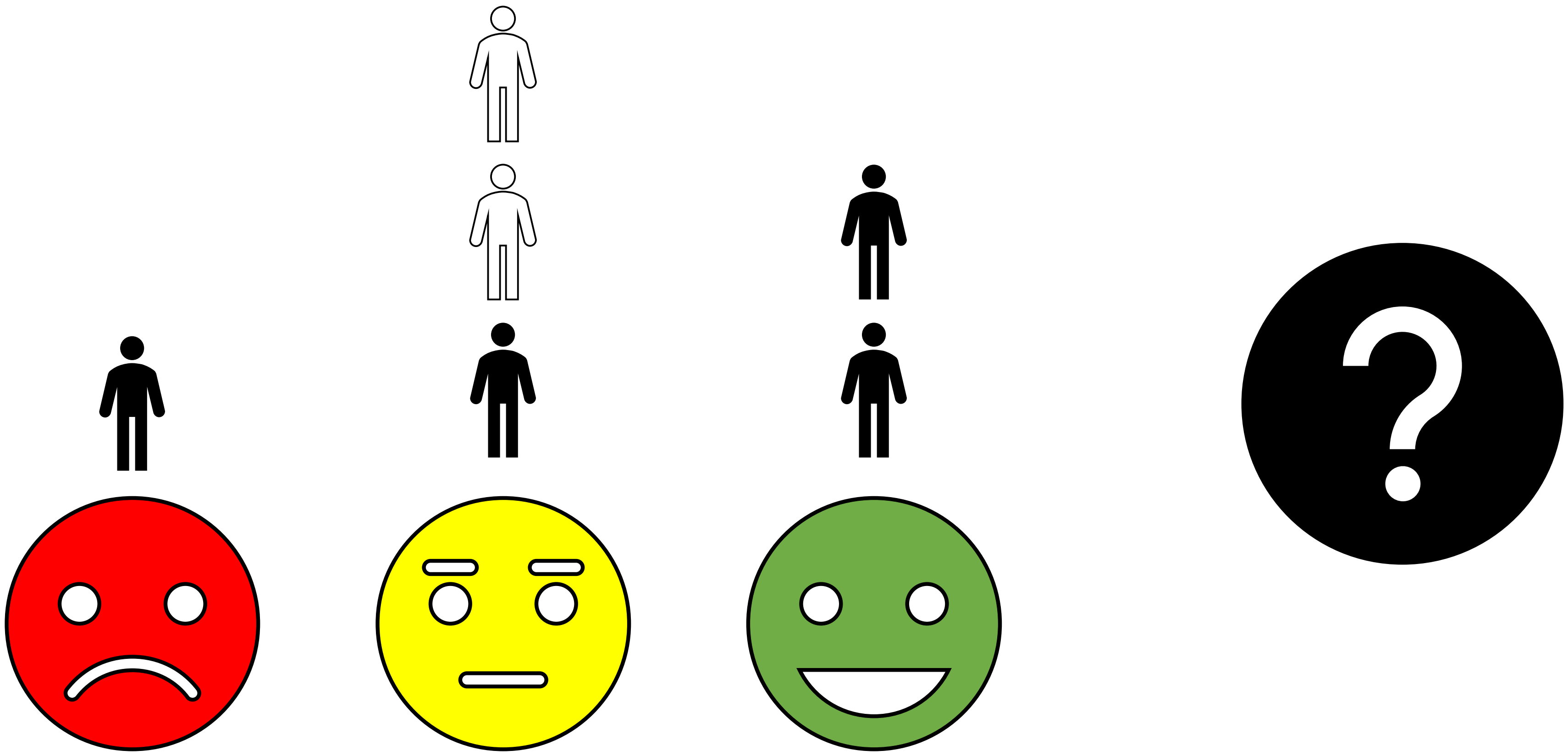
*The Quarterly Journal of Economics*, Volume 133, Issue 2, May 2018, Pages 765–800, https://doi.org/10.1093/qje/qjx042
**Published:** 10 October 2017



# Discretion in Hiring*

Mitchell Hoffman, Lisa B Kahn, Danielle Li

*The Quarterly Journal of Economics*, Volume 133, Issue 2, May 2018, Pages 765–800, https://doi.org/10.1093/qje/qjx042
**Published:** 10 October 2017

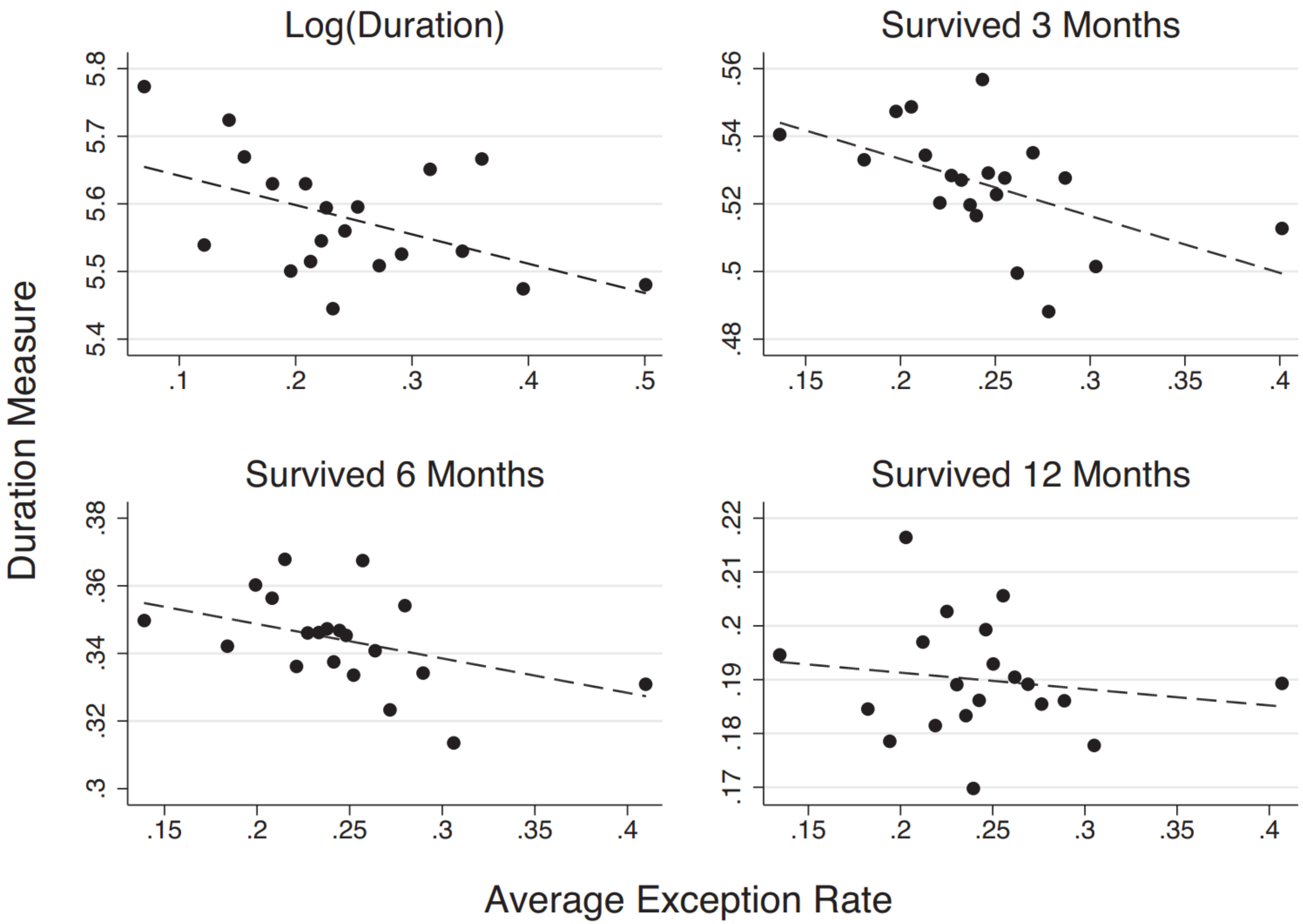


# Discretion in Hiring*

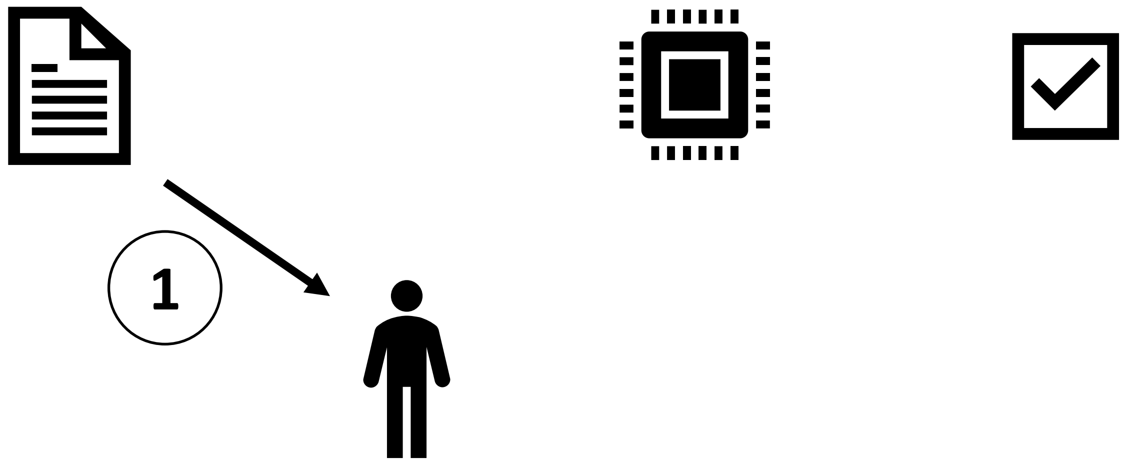
Mitchell Hoffman, Lisa B Kahn, Danielle Li

*The Quarterly Journal of Economics*, Volume 133, Issue 2, May 2018, Pages 765–800, https://doi.org/10.1093/qje/qjx042

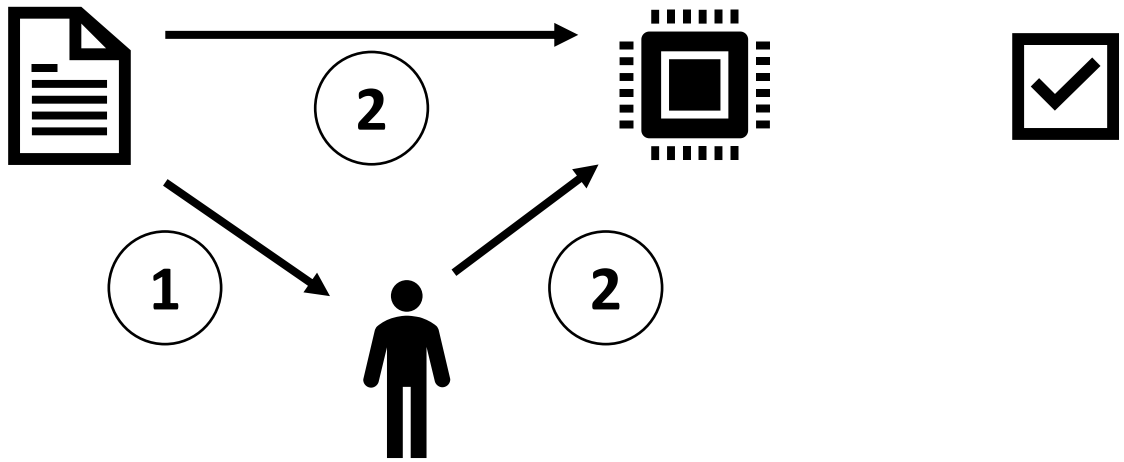
**Published:** 10 October 2017



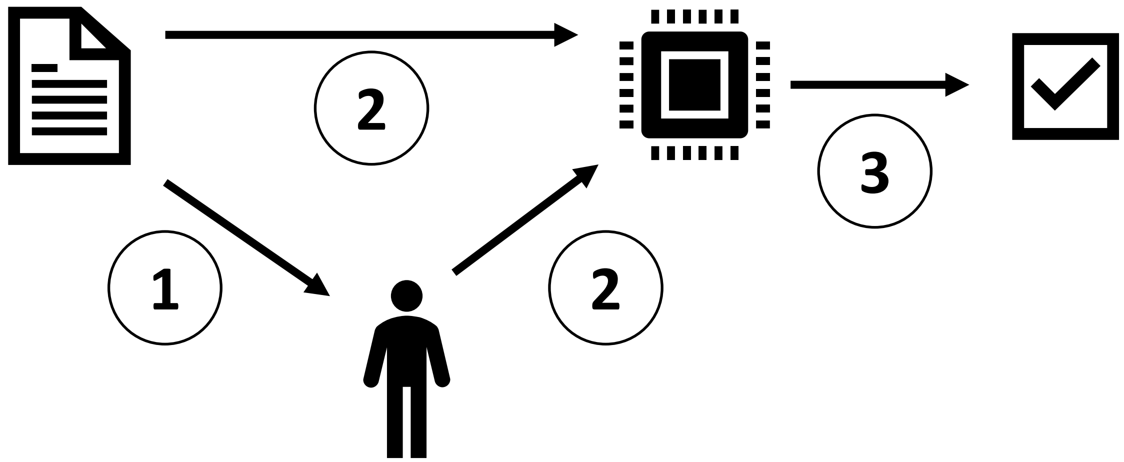
# Algorithms-as-Decider model



# Algorithms-as-Decider model

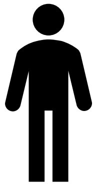


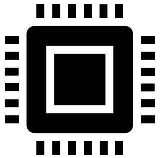
# Algorithms-as-Decider model



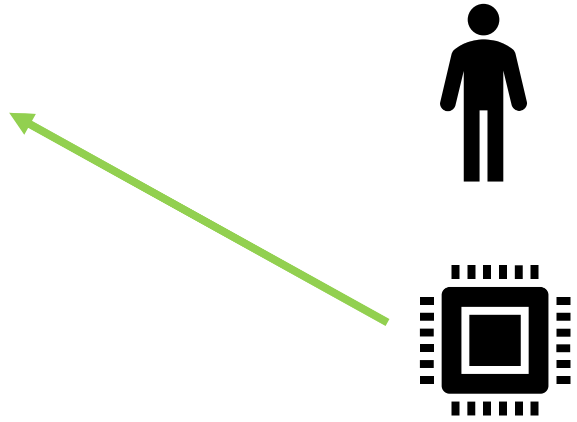
| Name | Extraversion | Integrity |
|---|---|---|
| Gunter | 3 | -1 |
| Janice | -1 | 1 |
| Carol | 2.5 | -1 |

| Like? |
|---|
| 1 |
| 0 |
| 0 |

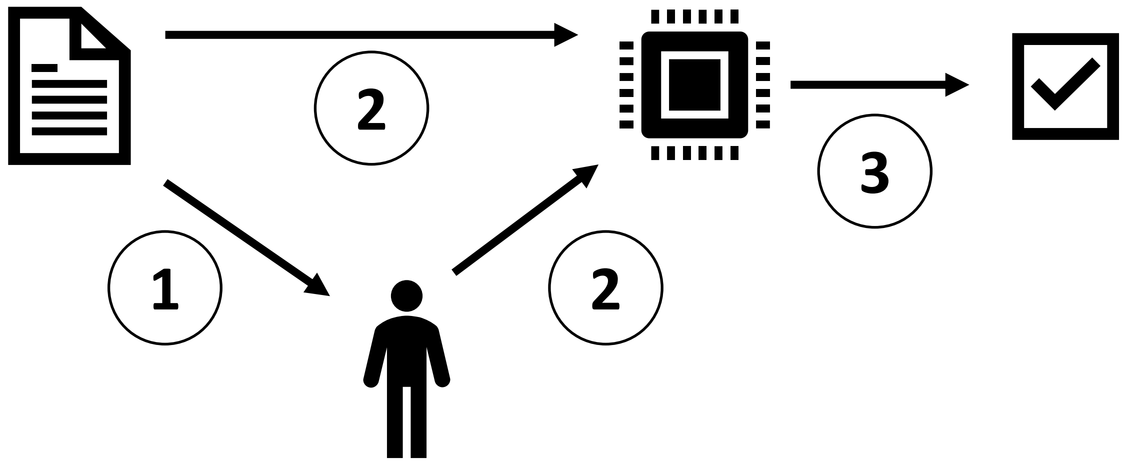




| Name | Extraversion | Integrity | Like? |
|---|---|---|---|
| Gunter | 3 | -1 | 1 |
| Janice | -1 | 1 | 0 |
| Carol | 2.5 | -1 | 0 |

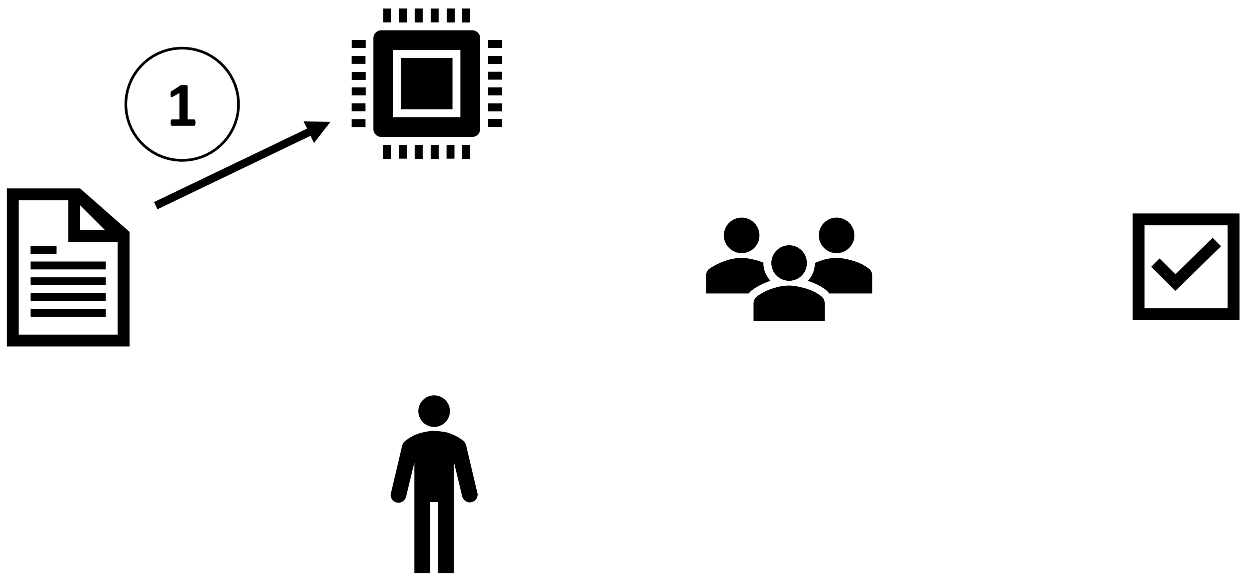


# Algorithms-as-Decider model

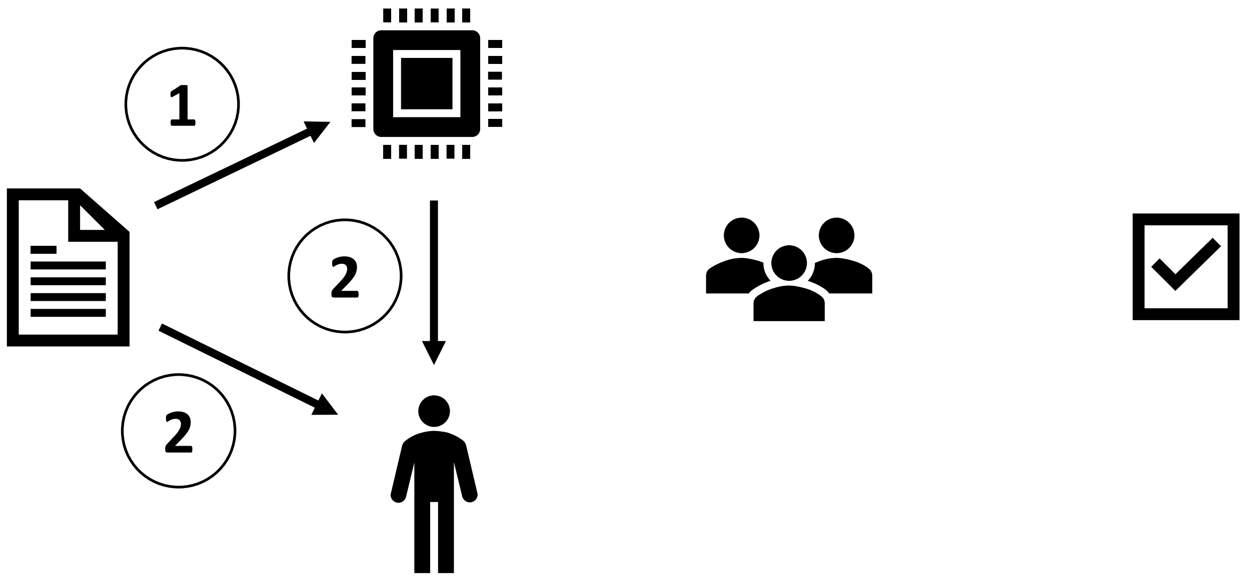



- Could average several/many managers opinions or put in each separately
- Gives the most authority to algorithm
- Downside: if algorithm re-creates or intensifies bias, this will become cemented across the all hires
- Susceptible to statistical and stereotypes/status discrimination

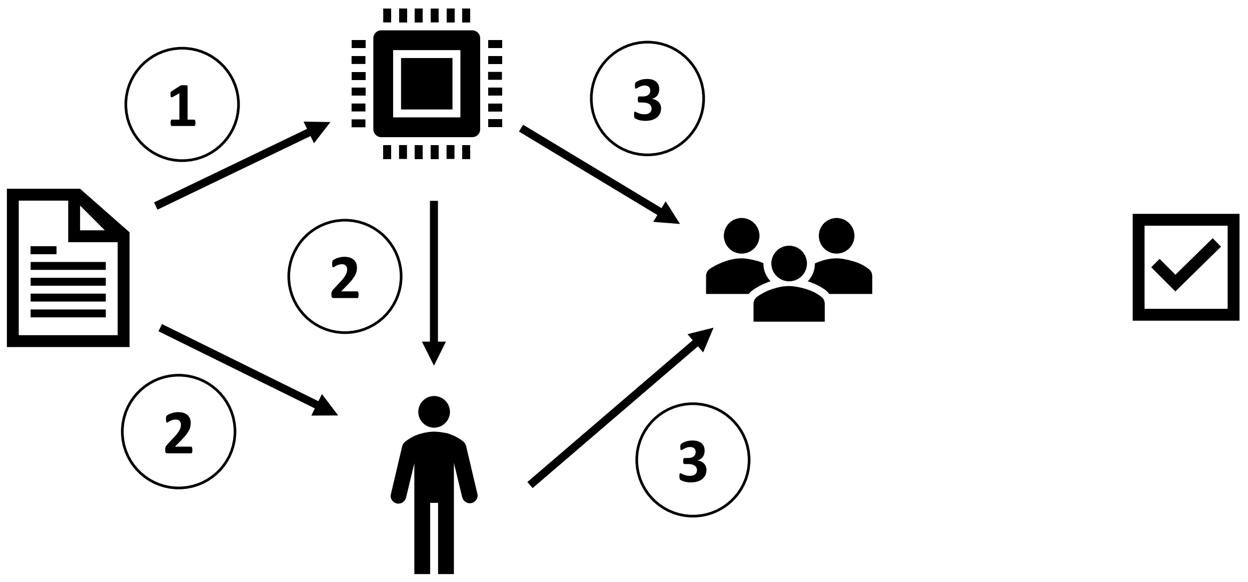
# Justified deviation model



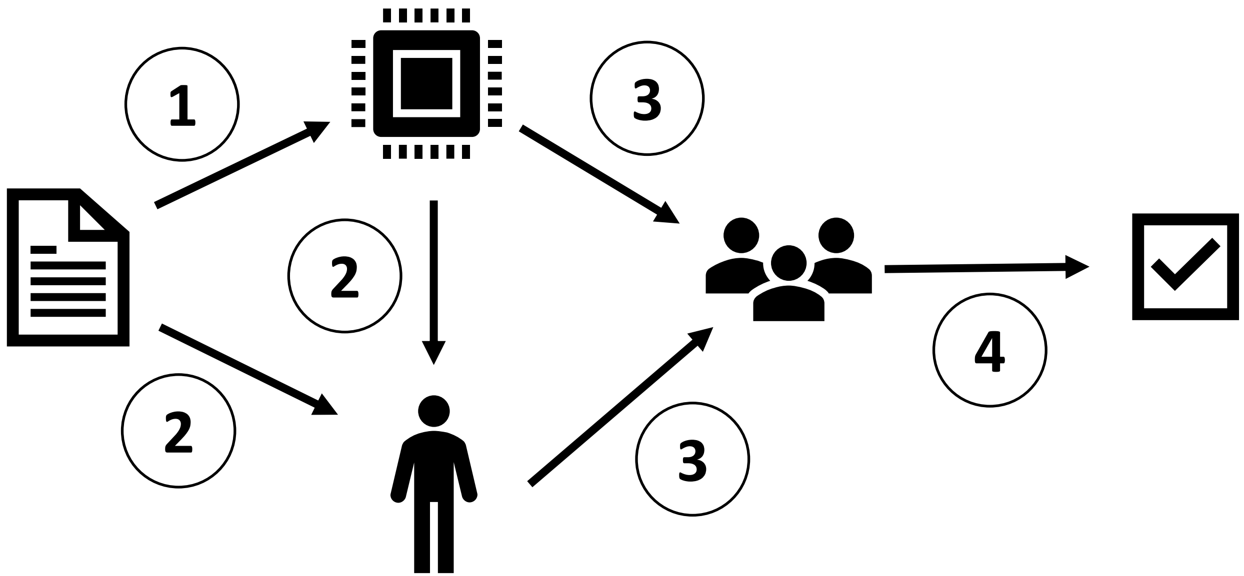
# Justified deviation model



# Justified deviation model

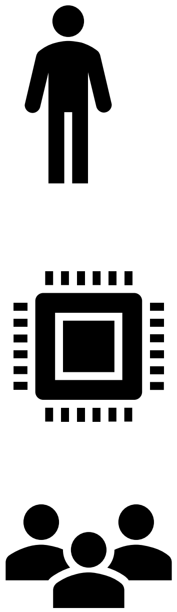


# Justified deviation model

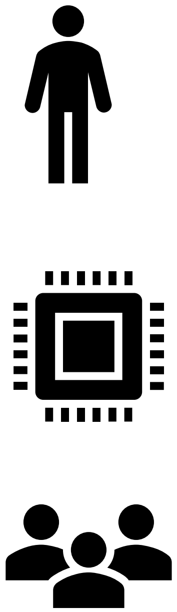


| Name | Extraversion | Integrity |
|---|---|---|
| Gunter | 3 | -1 |
| Janice | -1 | 1 |
| Carol | 2.5 | -1 |

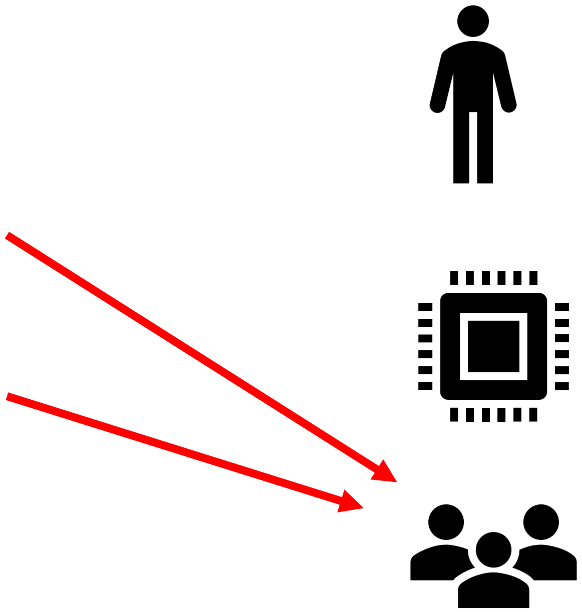
| Algorithm | Manager |
|---|---|
| 0 | ? |
| 1 | ? |
| 1 | ? |



| Name | Extraversion | Integrity | Algorithm | Manager |
|---|---|---|---|---|
| Gunter | 3 | -1 | 0 | 1 |
| Janice | -1 | 1 | 1 | 1 |
| Carol | 2.5 | -1 | 1 | 0 |

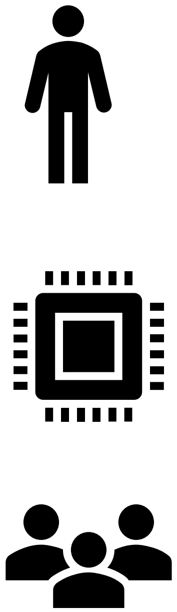


| Name | Extraversion | Integrity | Algorithm | Manager |
|---|---|---|---|---|
| Gunter | 3 | -1 | 0 | 1 |
| Janice | -1 | 1 | 1 | 1 |
| Carol | 2.5 | -1 | 1 | 0 |

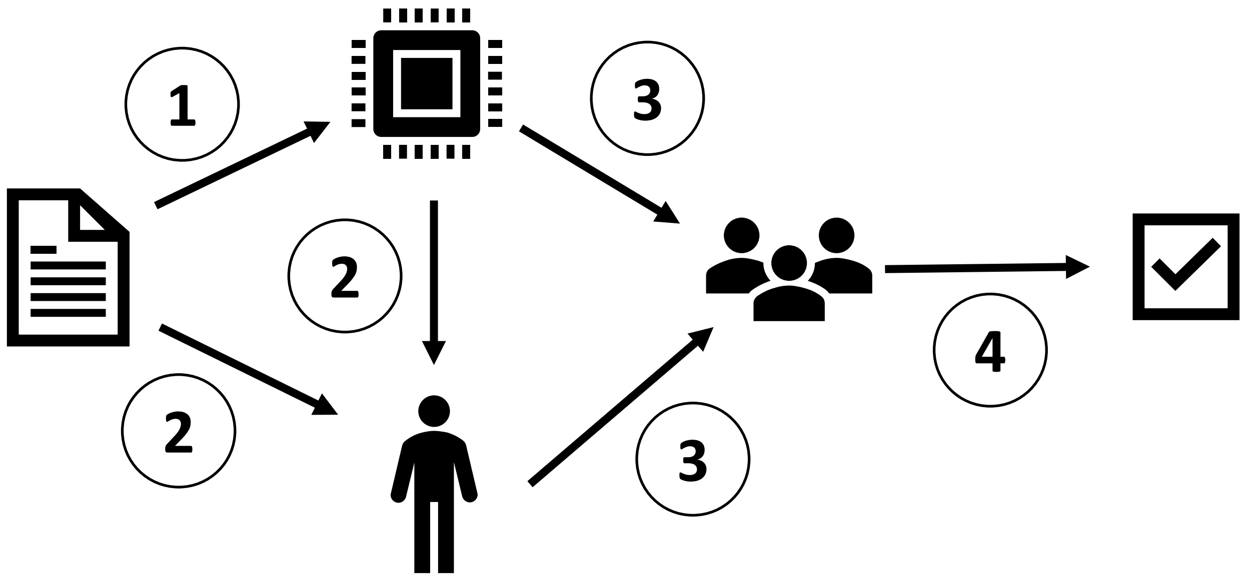


| Name | Extraversion | Integrity |
|---|---|---|
| Gunter | 3 | -1 |
| Janice | -1 | 1 |
| Carol | 2.5 | -1 |

| Algorithm | Manager |
|---|---|
| 0 | 1 |
| 1 | 1 |
| 1 | 0 |



# Justified deviation model




- Algorithm and group minimize taste-based discrimination
- Public justifications minimize stereotype/status
- Downside: depending on deviation rate, this can be the most labor-intensive
- Susceptible to (taste-based x statistical) discrimination

# Please fill out the survey at: tinyurl.com/paweek3

When you’re done, you’re free to go.

Have a great weekend!