

Neuroethics: From Moral Concern to Moral Constraint

Fiery Cushman



\$

vs

M



Neuroscience

$$\text{\$} = \text{M}$$

Neuroscience



\$ = M

Neuroscience

Psychology



$\$ = \overset{\circ}{M}$

$\$ \neq \overset{\circ}{M}$



$$Q_{(s,a)} = Q_{(s,a)} + \alpha(R_s + \gamma Q_{(s',a')} - Q_{(s,a)})$$

experience

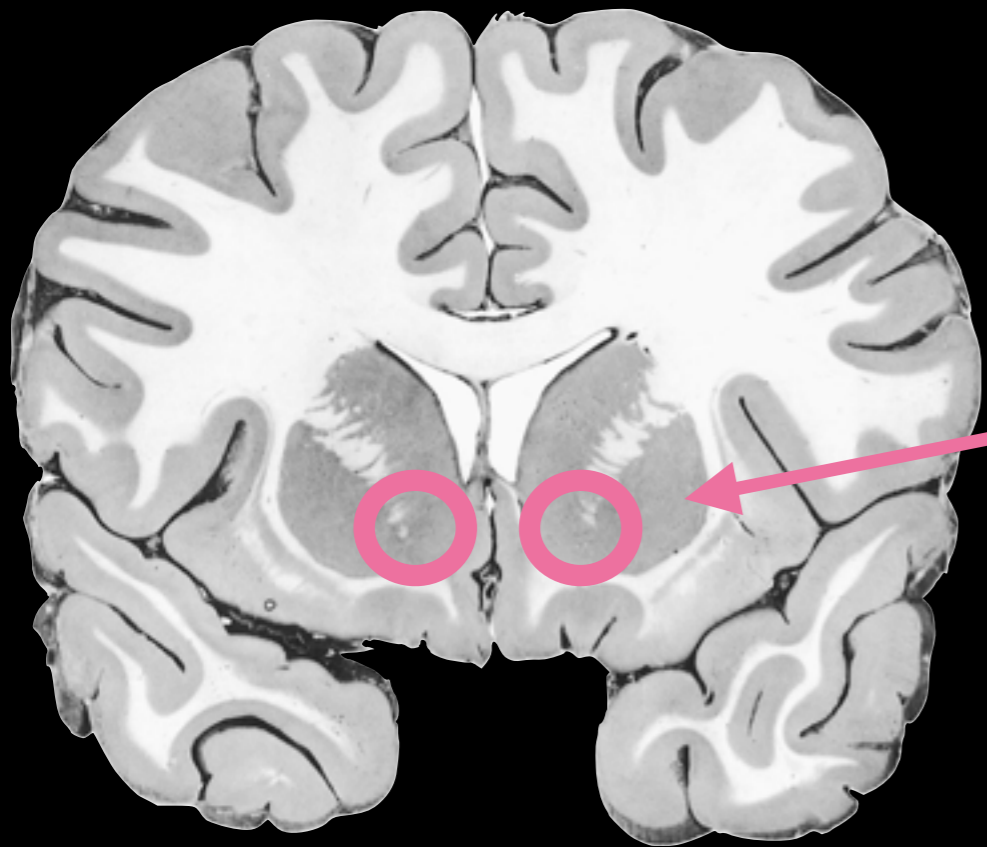


$$Q_{(s,a)} = Q_{(s,a)} + \alpha(R_s + \gamma Q_{(s',a')} - Q_{(s,a)})$$



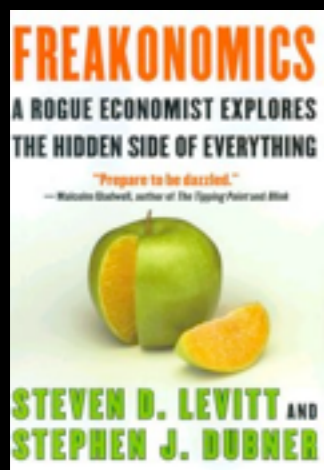
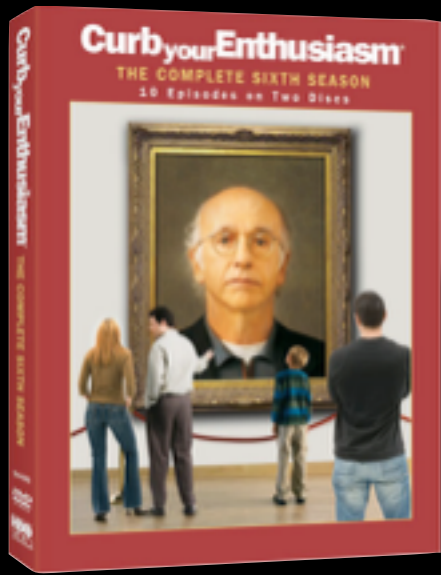
behavior

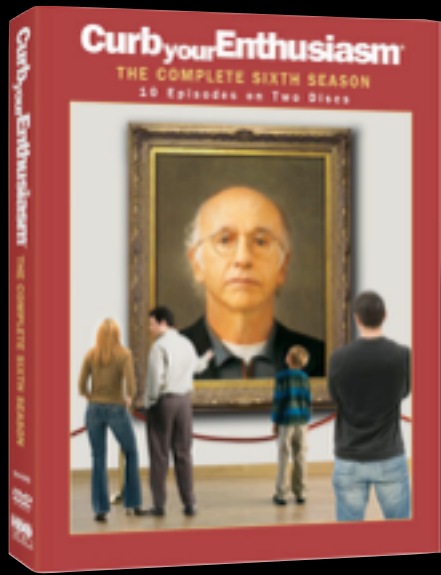
experience



$$Q_{(s,a)} = Q_{(s,a)} + \alpha(R_s + \gamma Q_{(s',a')} - Q_{(s,a)})$$

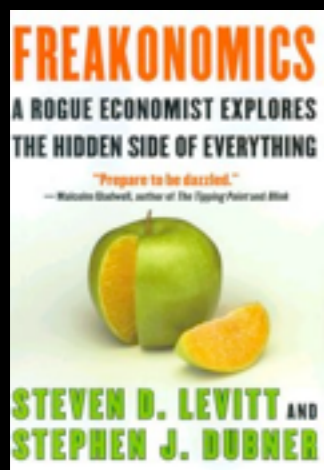
behavior





\$8

\$5



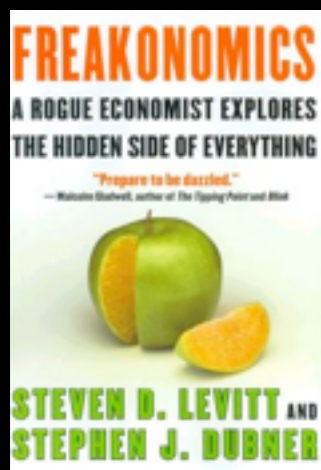
\$4



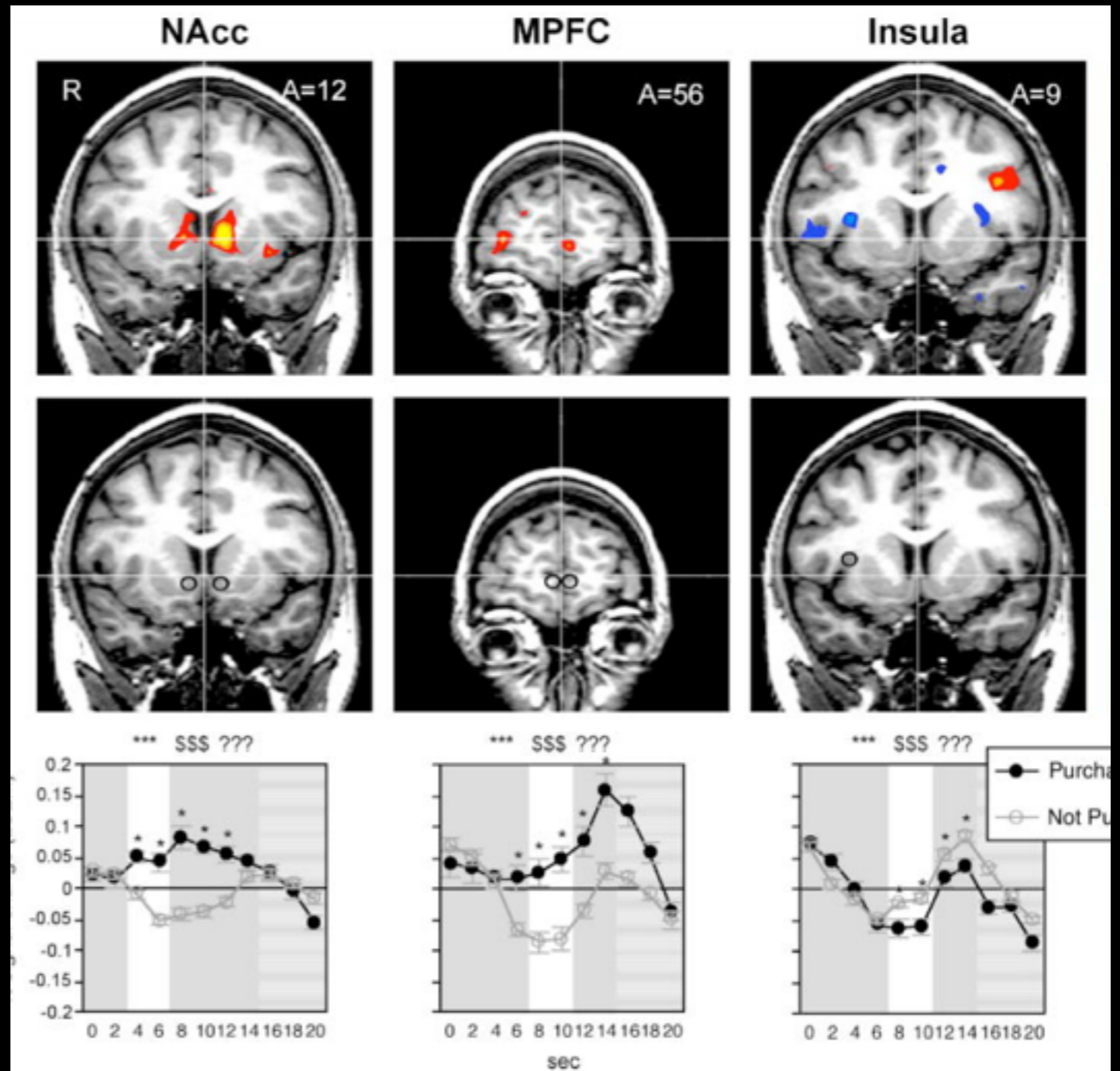
\$8



\$5



\$4



VMPFC

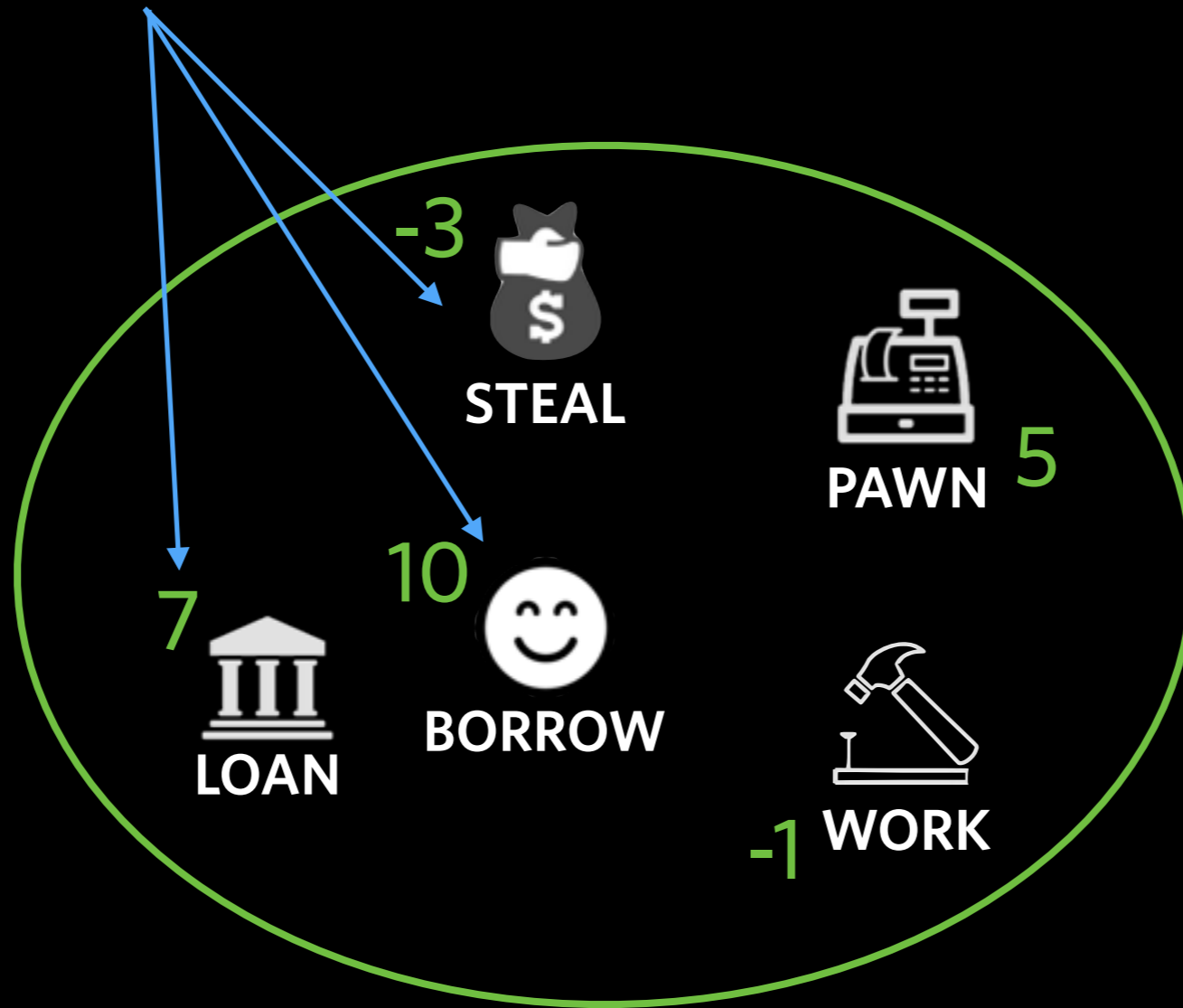


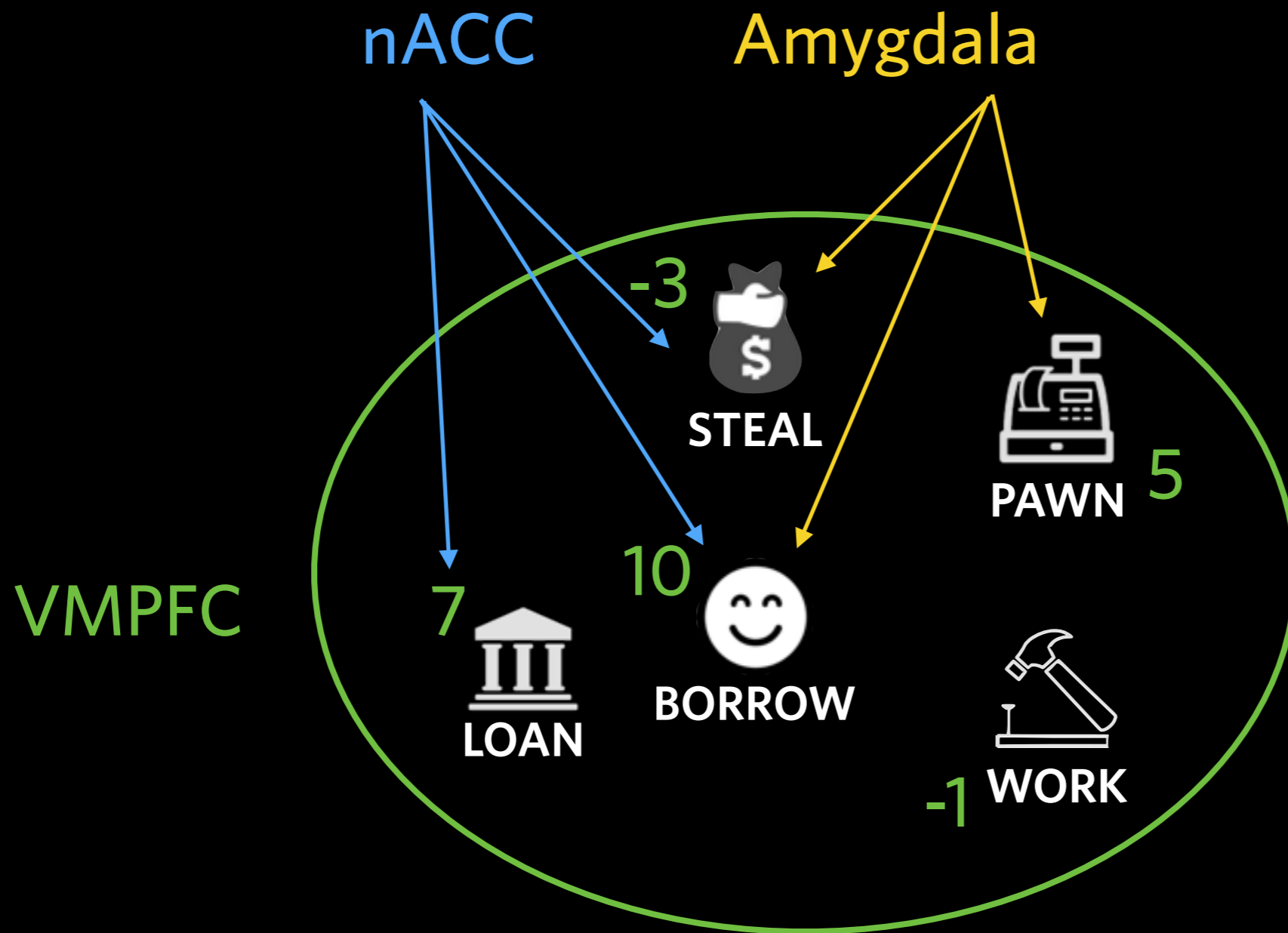
VMPFC

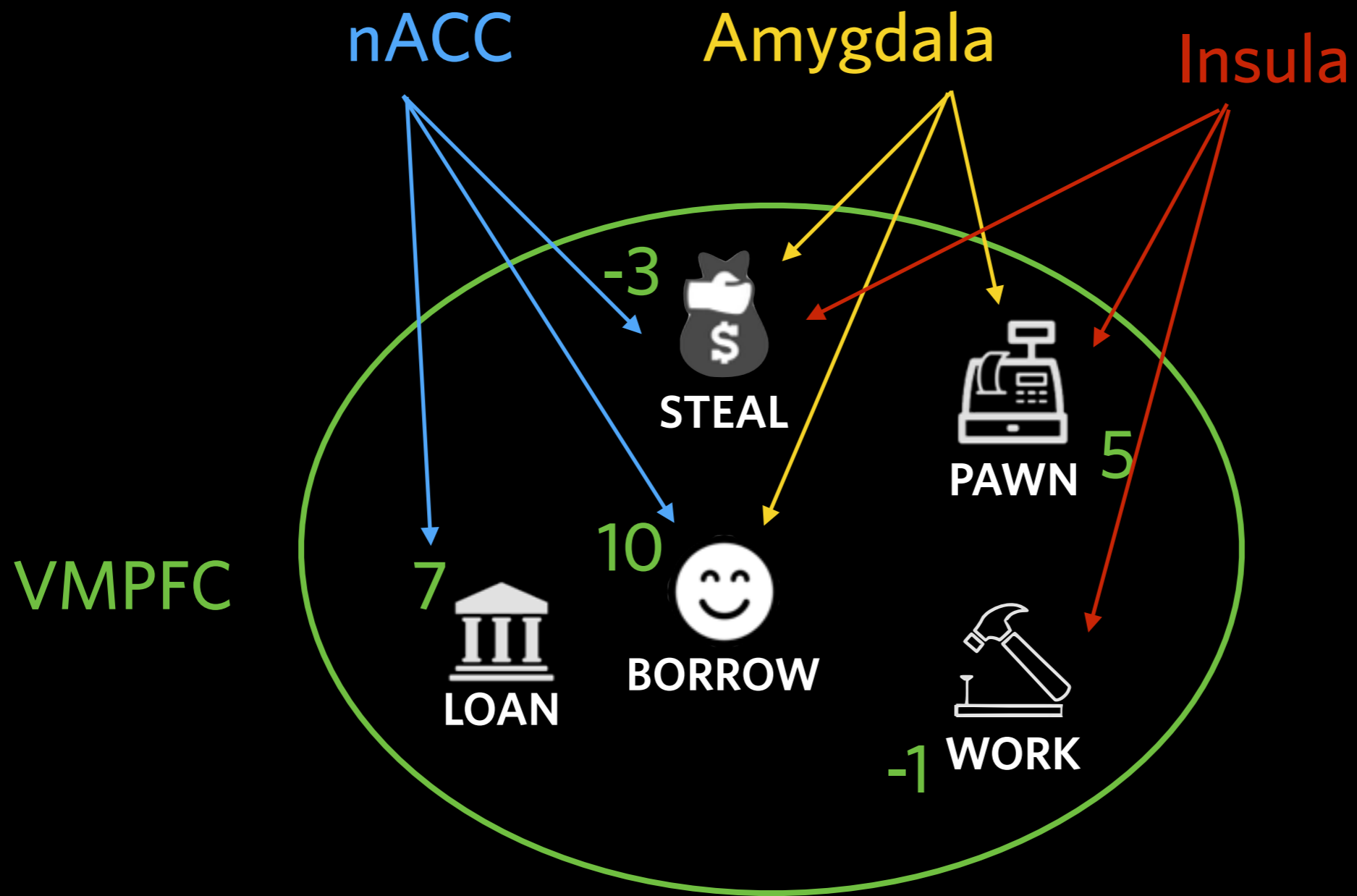


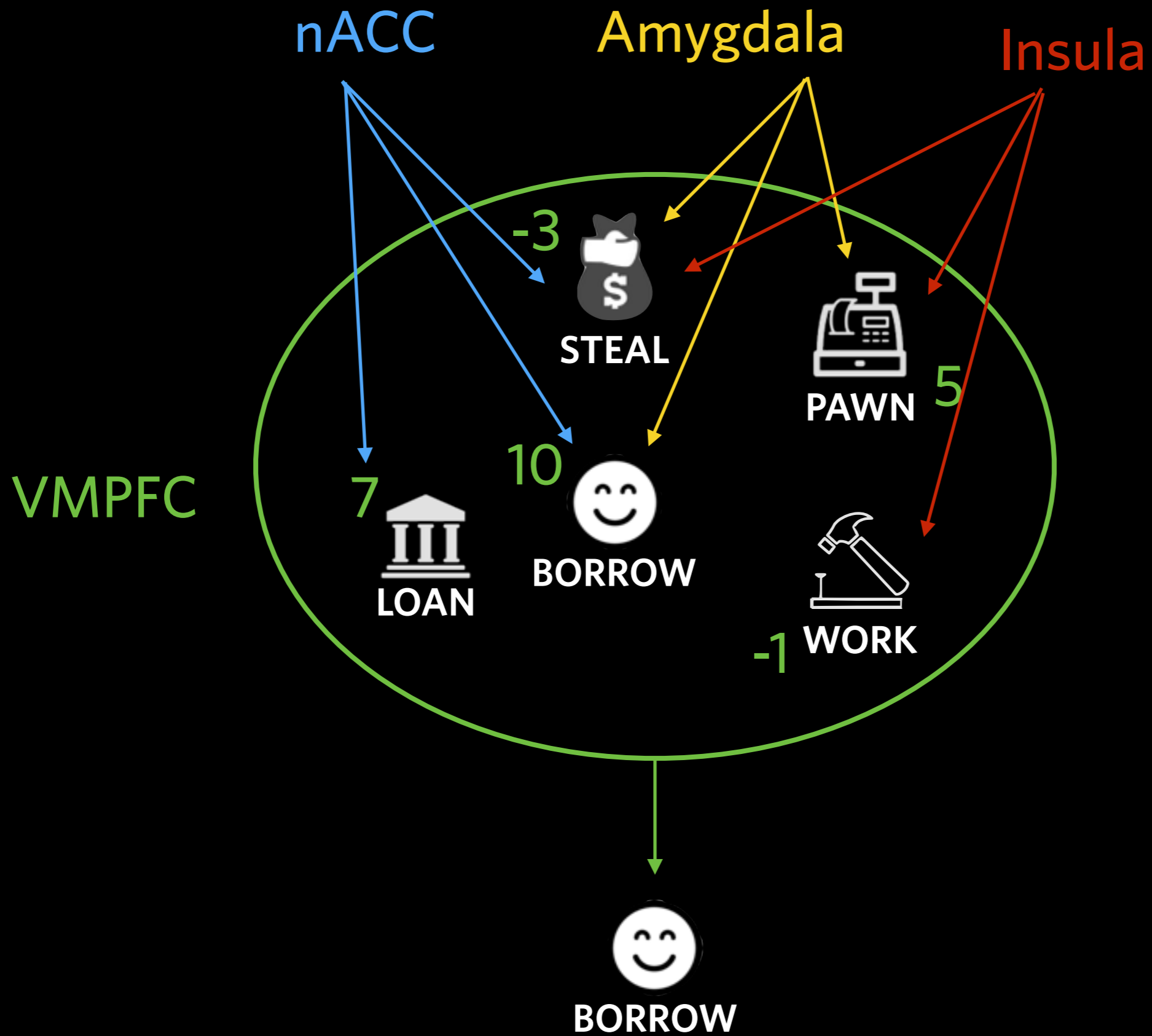
nACC

VMPFC



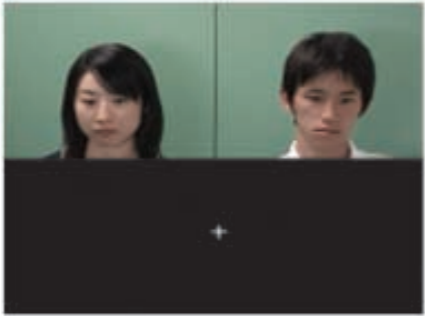
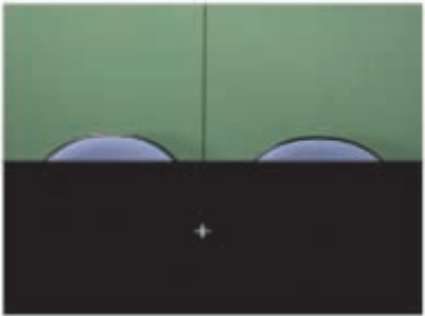






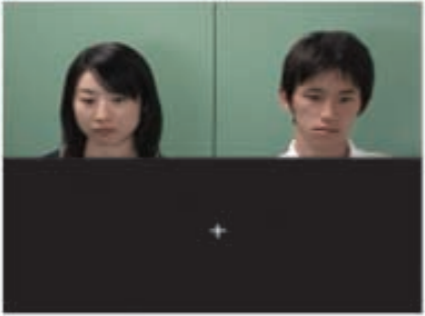
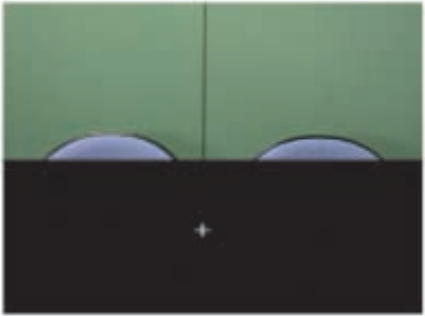
Application: Finding Happiness

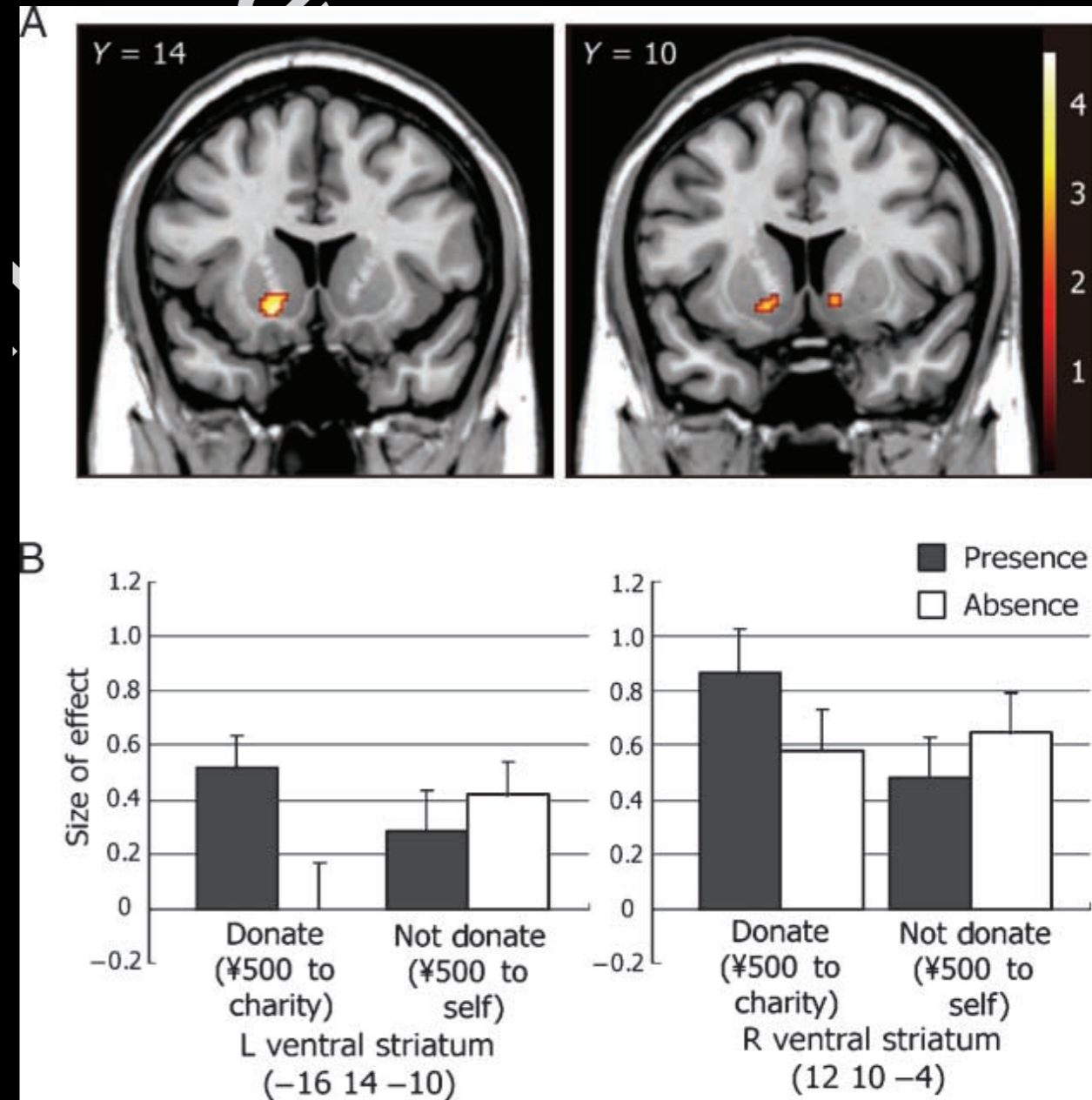
Izuma et al 2009

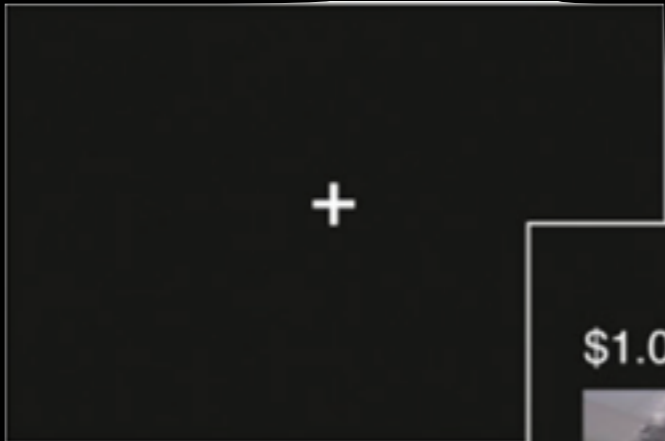
		Presence/absence of observers	
		Presence condition	Absence condition
Subject's choice	Donate (¥500 to charity)	 <ul style="list-style-type: none">• Extrinsic social reward (social approval)• Intrinsic benefit	 <ul style="list-style-type: none">• Intrinsic benefit
	Not donate (¥500 to self)	<ul style="list-style-type: none">• Monetary reward (¥500)• Extrinsic social cost (social disapproval)	<ul style="list-style-type: none">• Monetary reward (¥500)


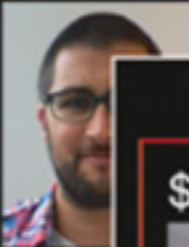
Application: Finding Happiness

Izuma et al 2009



	Presence/absence of observers	
	Presence condition	Absence condition
		
Donate (¥500 to charity)	<ul style="list-style-type: none"> • Extrinsic social reward (social approval) • Intrinsic benefit 	<ul style="list-style-type: none"> • Intrinsic benefit
Not donate (¥500 to self)	<ul style="list-style-type: none"> • Monetary reward (¥500) • Extrinsic social cost (social disapproval) 	<ul style="list-style-type: none"> • Monetary reward (¥500)

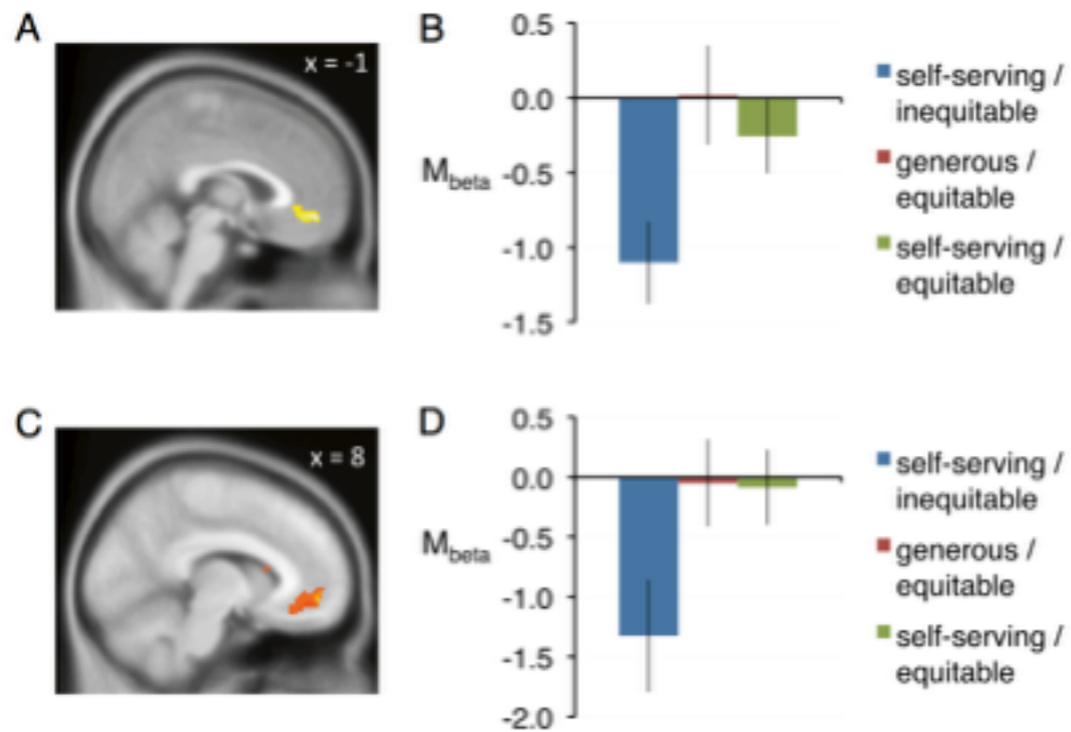
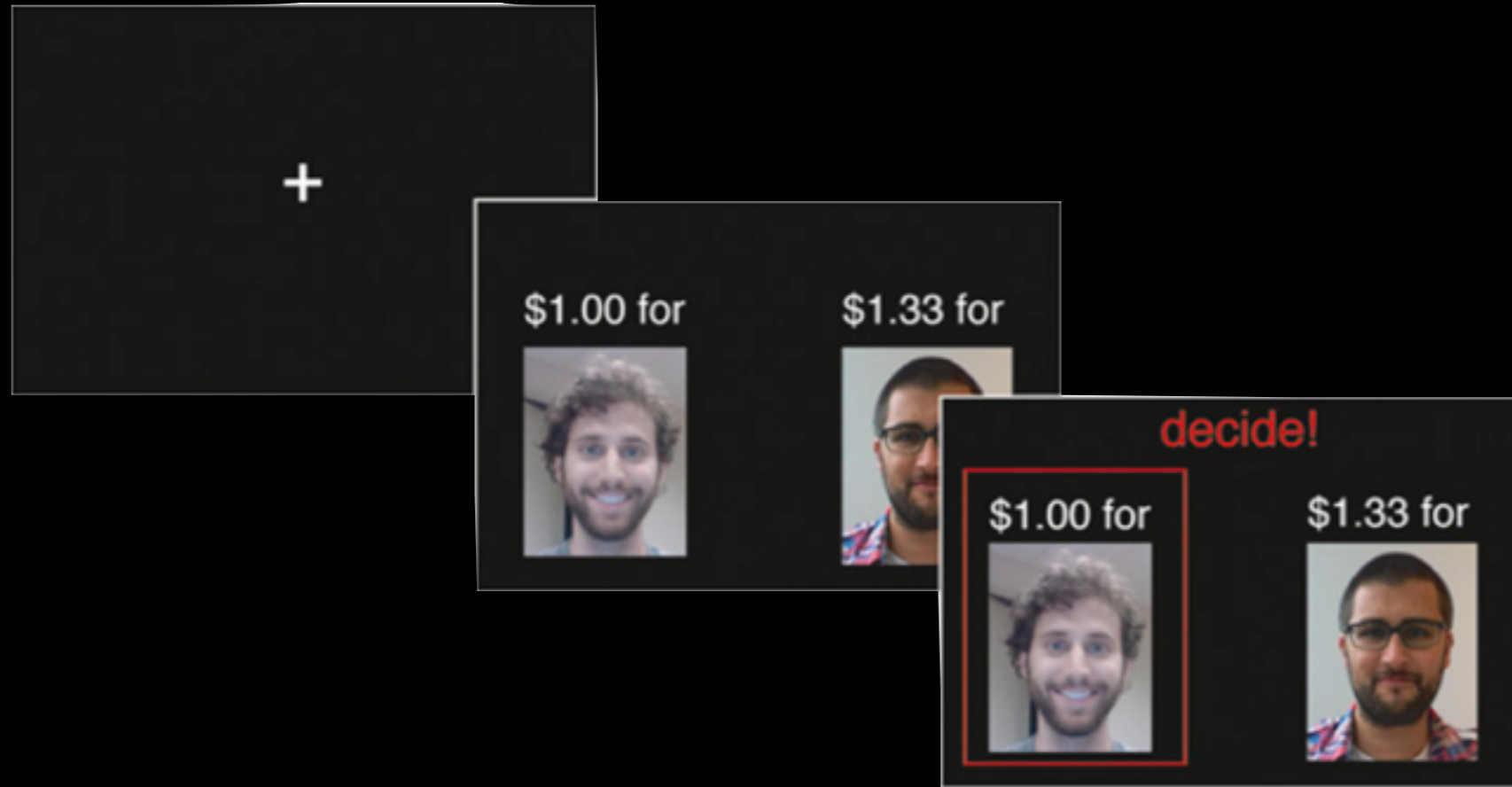


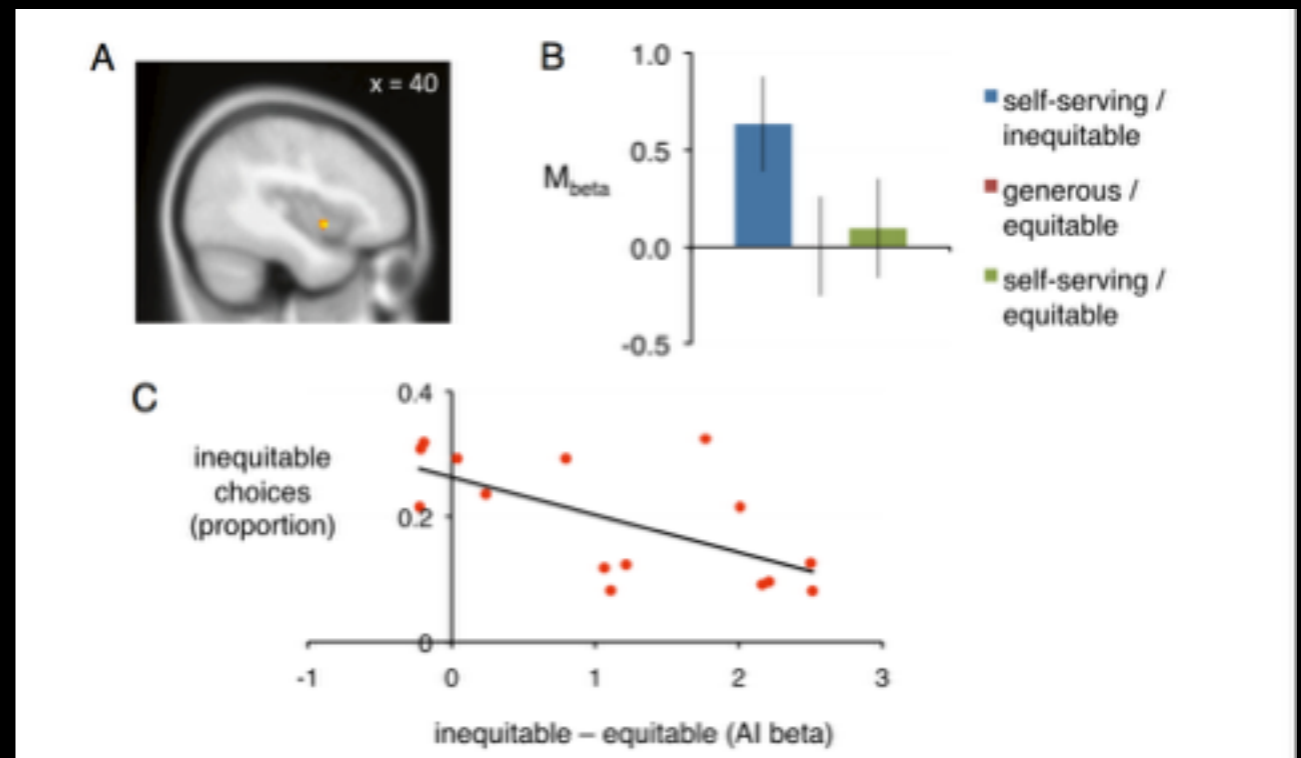
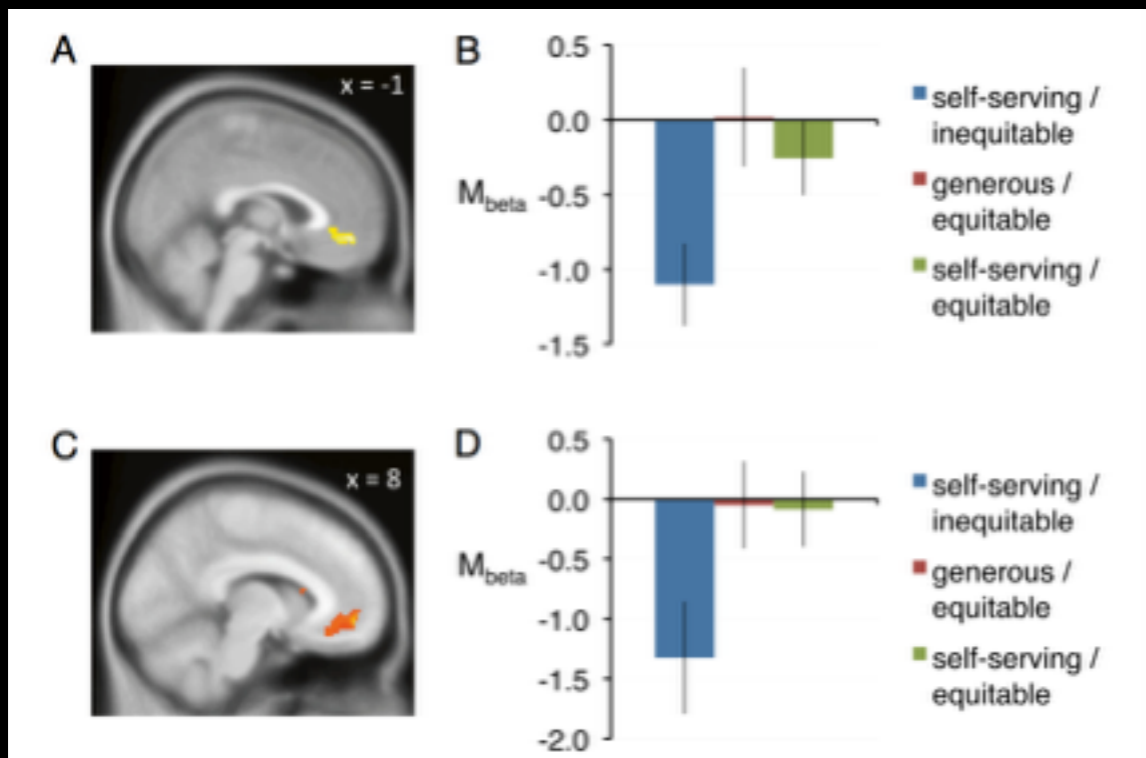
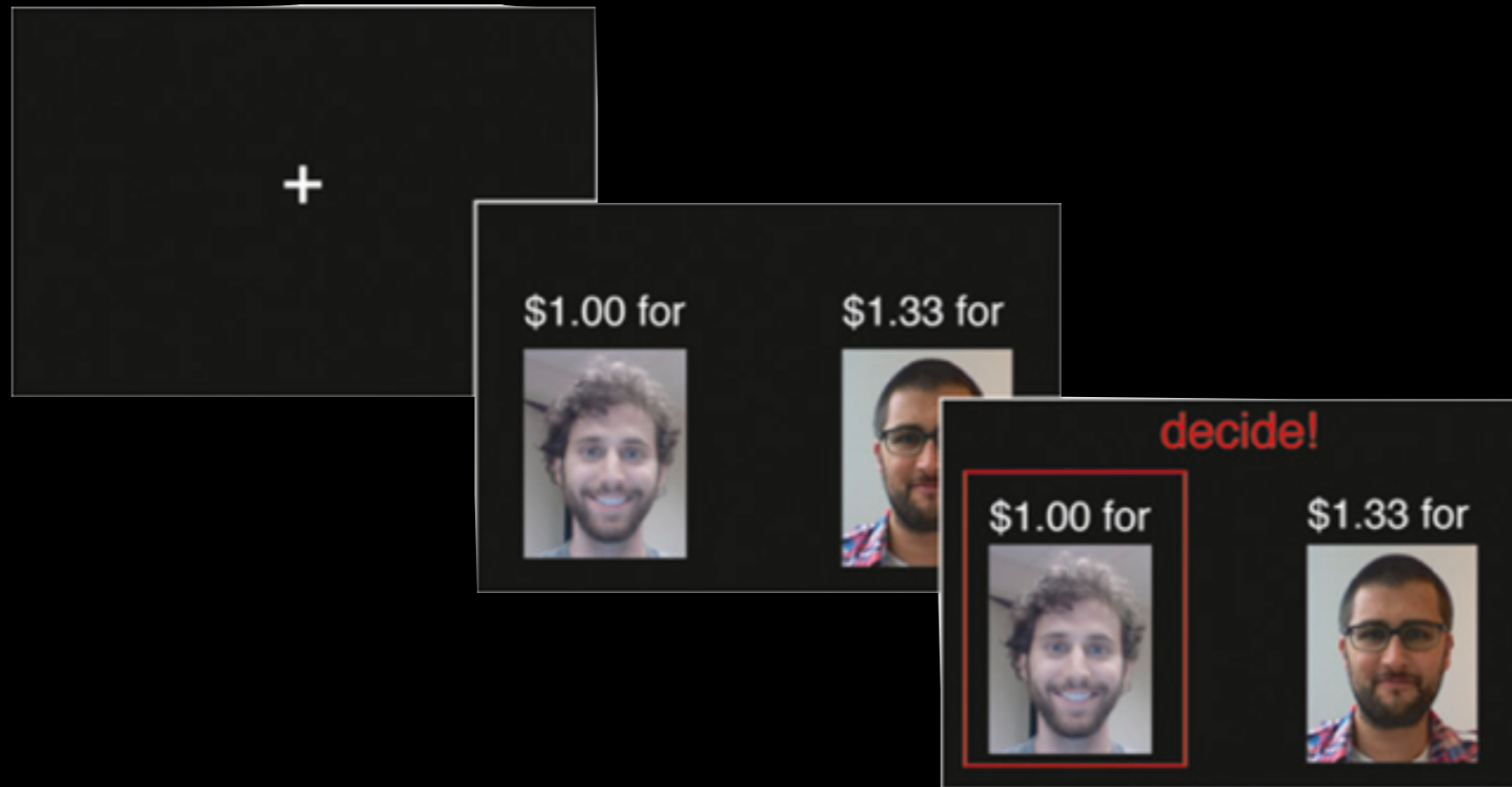


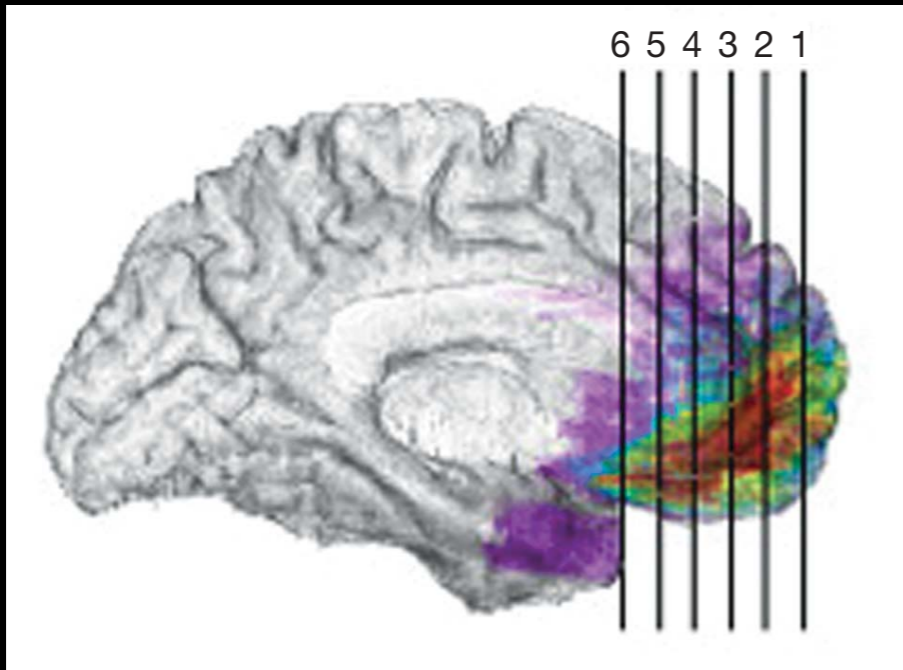
\$1.00 for  \$1.33 for 

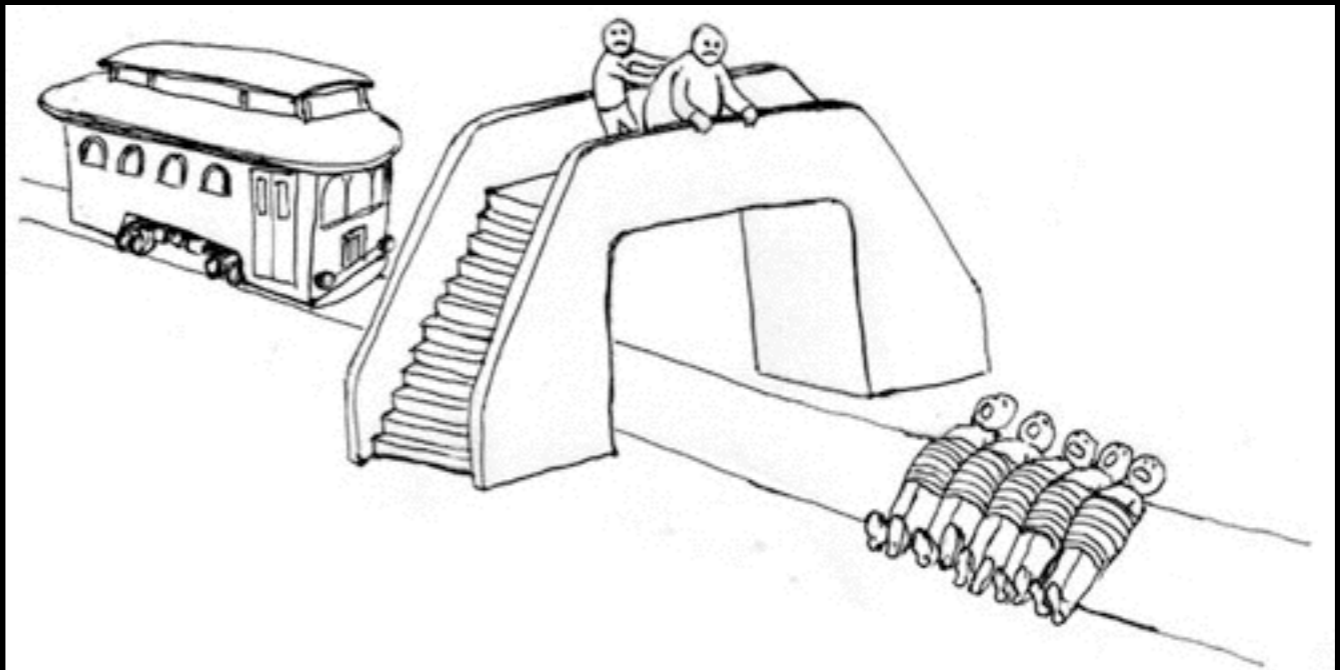
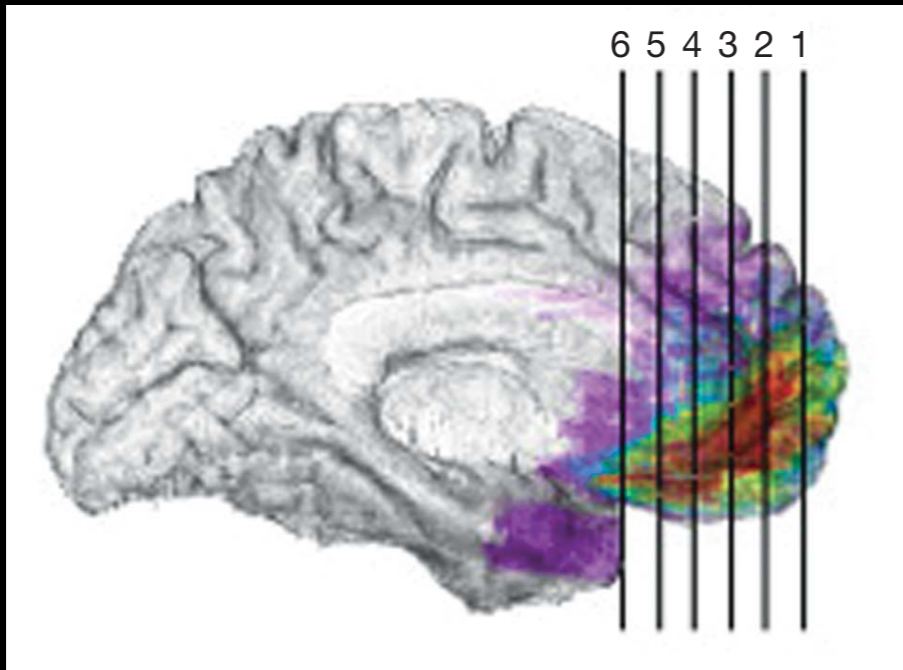
decide!

\$1.00 for  \$1.33 for 









Anderson et al. 1999:

Anderson et al. 1999:

Subject A: Run over by vehicle at age 15 months

Anderson et al. 1999:

Subject A: Run over by vehicle at age 15 months

“... stole from her family and from other children and shoplifted frequently, leading to multiple arrests.”

Anderson et al. 1999:

Subject A: Run over by vehicle at age 15 months

“... stole from her family and from other children and shoplifted frequently, leading to multiple arrests.”

“... verbally and physically abusive to others.”

Anderson et al. 1999:

Subject A: Run over by vehicle at age 15 months

“... stole from her family and from other children and shoplifted frequently, leading to multiple arrests.”

“... verbally and physically abusive to others.”

“... lied chronically.”

Anderson et al. 1999:

Subject A: Run over by vehicle at age 15 months

“... stole from her family and from other children and shoplifted frequently, leading to multiple arrests.”

“... verbally and physically abusive to others.”

“... lied chronically.”

“... little or no evidence that she experienced empathy, and her maternal behavior was marked by dangerous insensitivity to the infant’s needs”

Anderson et al. 1999:

Anderson et al. 1999:

“... largely unresponsive to verbal or physical punishment...”

Anderson et al. 1999:

“... largely unresponsive to verbal or physical punishment...”

“... early and risky sexual behavior leading to pregnancy at age 18”

Anderson et al. 1999:

“... largely unresponsive to verbal or physical punishment...”

“... early and risky sexual behavior leading to pregnancy at age 18”

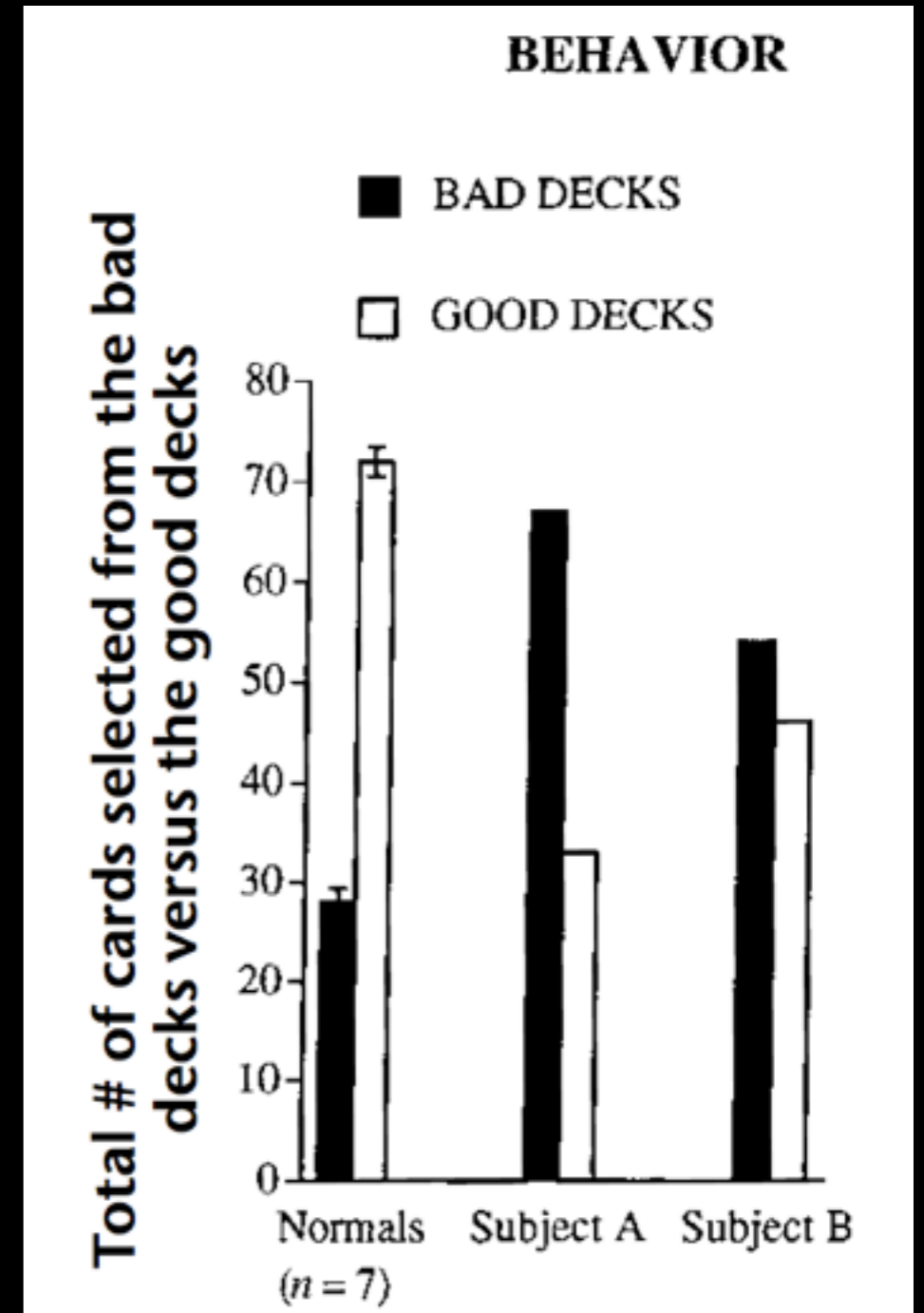
“... repeatedly [put] herself at physical and financial risk ...”

Anderson et al. 1999:

“... largely unresponsive to verbal or physical punishment...”

“... early and risky sexual behavior leading to pregnancy at age 18”

“... repeatedly [put] herself at physical and financial risk ...”

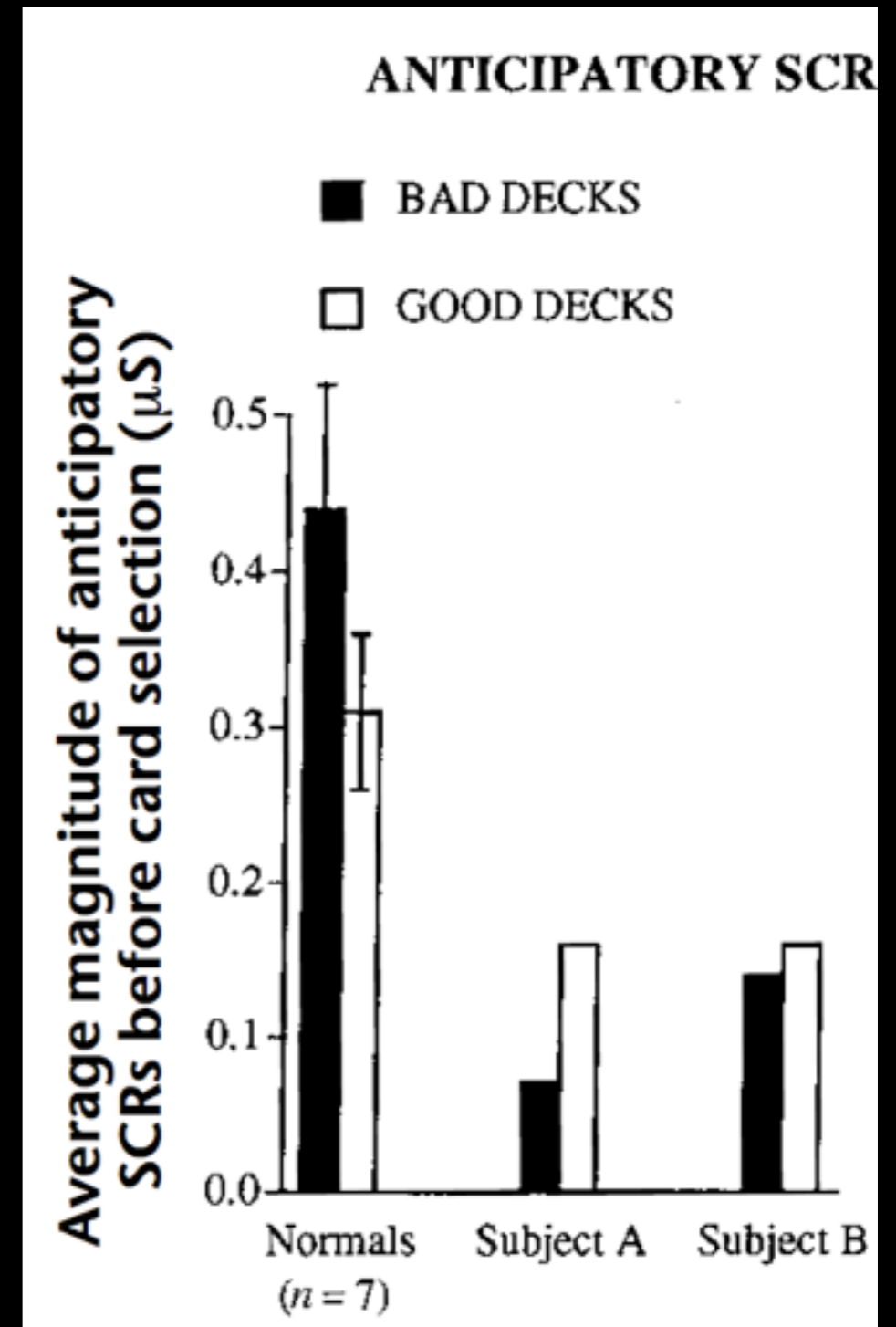


Anderson et al. 1999:

“... largely unresponsive to verbal or physical punishment...”

“... early and risky sexual behavior leading to pregnancy at age 18”

“... repeatedly [put] herself at physical and financial risk ...”



The neurobiology of rewards and values in social decision making

Christian C. Ruff and Ernst Fehr

Abstract | How does our brain choose the best course of action? Choices between material goods are thought to be steered by neural value signals that encode the rewarding properties of the choice options. Social decisions, by contrast, are traditionally thought to rely on neural representations of the self and others. However, recent studies show that many types of social decisions may also involve neural value computations. This suggests a unified mechanism for motivational control of behaviour that may incorporate both social and non-social factors. In this Review, we outline a theoretical framework that may help to identify possible overlaps and differences between the neural processes that guide social and non-social decision making.

The neurobiology of rewards and values in social decision making

Christian C. Ruff and Ernst Fehr

Abstract | How does our brain choose the best course of action? Choices between material goods are thought to be steered by neural value signals that encode the rewarding properties of the choice options. Social decisions, by contrast, are traditionally thought to rely on neural representations of the self and others. However, recent studies show that many types of social decisions may also involve neural value computations. This suggests a unified

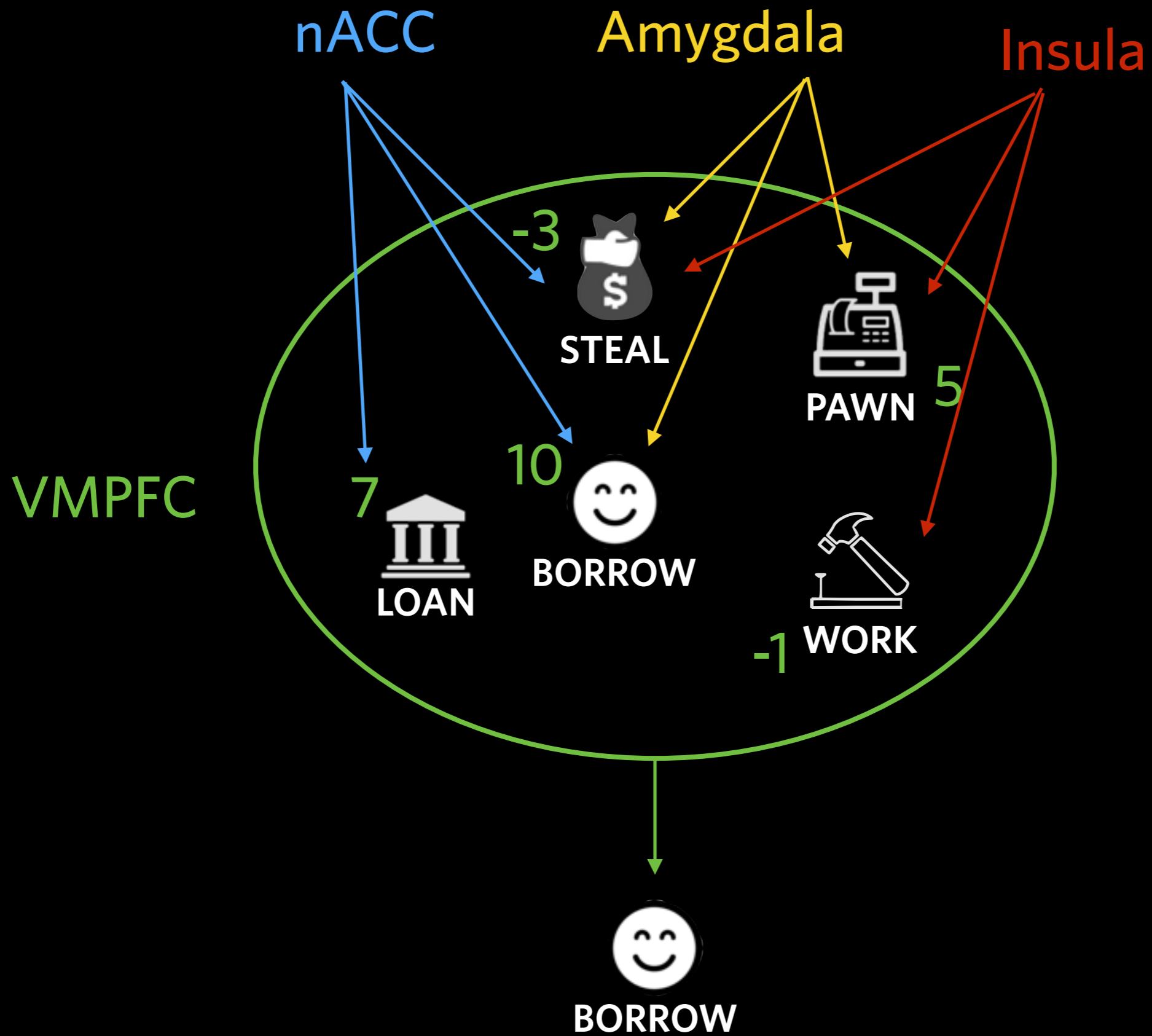
mechanism for motivation and non-social factors. In this review, we identify possible overlaps between social and non-social decision making.

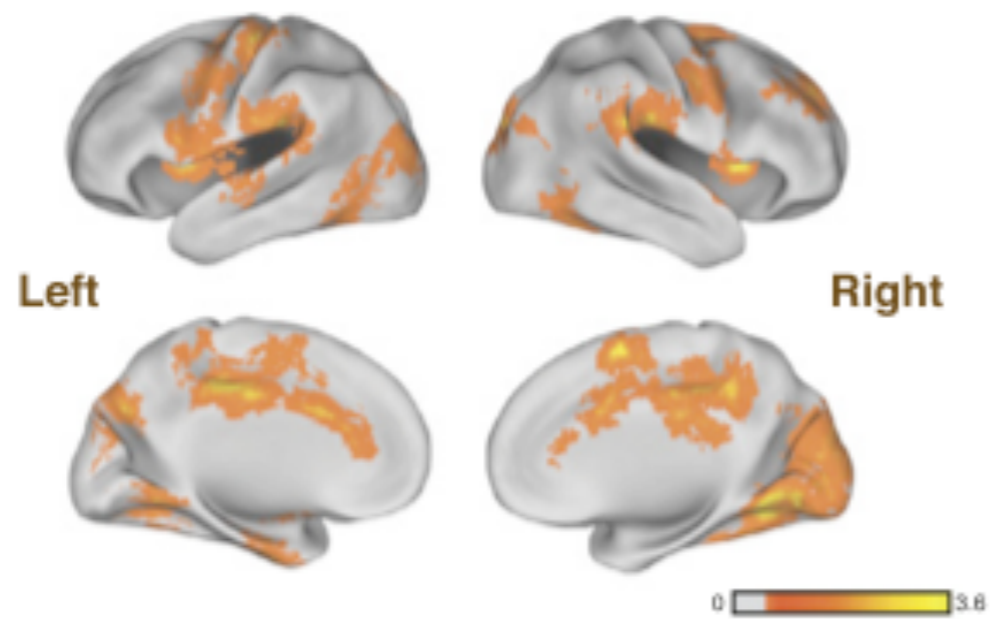
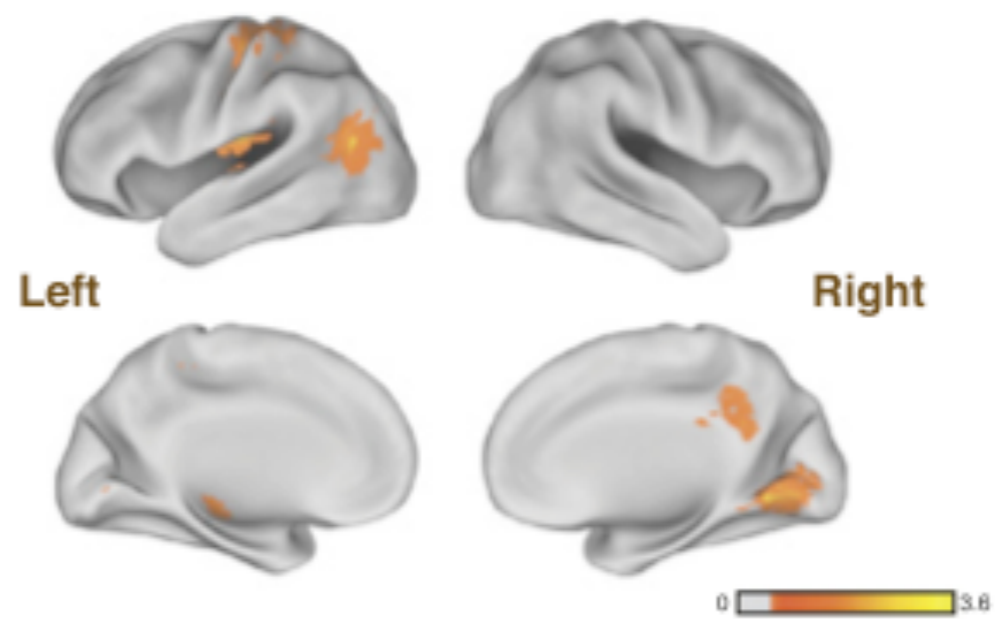
Where in the brain is morality? Everywhere and maybe Nowhere

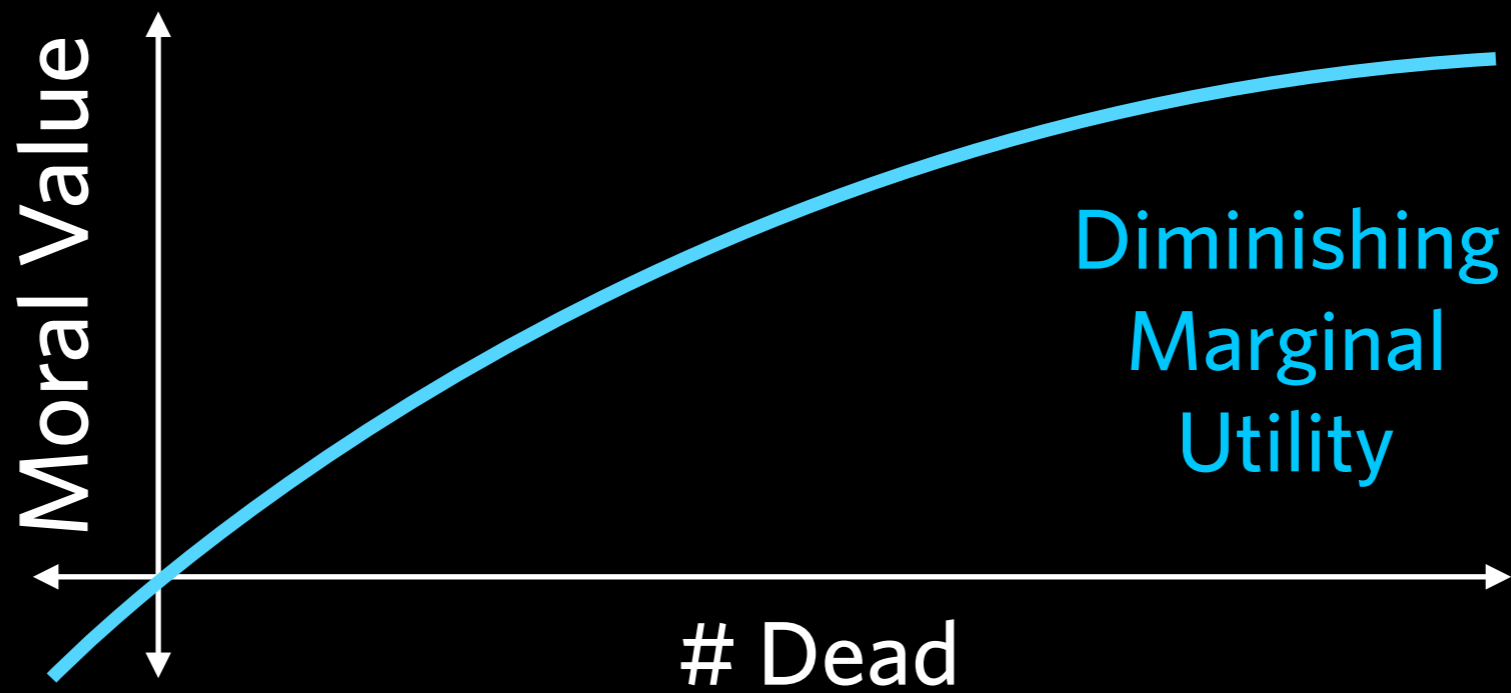
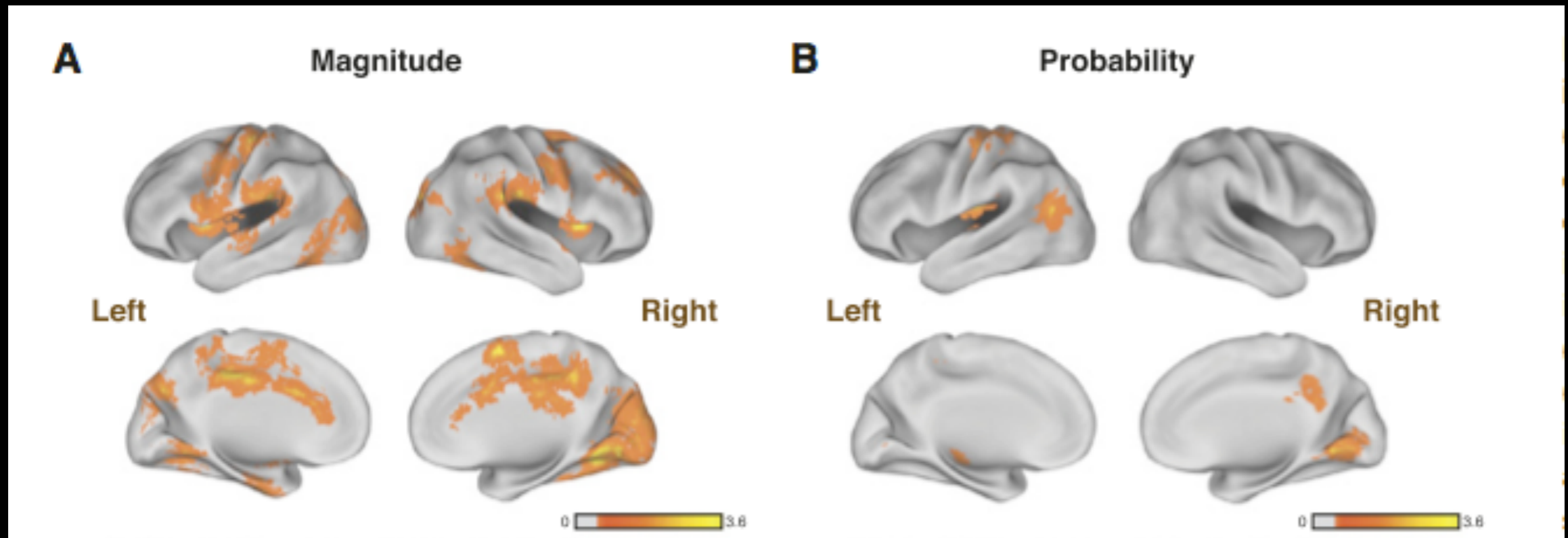
Liane Young and James Dungan

Department of Psychology, Boston College, Chestnut Hill, MA, USA

The neuroscience of morality has focused on how morality works and where it is in the brain. In tackling these questions, researchers have taken both domain-specific and domain-general approaches—searching for neural substrates and systems dedicated to moral cognition versus characterizing the contributions of domain-general processes. Where in the brain is morality? On one hand, morality is made up of complex cognitive processes, deployed across many domains and housed all over the brain. On the other hand, no neural substrate or system that uniquely supports moral cognition has been found. In this review, we will discuss early assumptions of domain-specificity in moral neuroscience as well as subsequent investigations of domain-general contributions, taking emotion and social cognition (i.e., theory of mind) as case studies. Finally, we will consider possible cognitive accounts of a domain-specific morality: Does uniquely moral cognition exist?

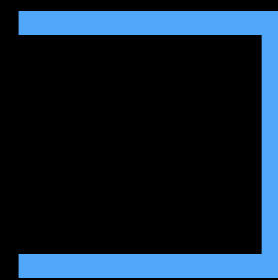


A**Magnitude****B****Probability**



Neuroscience

Psychology



$\$ = \overset{\circ}{M}$

$\$ \neq \overset{\circ}{M}$

Moral vs. Conventional

Turiel, Nucci & others

Moral vs. Conventional

Turiel, Nucci & others

Moral Violations

Conventional Violations

Moral vs. Conventional

Turiel, Nucci & others

Moral Violations

Hitting a classmate

Conventional Violations

Boy wearing dress to school

Moral vs. Conventional

Turiel, Nucci & others

Moral Violations

Hitting a classmate

Very bad

Conventional Violations

Boy wearing dress to school

Moderately bad

Moral vs. Conventional

Turiel, Nucci & others

Moral Violations

Hitting a classmate

Very bad

Must be punished

Conventional Violations

Boy wearing dress to school

Moderately bad

Might be punished

Moral vs. Conventional

Turiel, Nucci & others

Moral Violations

Hitting a classmate

Very bad

Must be punished

Universal

Conventional Violations

Boy wearing dress to school

Moderately bad

Might be punished

Parochial

Moral vs. Conventional

Turiel, Nucci & others

Moral Violations

Hitting a classmate

Very bad

Must be punished

Universal

Authority-Independent

Conventional Violations

Boy wearing dress to school

Moderately bad

Might be punished

Parochial

Authority-dependent

Moral vs. Conventional

Turiel, Nucci & others

Moral Violations

Hitting a classmate

Very bad

Must be punished

Universal

Authority-Independent

Conventional Violations

Boy wearing dress to school

Moderately bad

Might be punished

Parochial

Authority-dependent



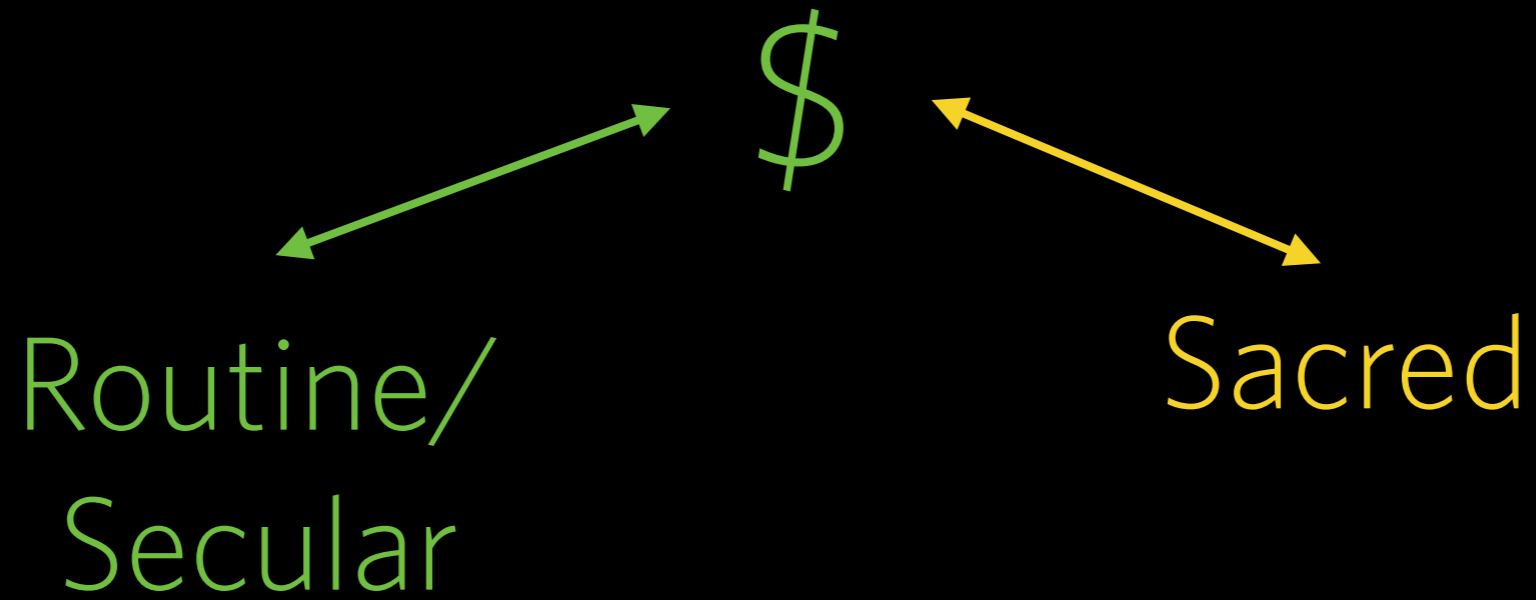
Tetlock, Kristel, Elson, Green & Lerner 2003



Sacred Values

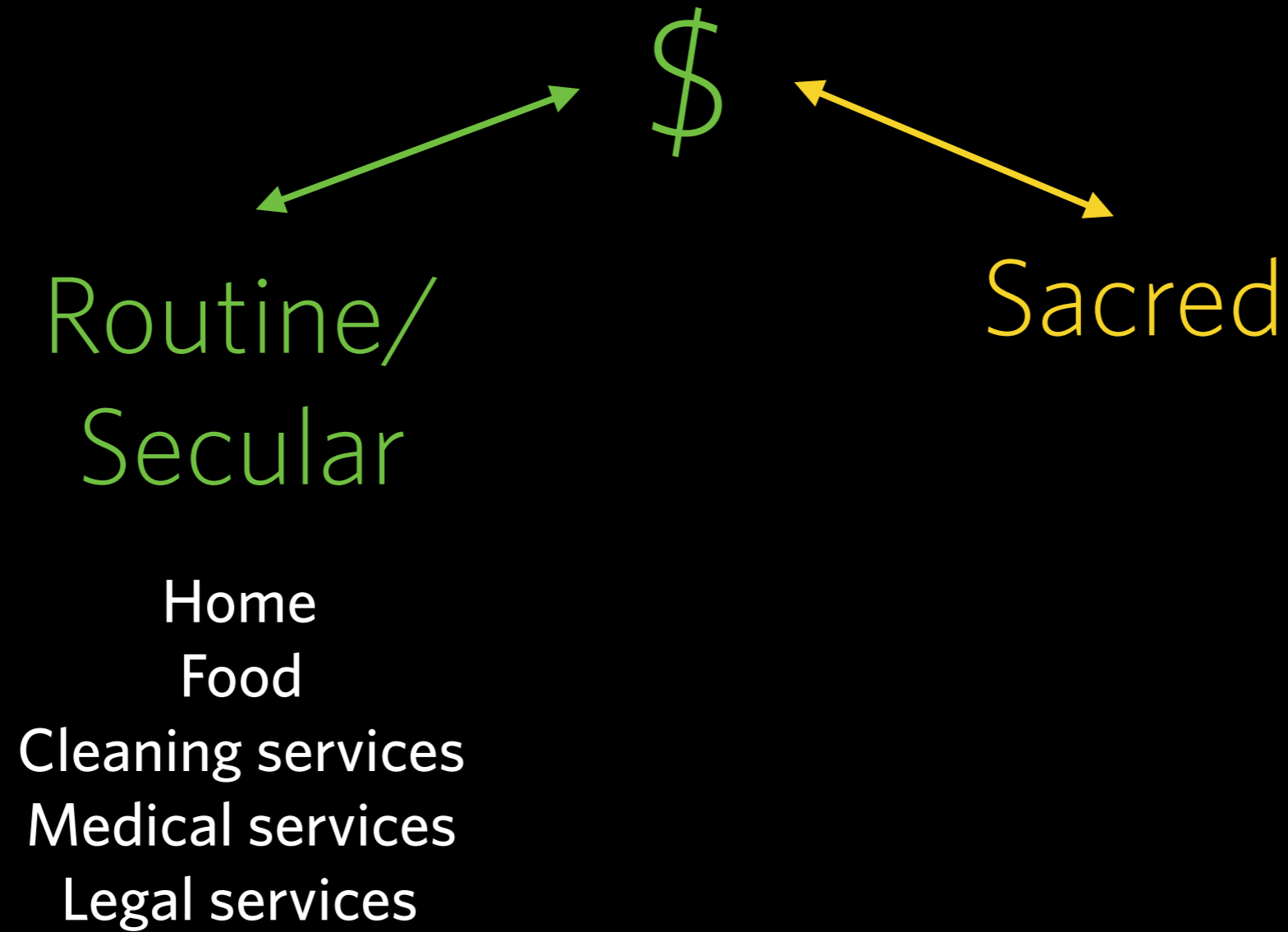


Sacred Values



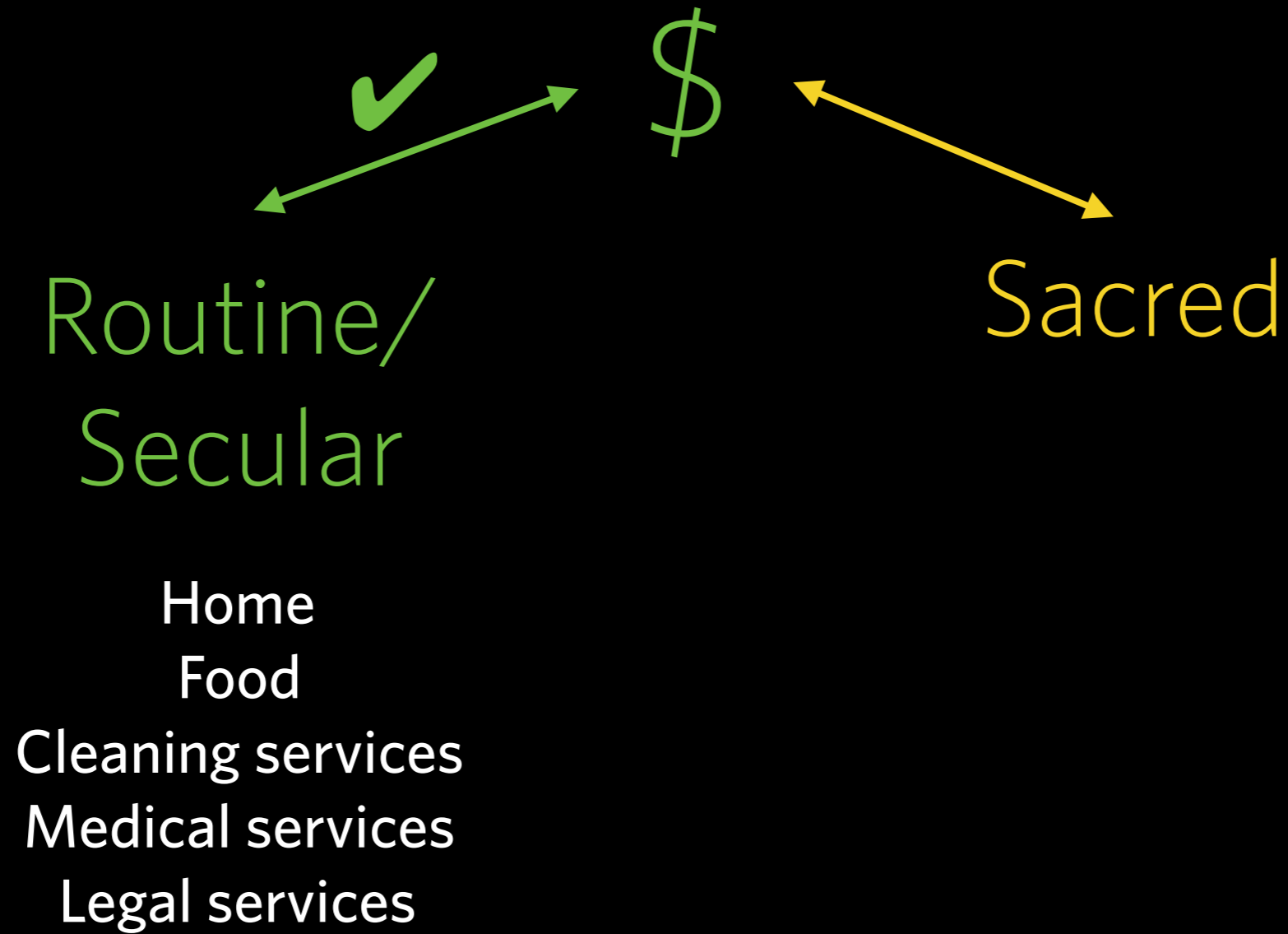


Sacred Values



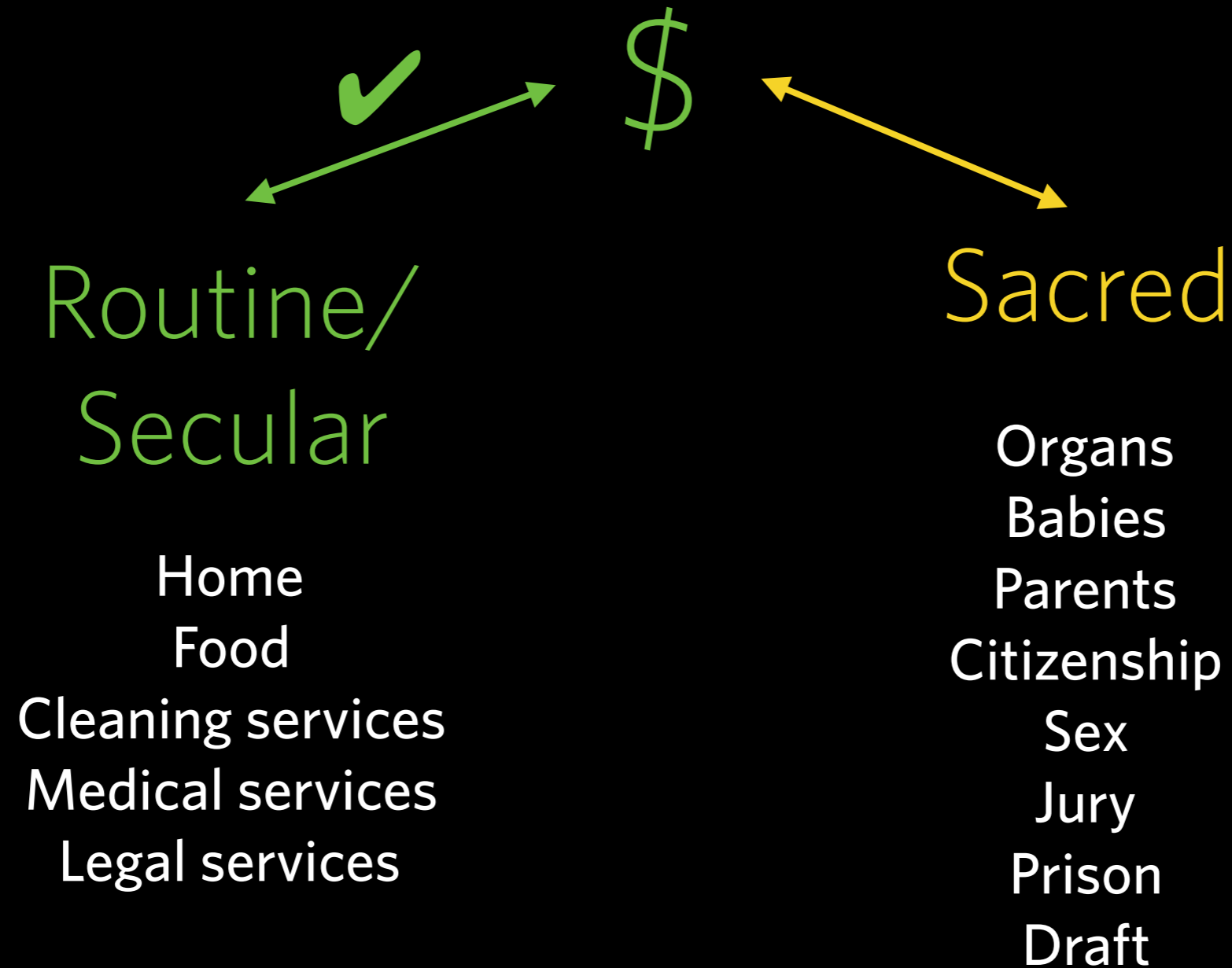


Sacred Values



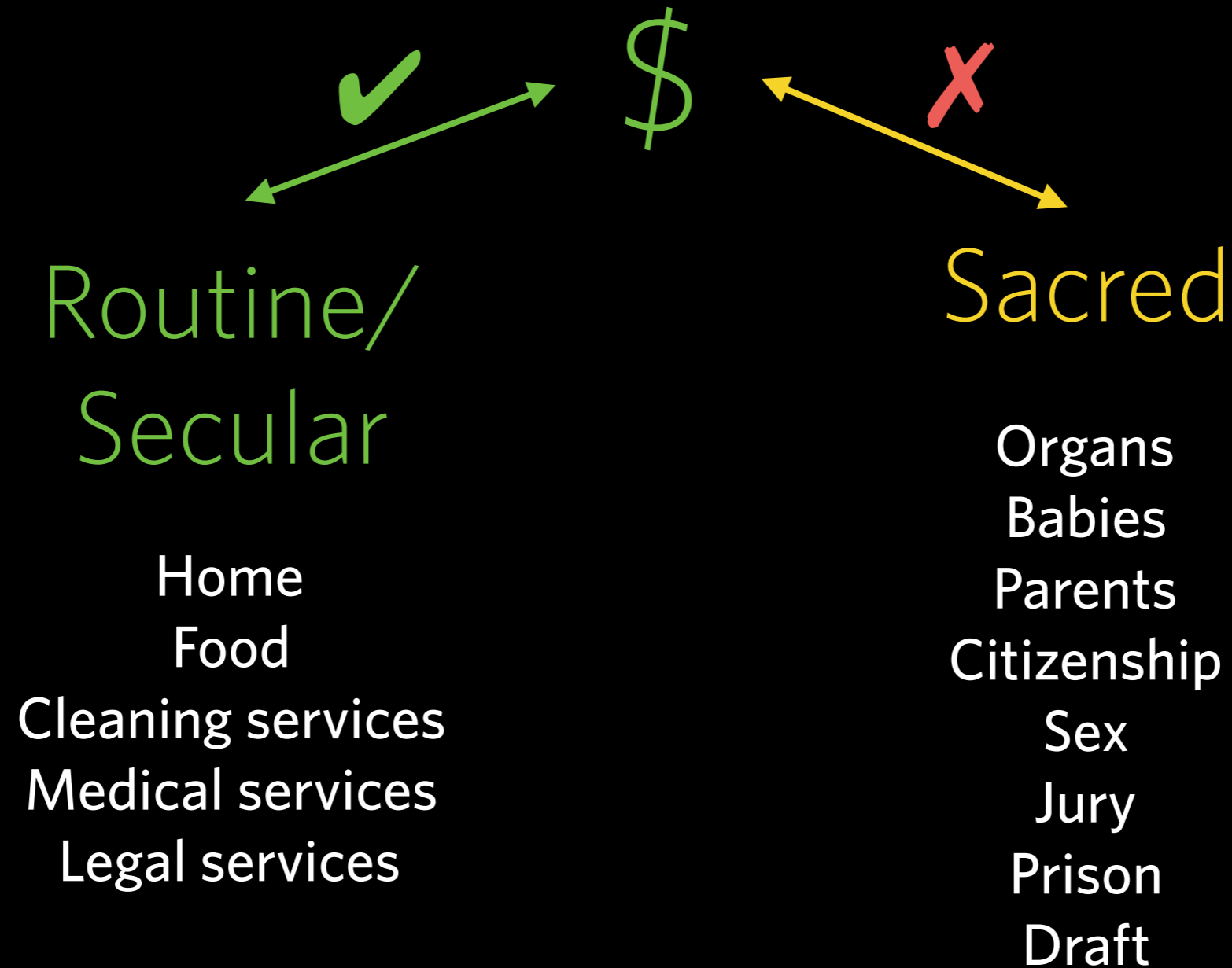


Sacred Values



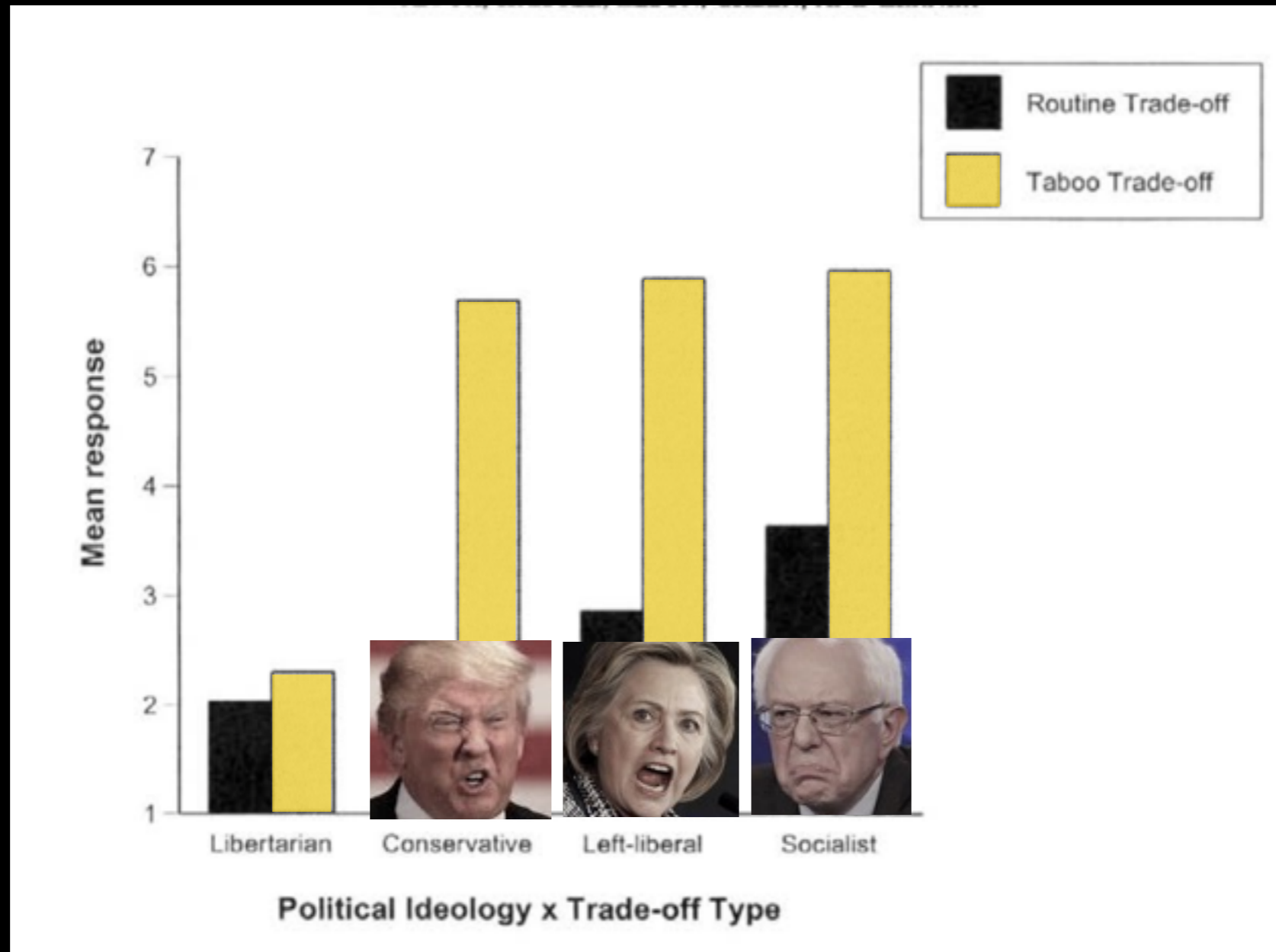


Sacred Values



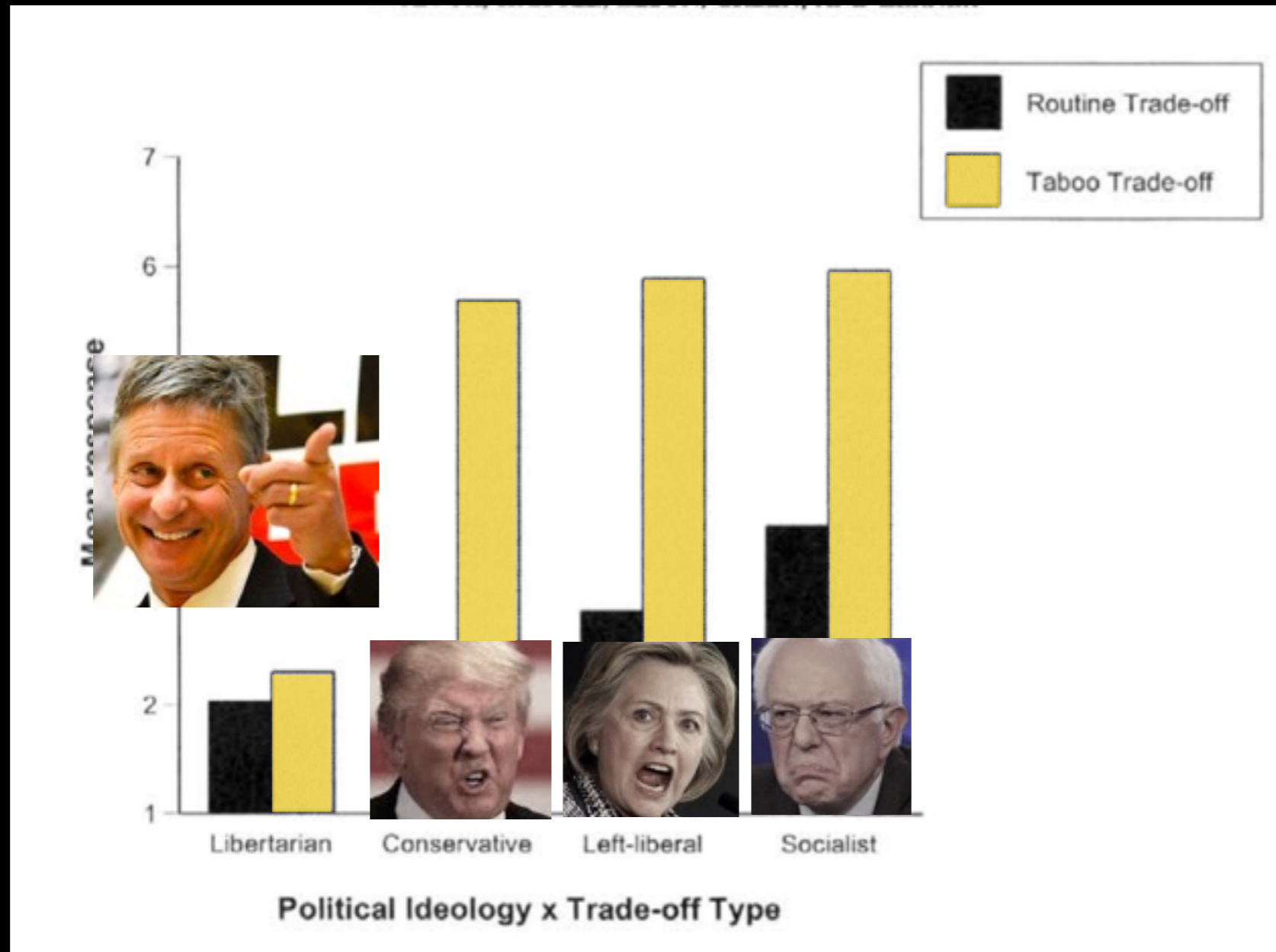


Sacred Values





Sacred Values



Couple billed \$39.35 to hold newborn son

Linda Hervieux, Newser staff

9:43 a.m. EDT October 5, 2016



(NEWSER) – A Utah couple got a surprise when they got the bill for the birth of their son, Samuel.

Tucked below the charge for "Delivery C Section" was a \$39.35 fee for "skin to skin," their fee for getting to hold the baby after delivery, explains [Vox](#).

In a post on [Reddit](#) that has gone viral, user "halfthrottle" writes: "During the C-section the nurse asked my wife if she would like to do skin to skin after the baby was born. Which of course anyone would say yes (to). We just noticed it in the bill today and had a laugh."

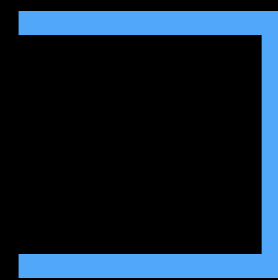


“If someone really thinks, in advance, that it is open to question whether such an action as procuring the judicial execution of the innocent should be quite excluded from consideration -- I do not want to argue with him; he shows a corrupt mind.”

- *Elizabeth Anscombe*

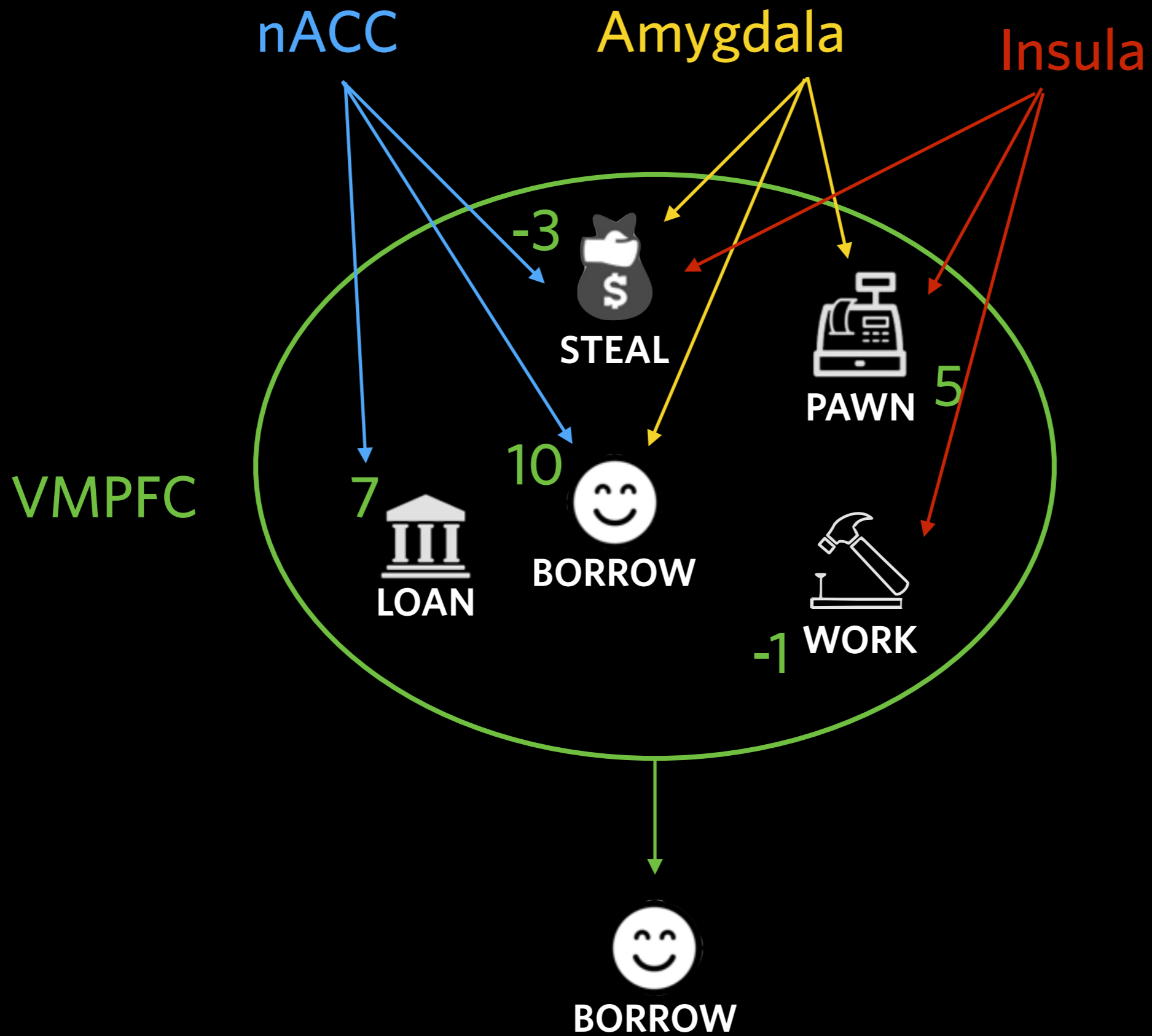
Neuroscience

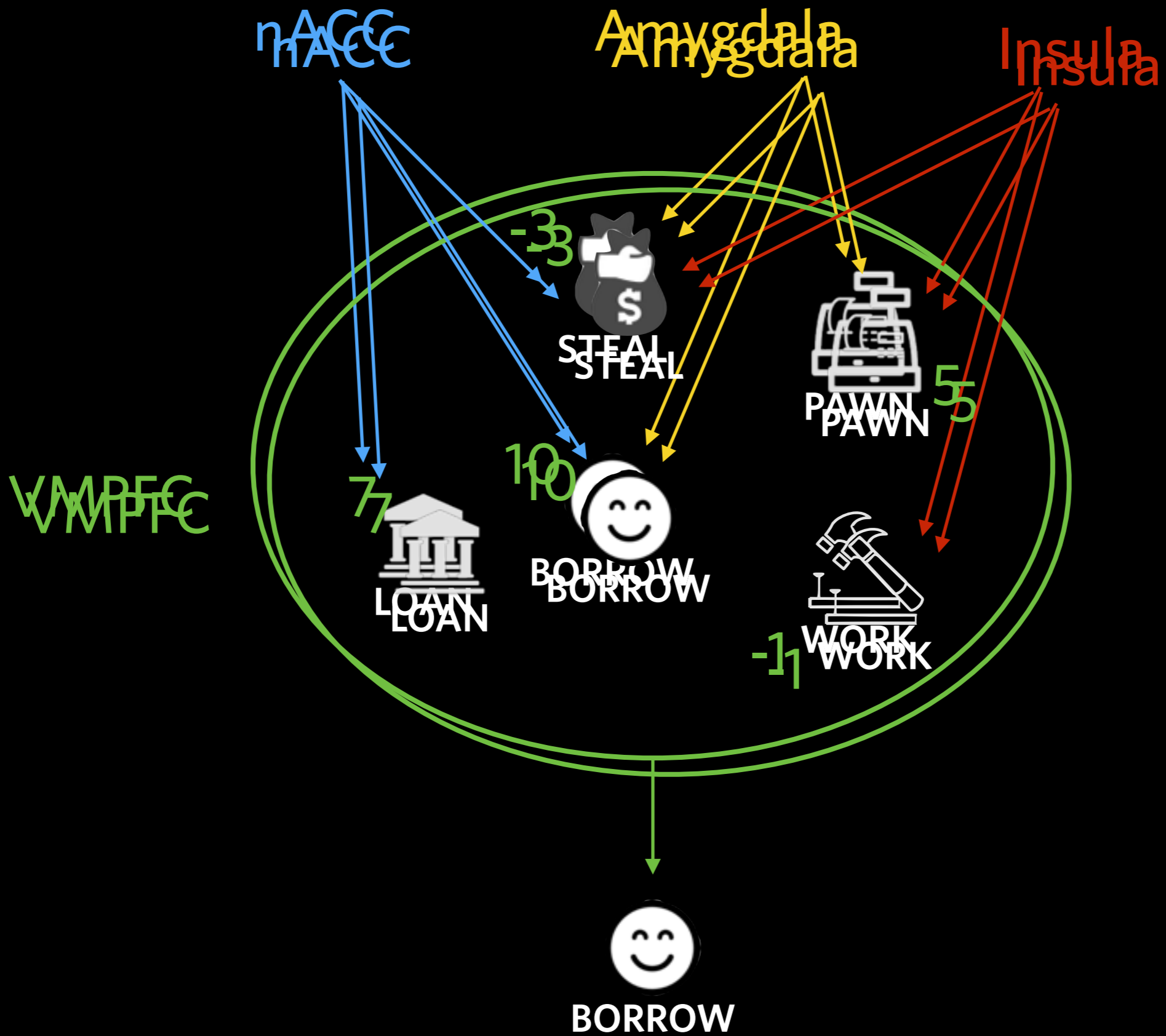
Psychology



$\$ = \overset{\circ}{M}$

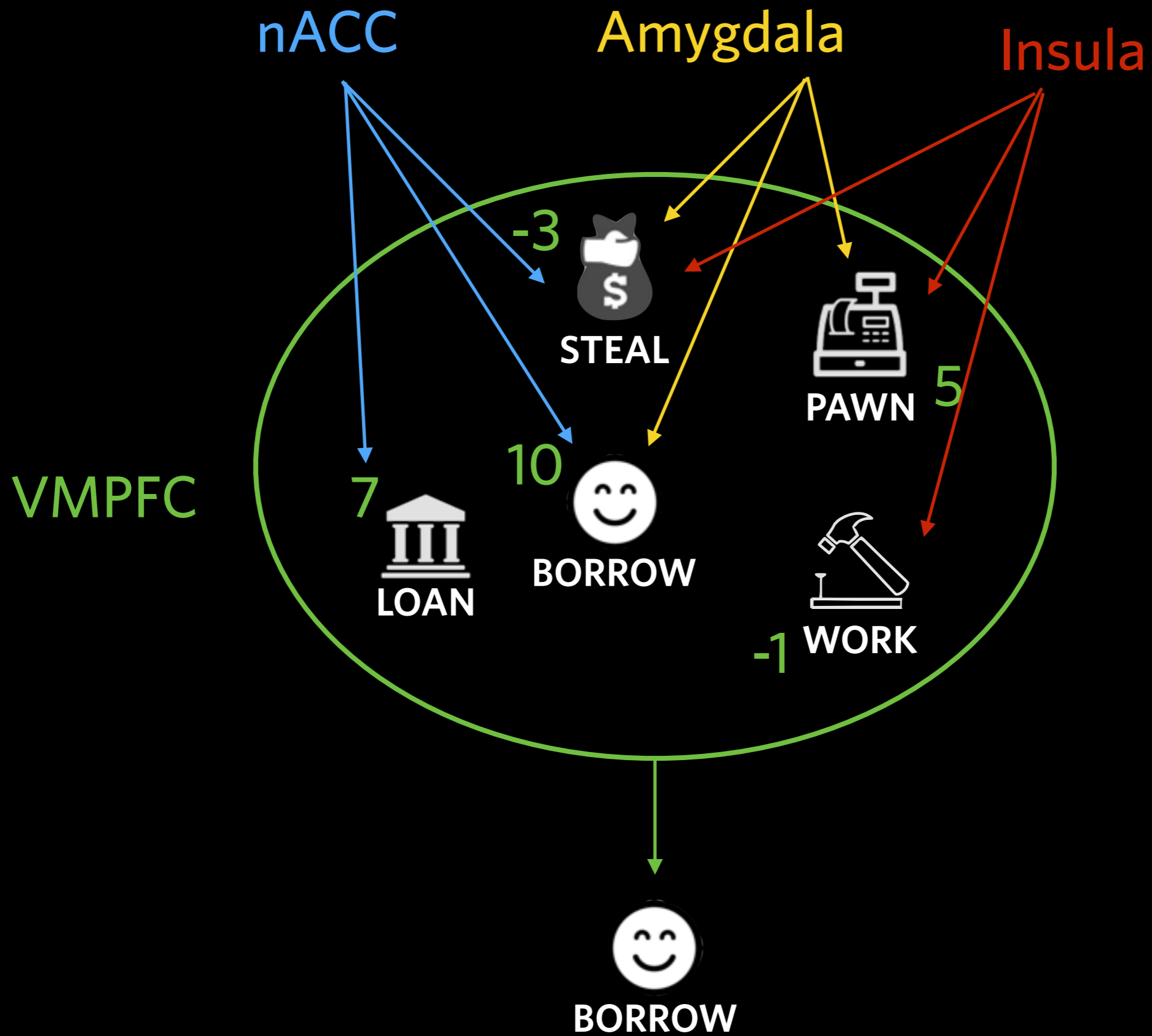
$\$ \neq \overset{\circ}{M}$

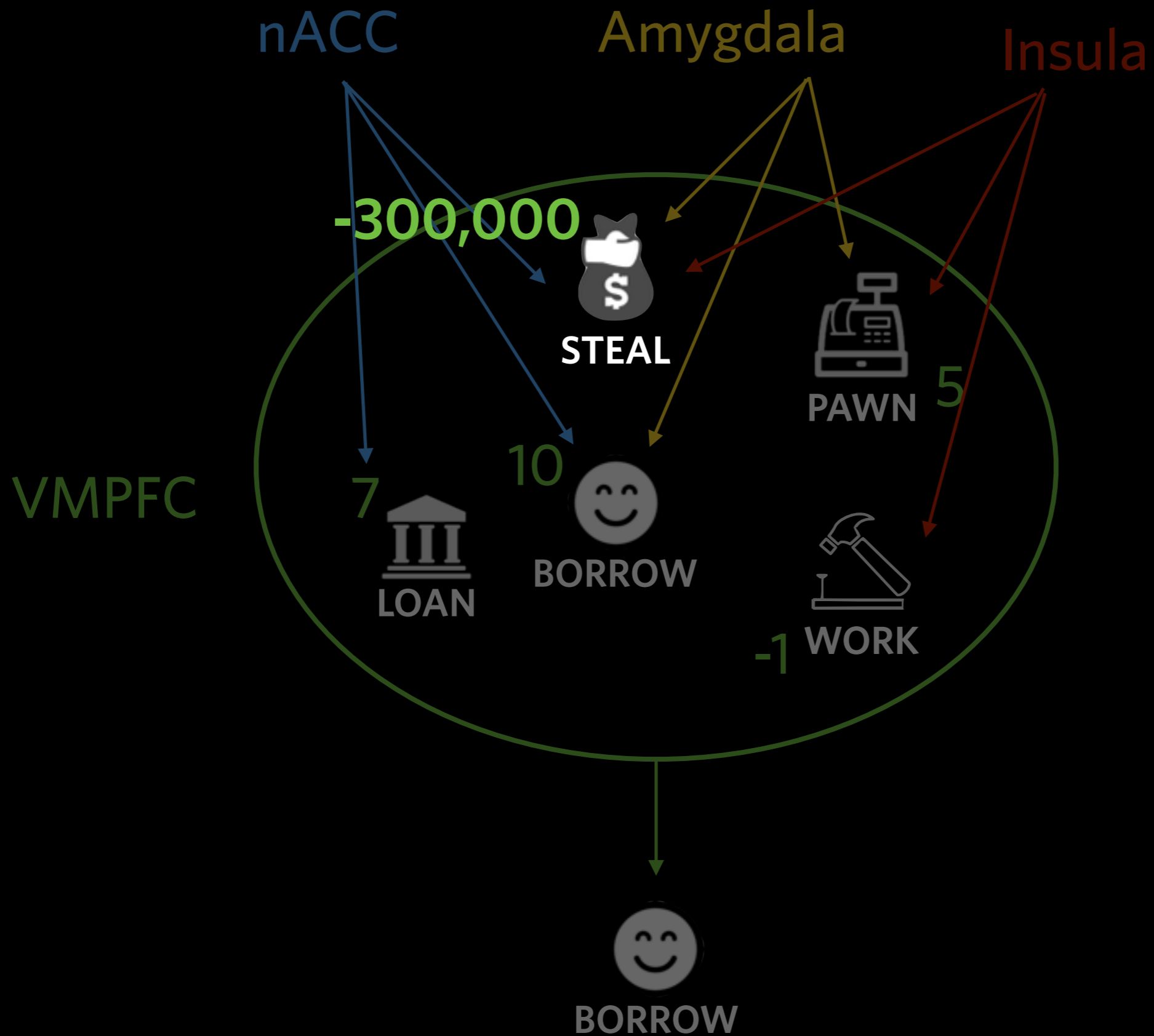




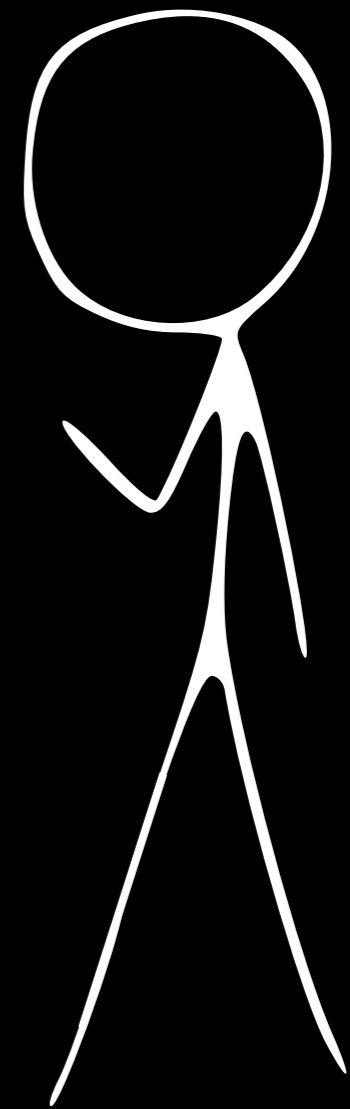
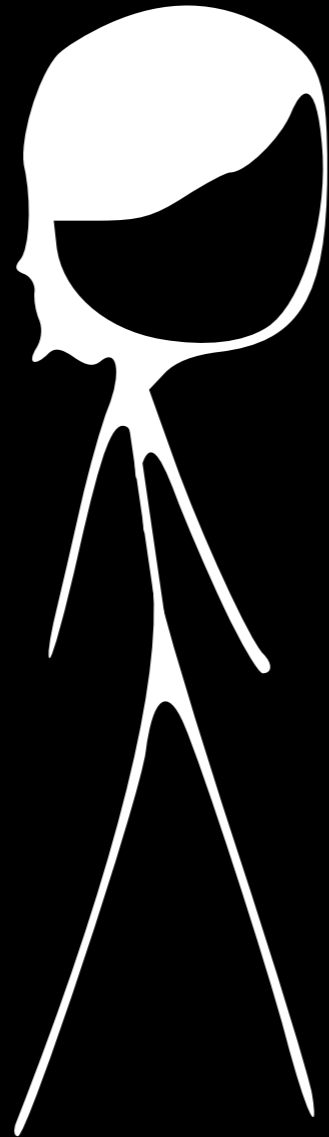


$$Q_{(s,a)} = Q_{(s,a)} + \alpha(R_s + \gamma Q_{(s',a')} - Q_{(s,a)})$$

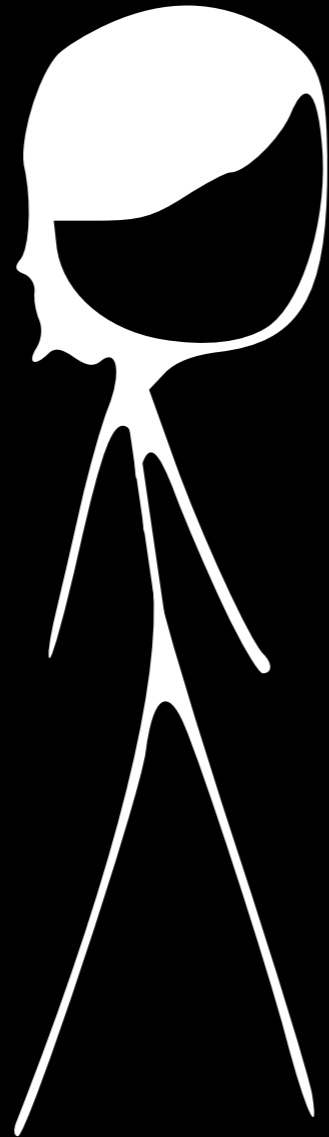
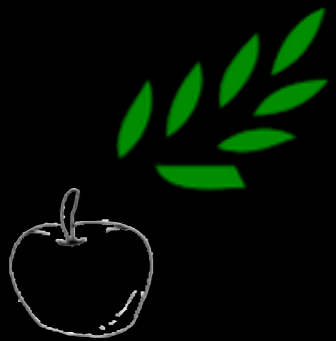




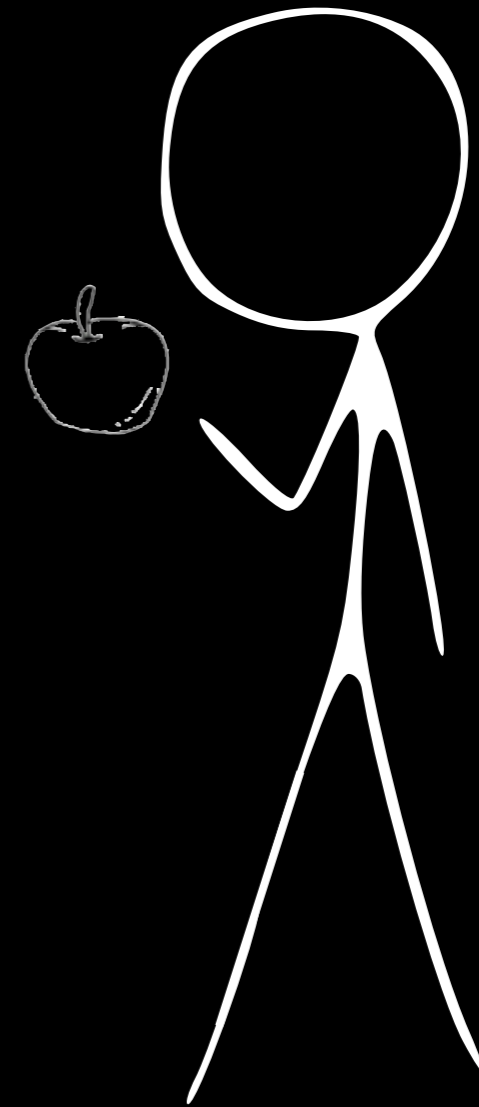
Evolution of Social Preferences



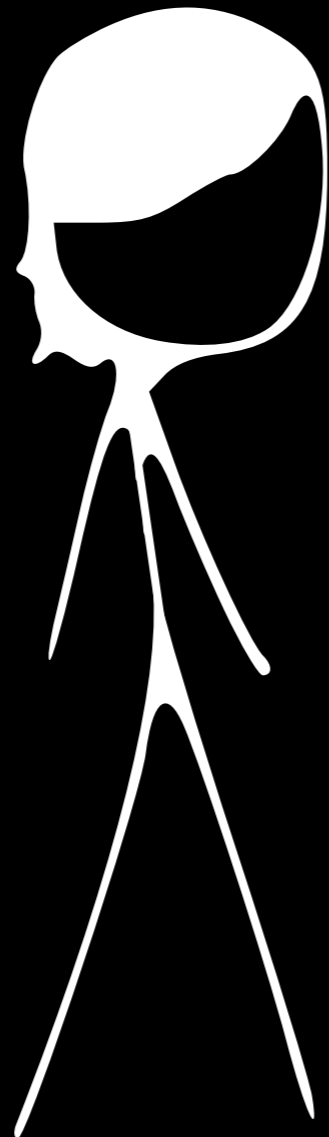
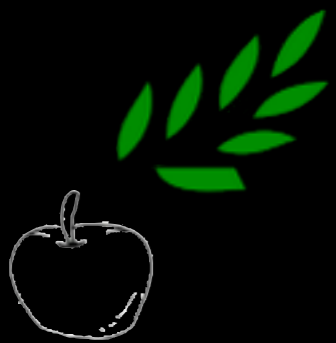
Evolution of Social Preferences



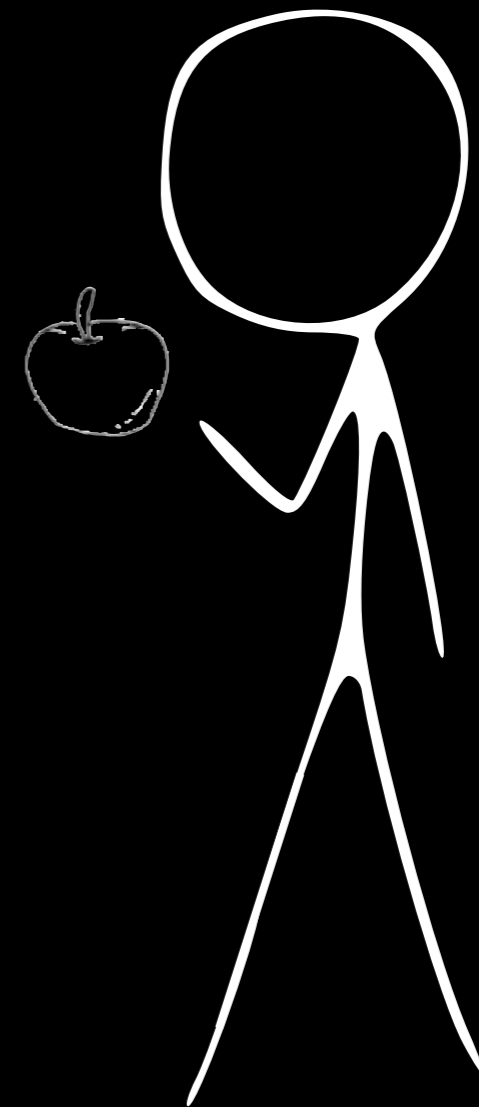
① Forage



Evolution of Social Preferences



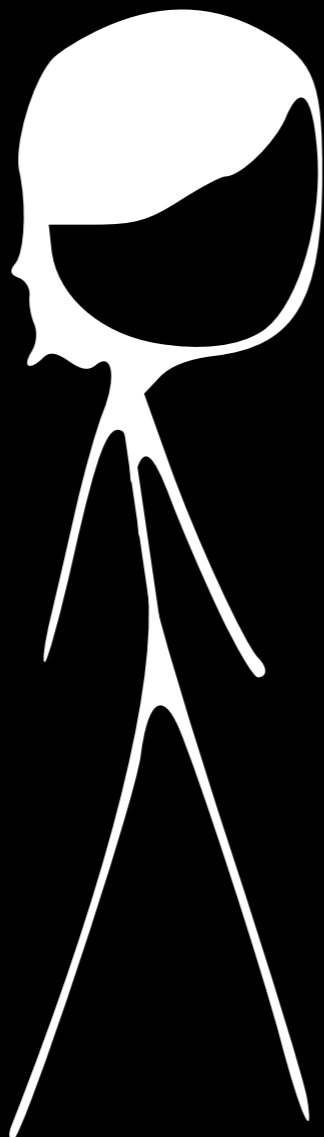
① Forage



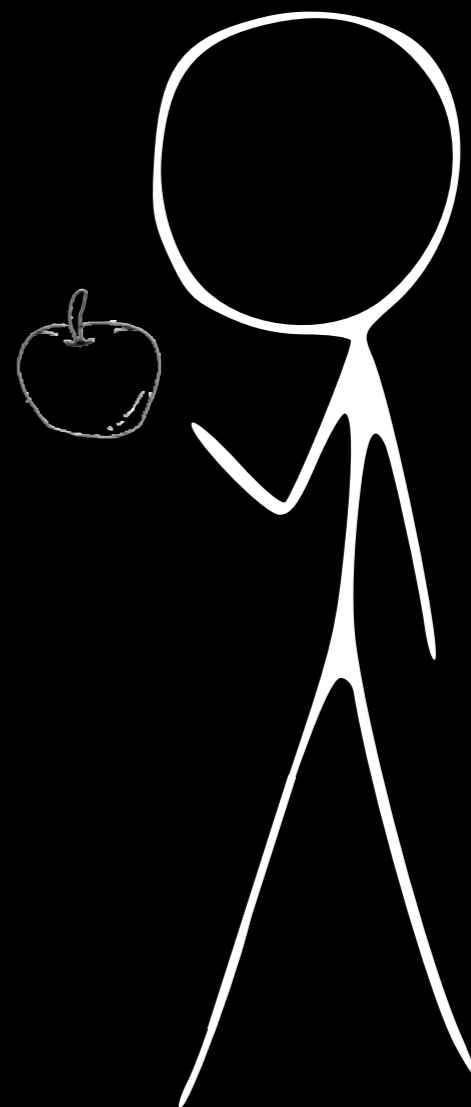
Evolution of Social Preferences



① Forage



① Forage



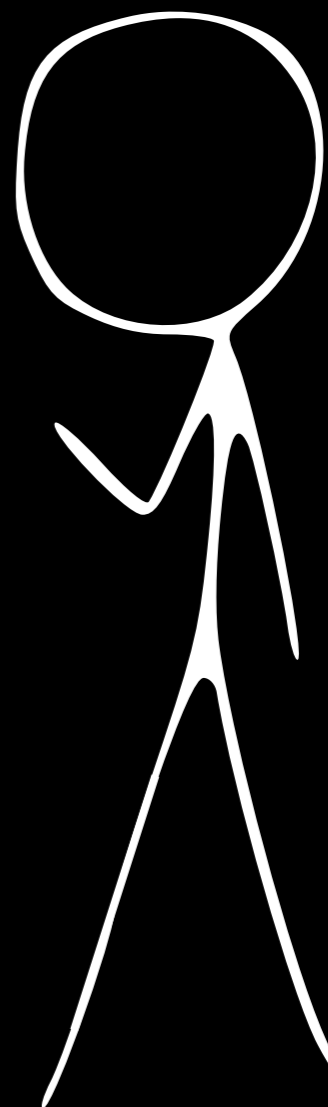
Evolution of Social Preferences



① Forage



② Steal



① Forage

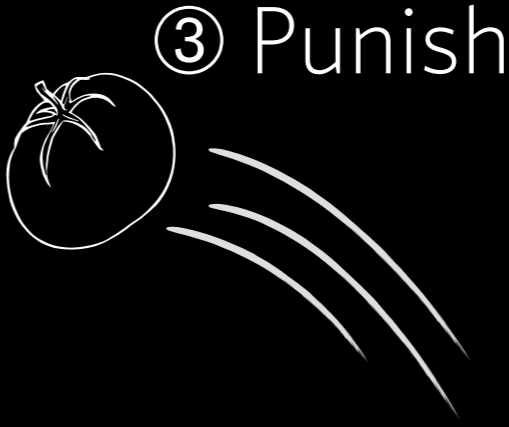
Evolution of Social Preferences



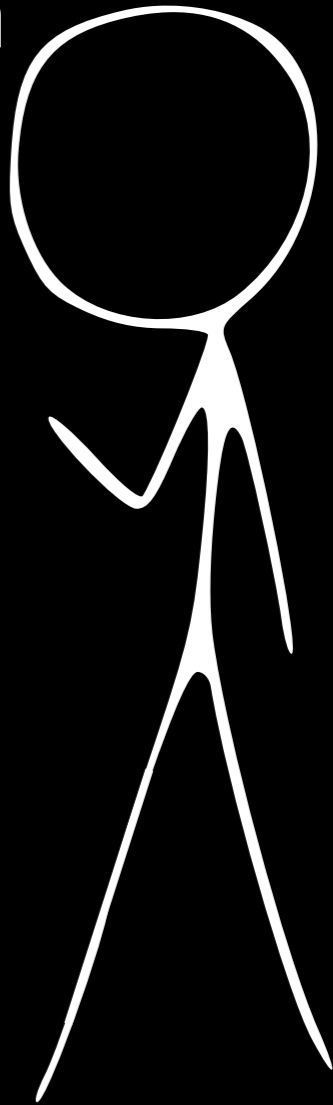
① Forage



② Steal



③ Punish



① Forage

Evolution of Social Preferences



① Forage



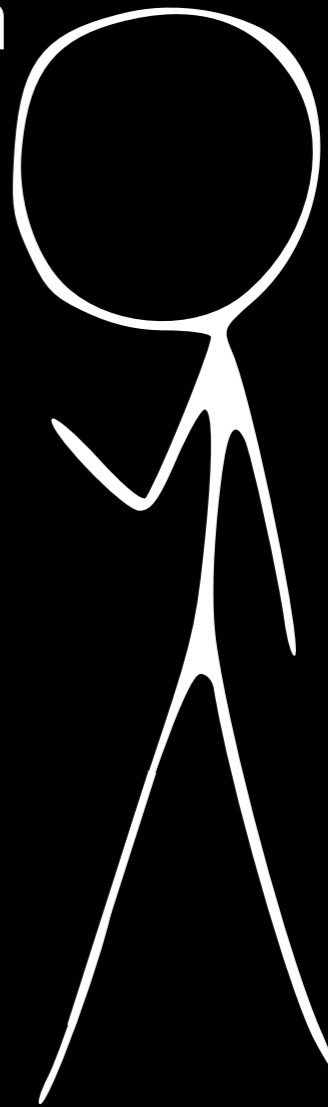
① Forage



② Steal



③ Punish



Evolution of Social Preferences




① Forage

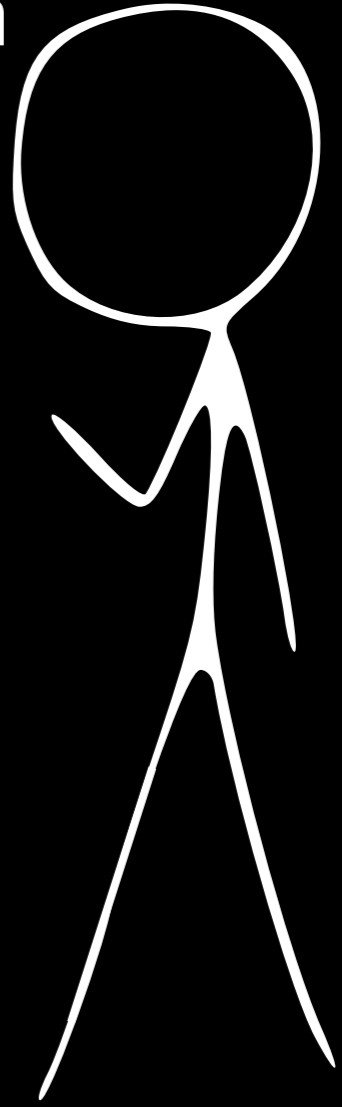


② Steal

+\$



③ Punish



\$

① Forage

Evolution of Social Preferences



① Forage



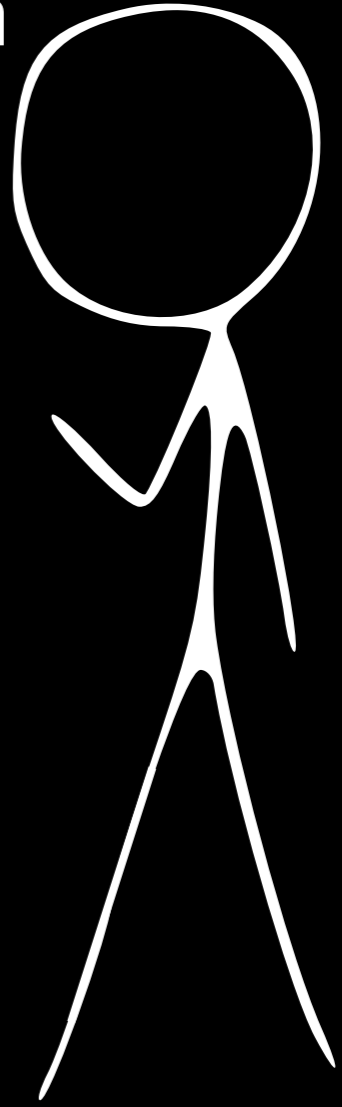
② Steal

- \$

+ \$



③ Punish



① Forage

\$



Evolution of Social Preferences



① Forage

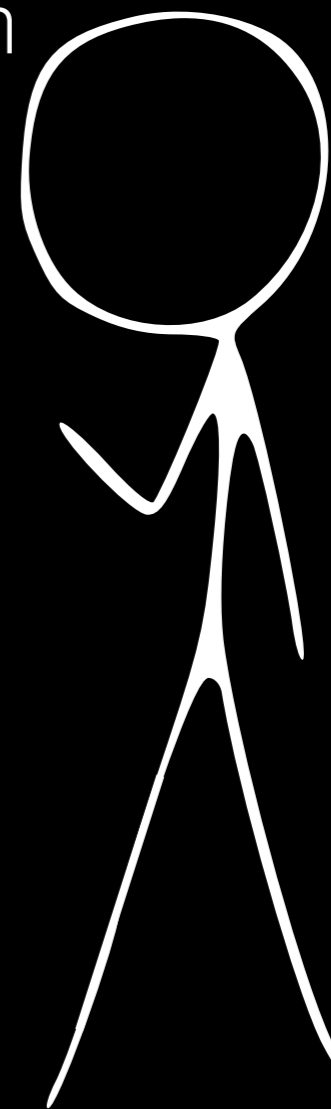


② Steal

-\$

+\$

③ Punish

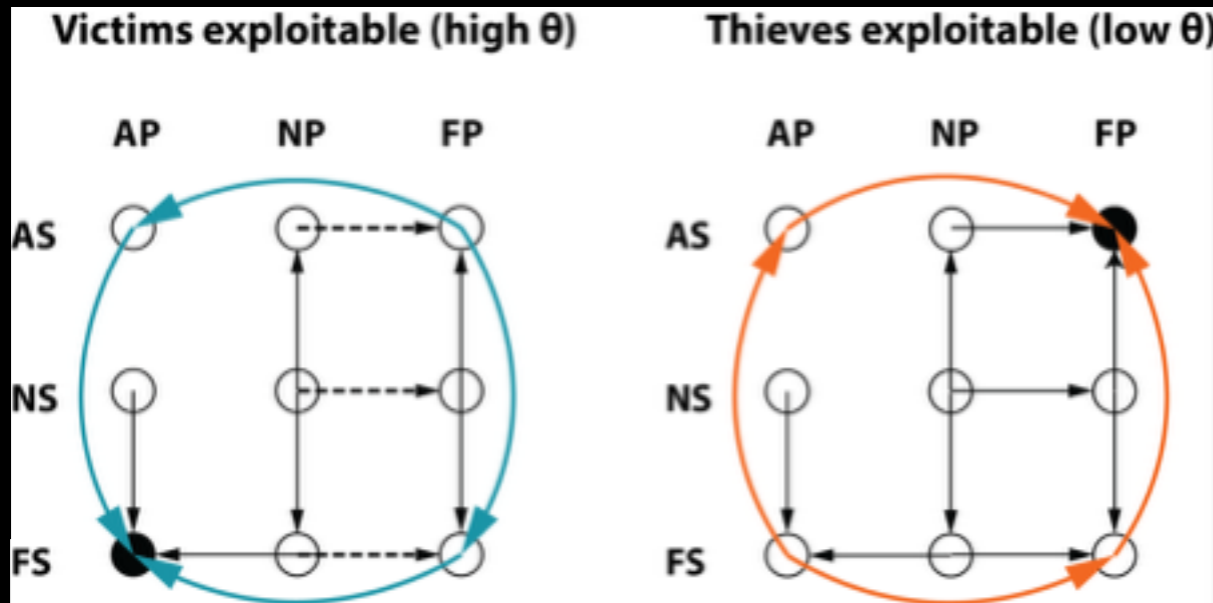


① Forage

\$



Evolution of Social Preferences



+\$

① Forage

③ Punish



①

Evolutionary Dynamics



② Steal

-\$

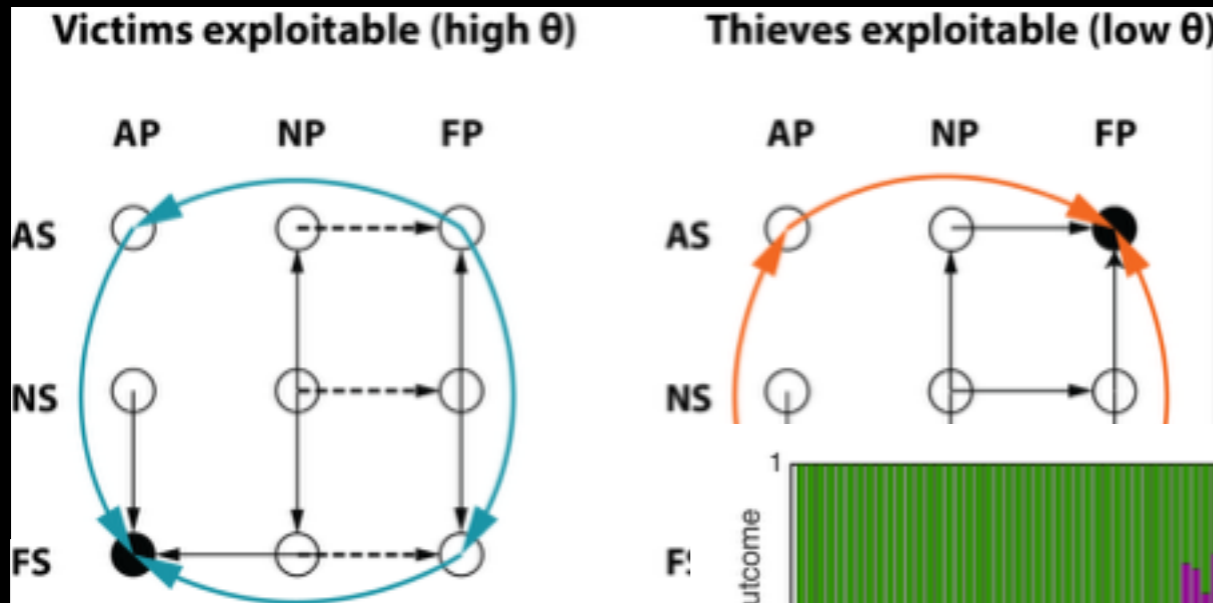


Flexible theft and resolution punishment:
Morris et al (under review)

Evolution of Social Preferences



① Forage

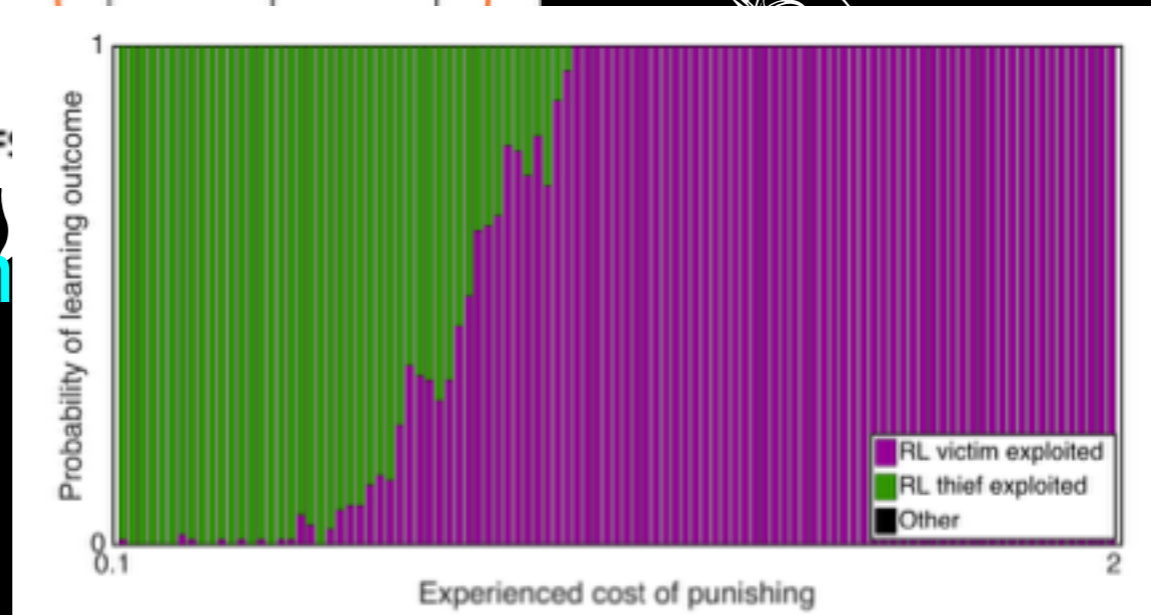


+\$

③ Punish

①

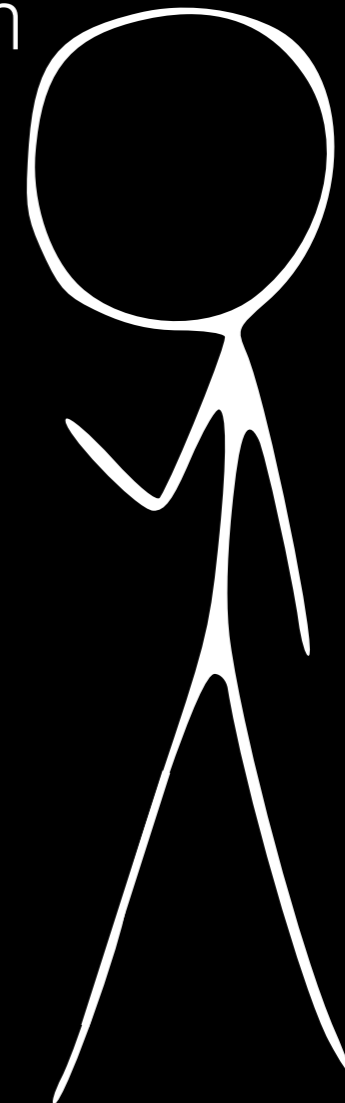
Evolutionary Dyn



Learning Dynamics

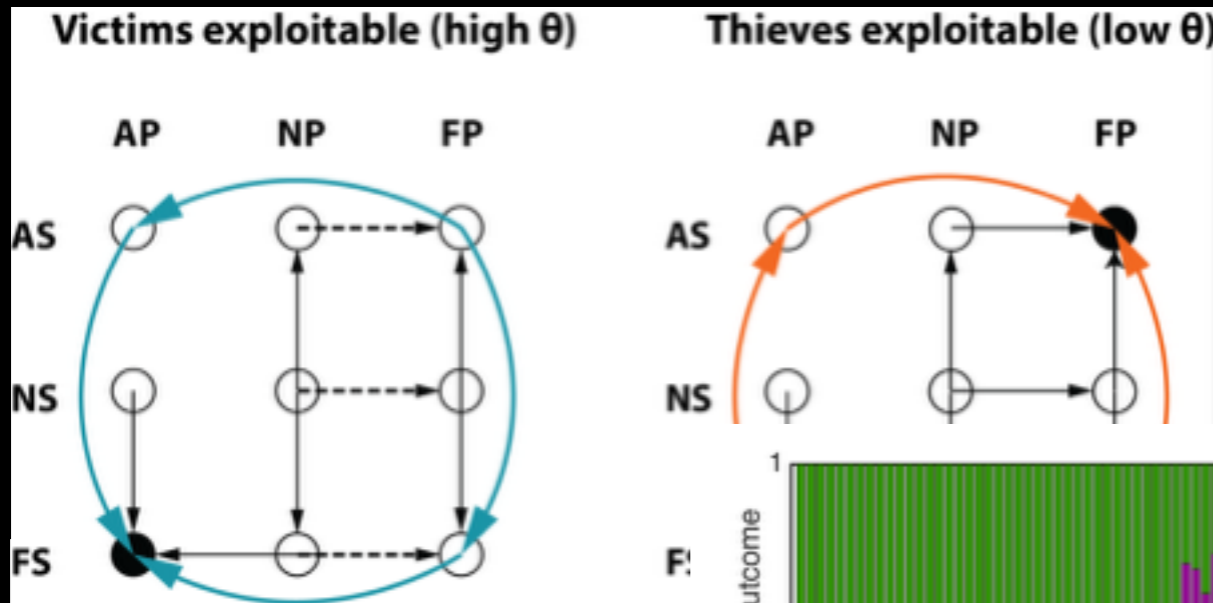
② Steal

-\$



Flexible theft and resolution punishment:
Morris et al (under review)

Evolution of Social Preferences



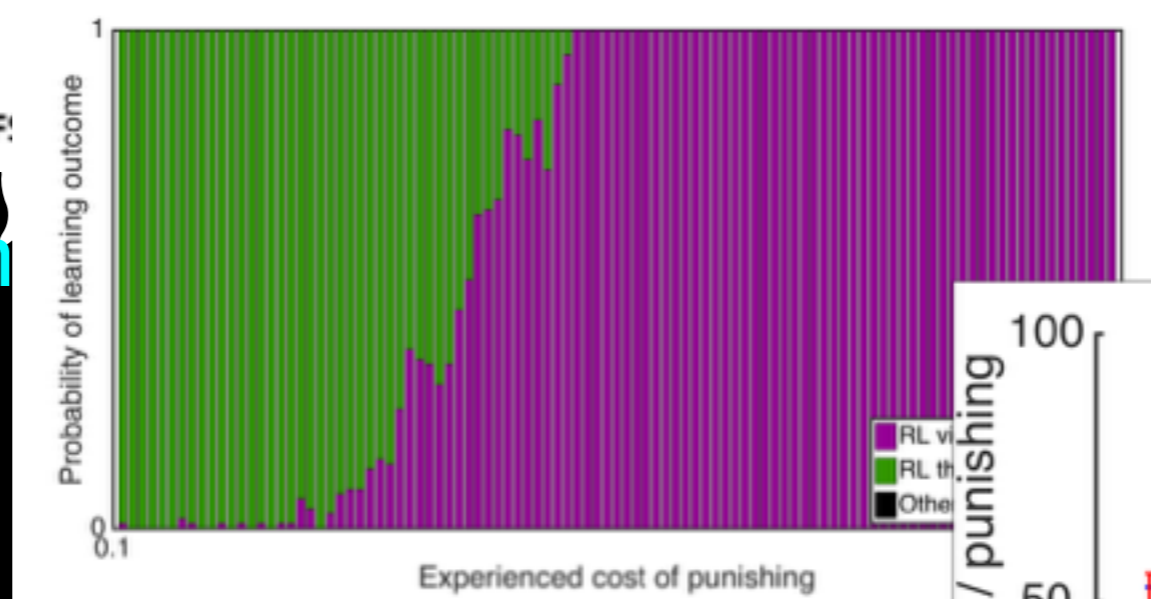
+\$

① Forage

③ Punish

①

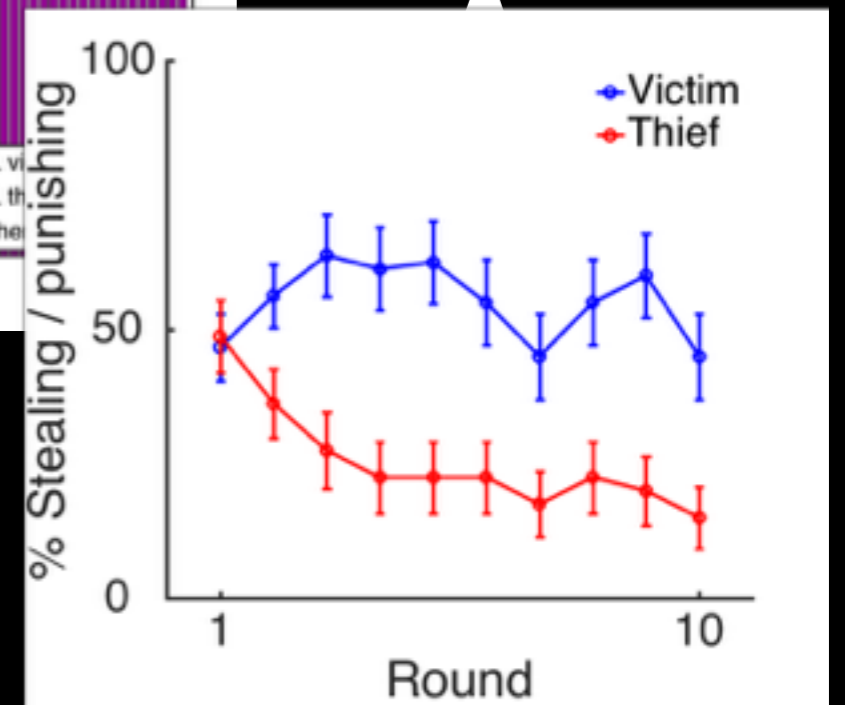
Evolutionary Dyn



Learning Dynamics

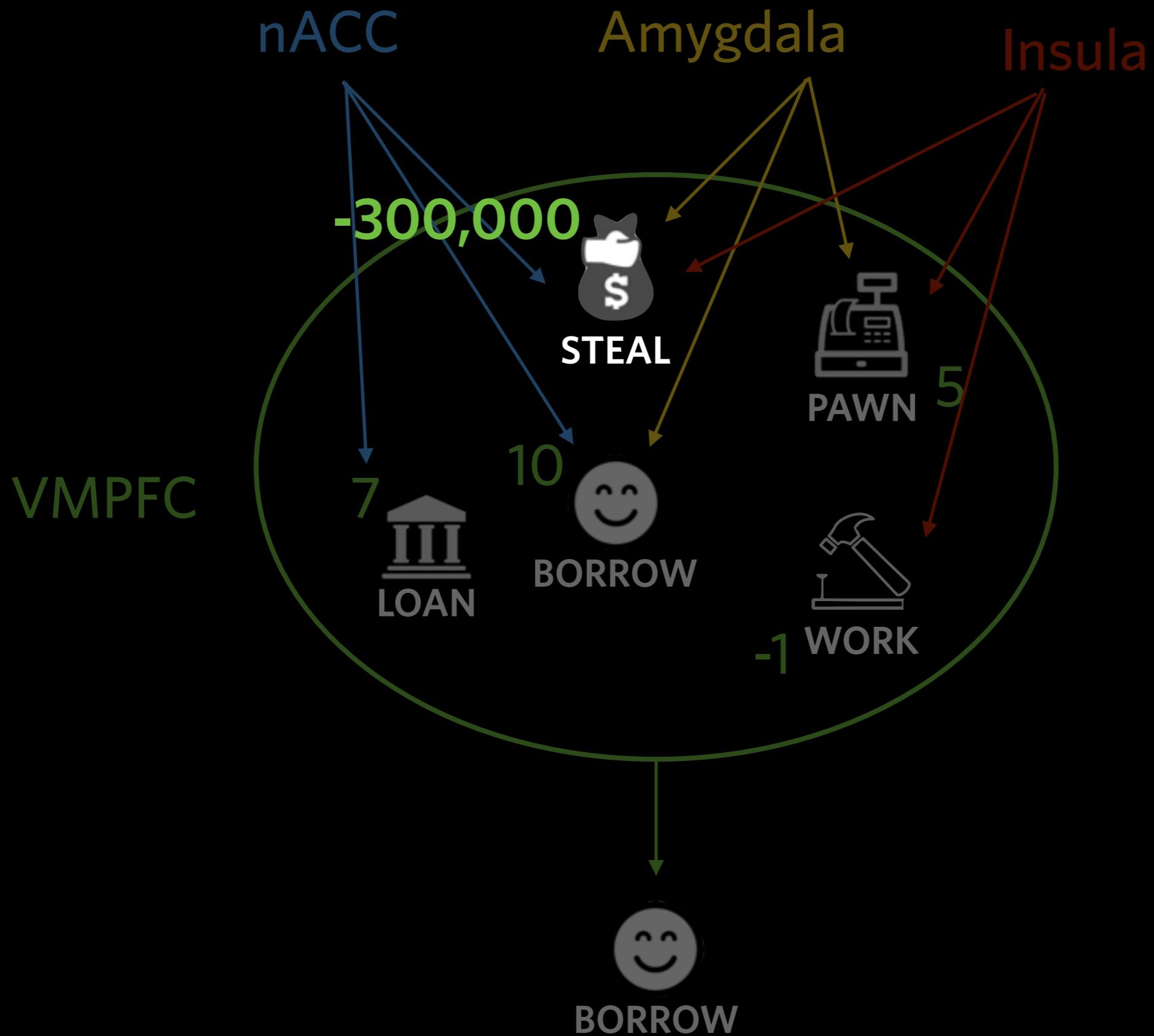
② Steal

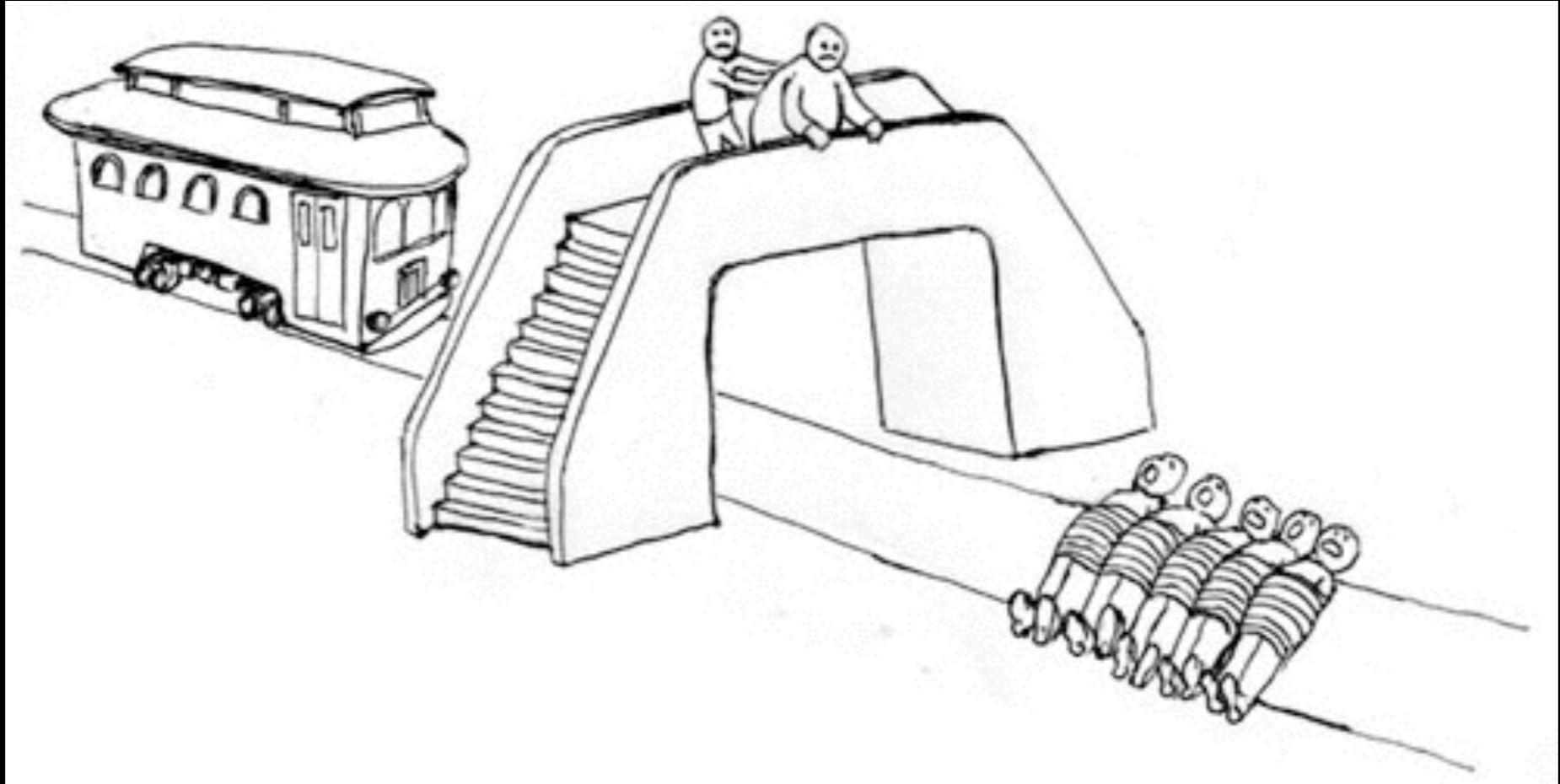
-\$



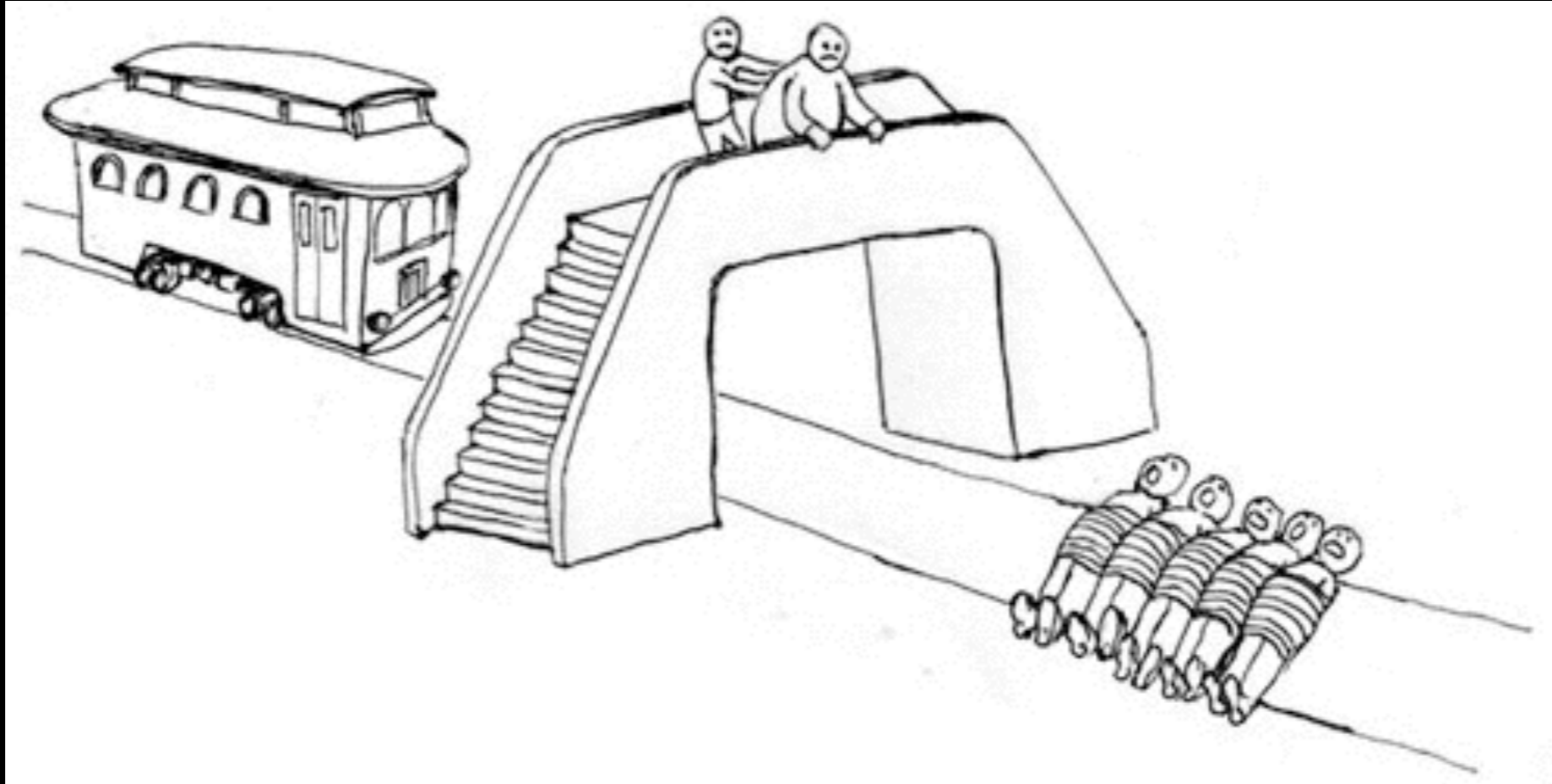
Human performance

Flexible theft and resolution punishment:
Morris et al (under review)

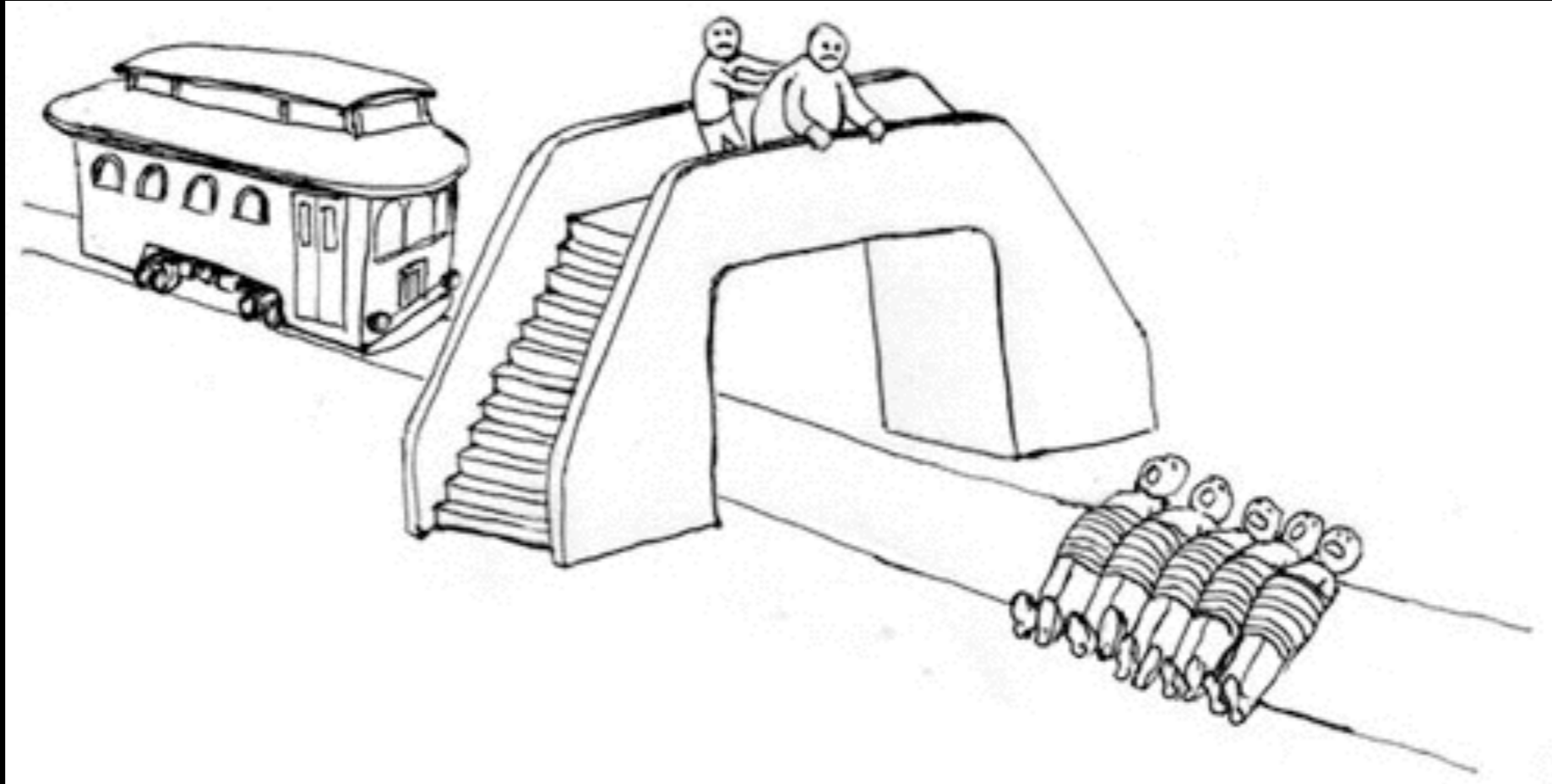




-300,000



-300,000



-300,000

-300,000

-300,000

-300,000

-300,000

Couple billed \$39.35 to hold newborn son

Linda Hervieux, Newser staff

9:43 a.m. EDT October 5, 2016



(NEWSER) – A Utah couple got a surprise when they got the bill for the birth of their son, Samuel.

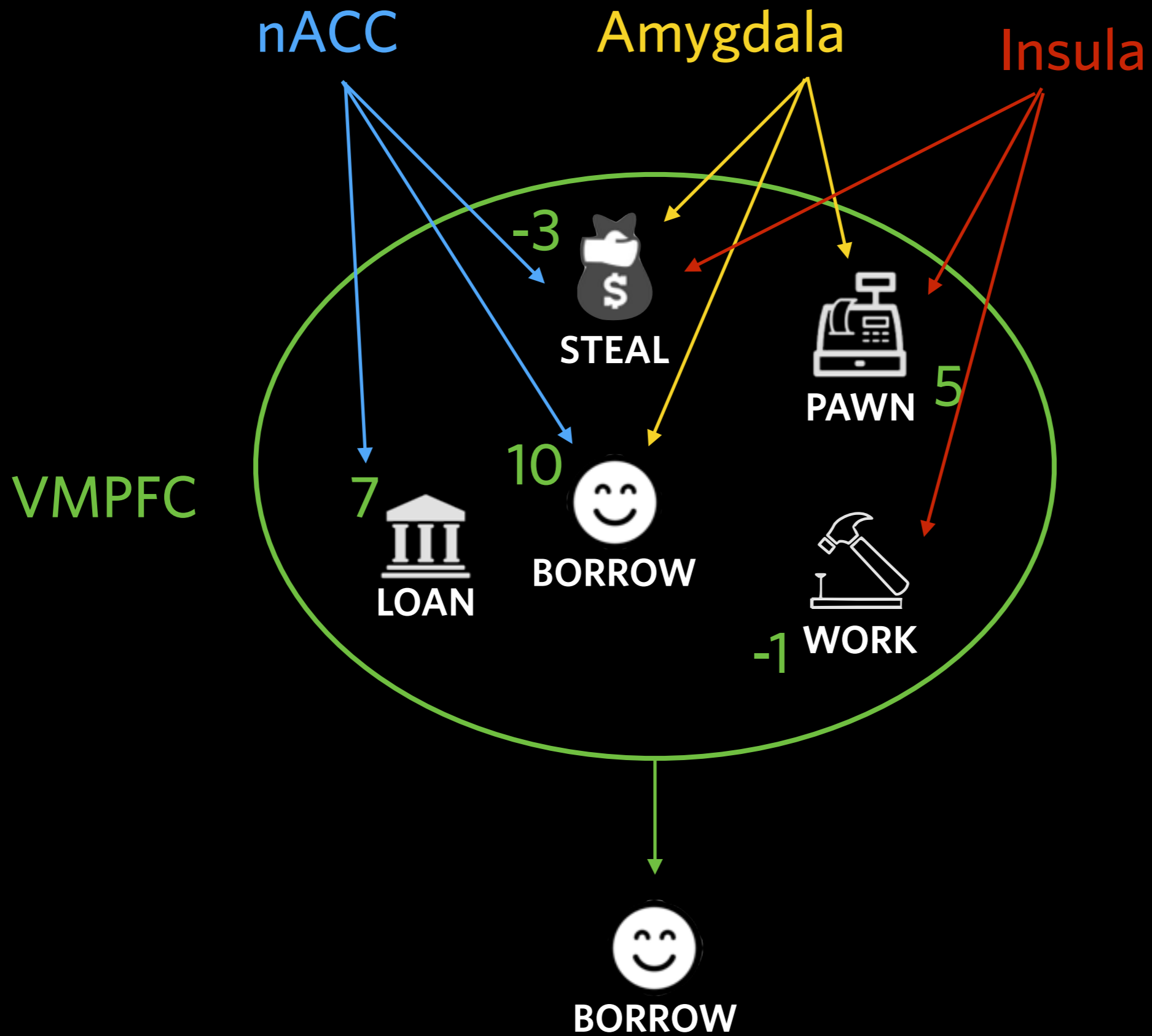
Tucked below the charge for "Delivery C Section" was a \$39.35 fee for "skin to skin," their fee for getting to hold the baby after delivery, explains [Vox](#).

In a post on [Reddit](#) that has gone viral, user "halfthrottle" writes: "During the C-section the nurse asked my wife if she would like to do skin to skin after the baby was born. Which of course anyone would say yes (to). We just noticed it in the bill today and had a laugh."

-300,000

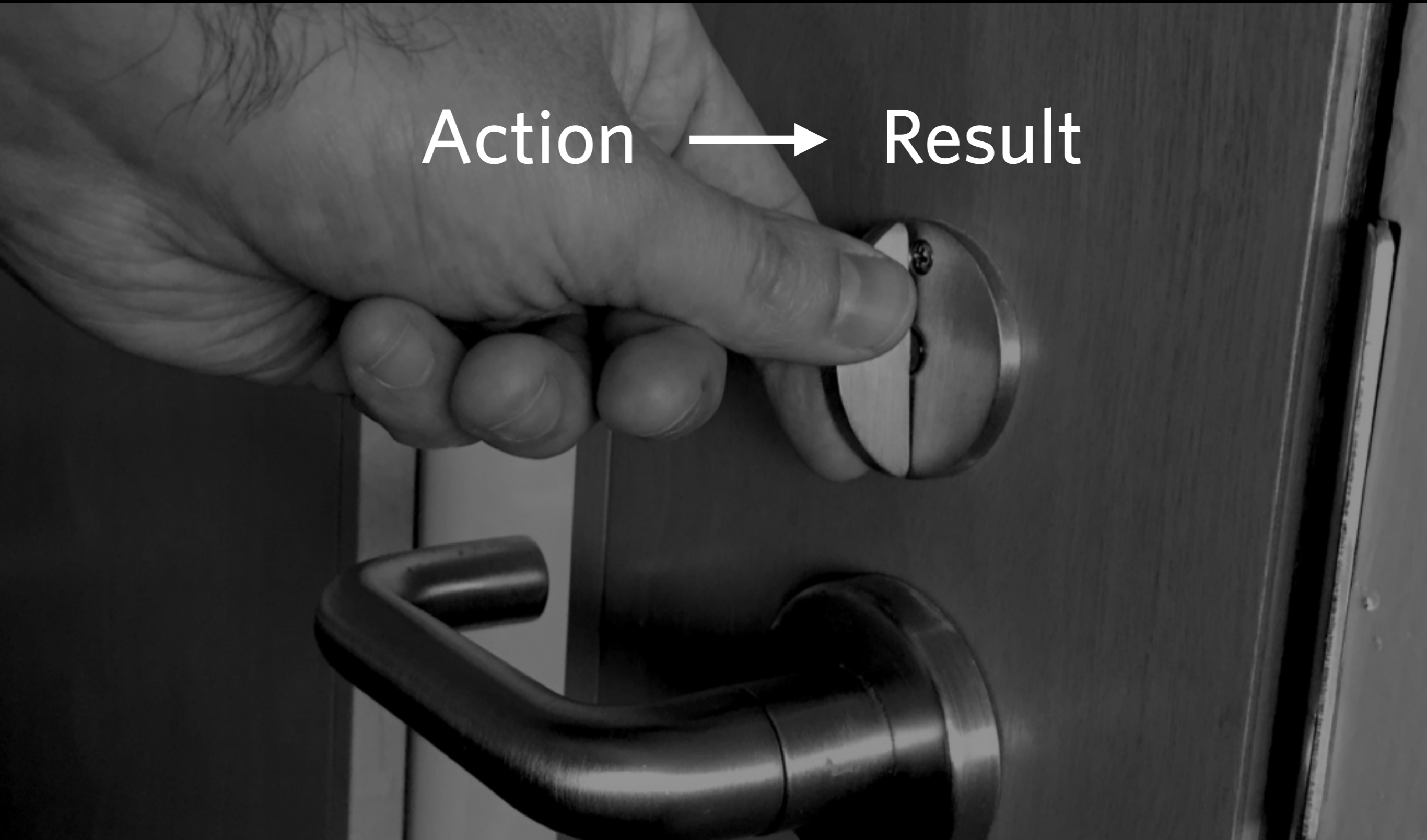
-300,000

-300,000



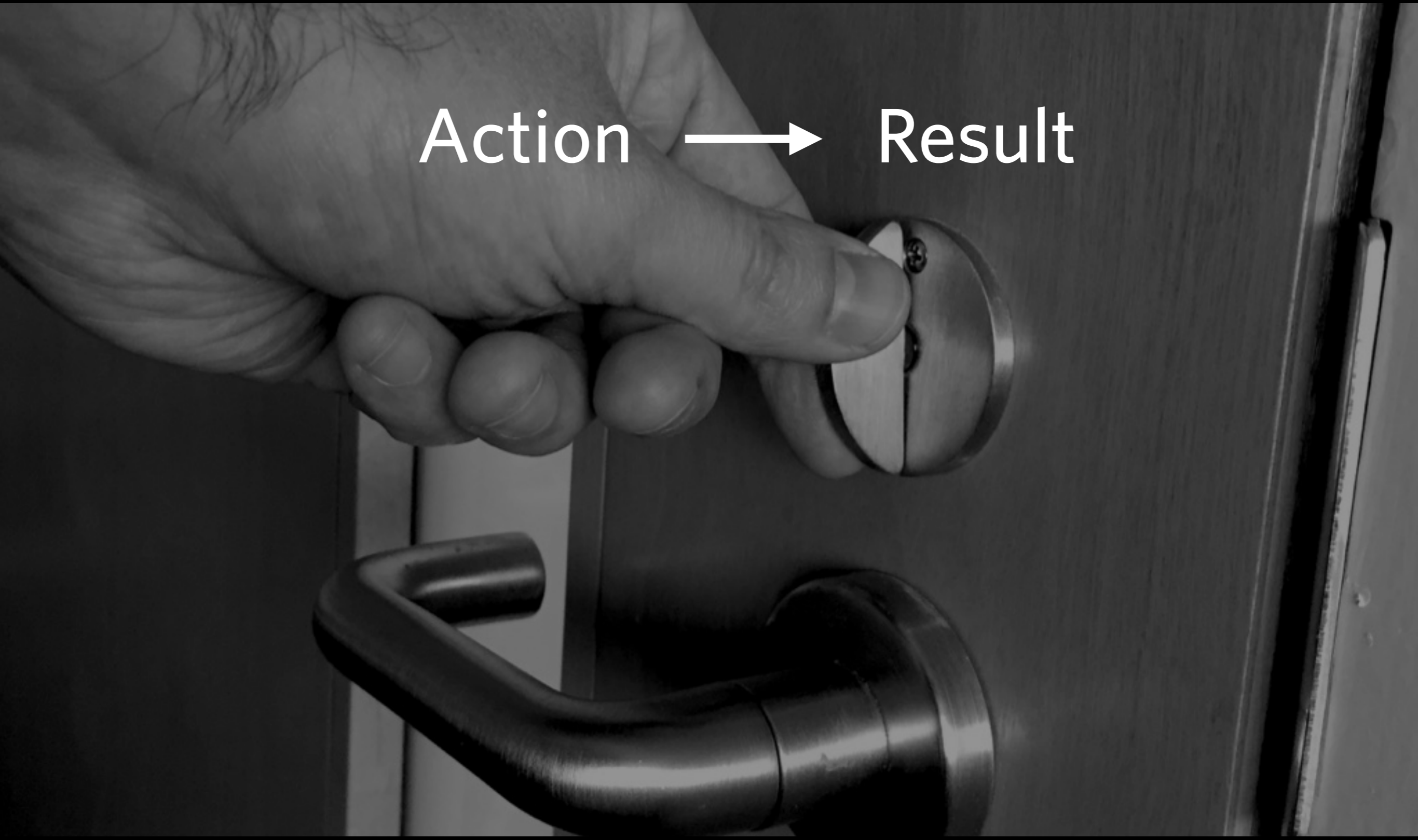


Action → Result



**GOAL
DIRECTED**

Action → Result



**GOAL
DIRECTED**

Action



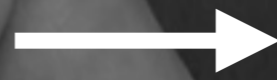
Result



HABITUAL



Action



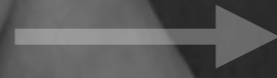
Result



HABITUAL



Action



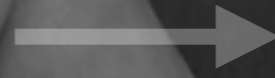
Result



HABITUAL



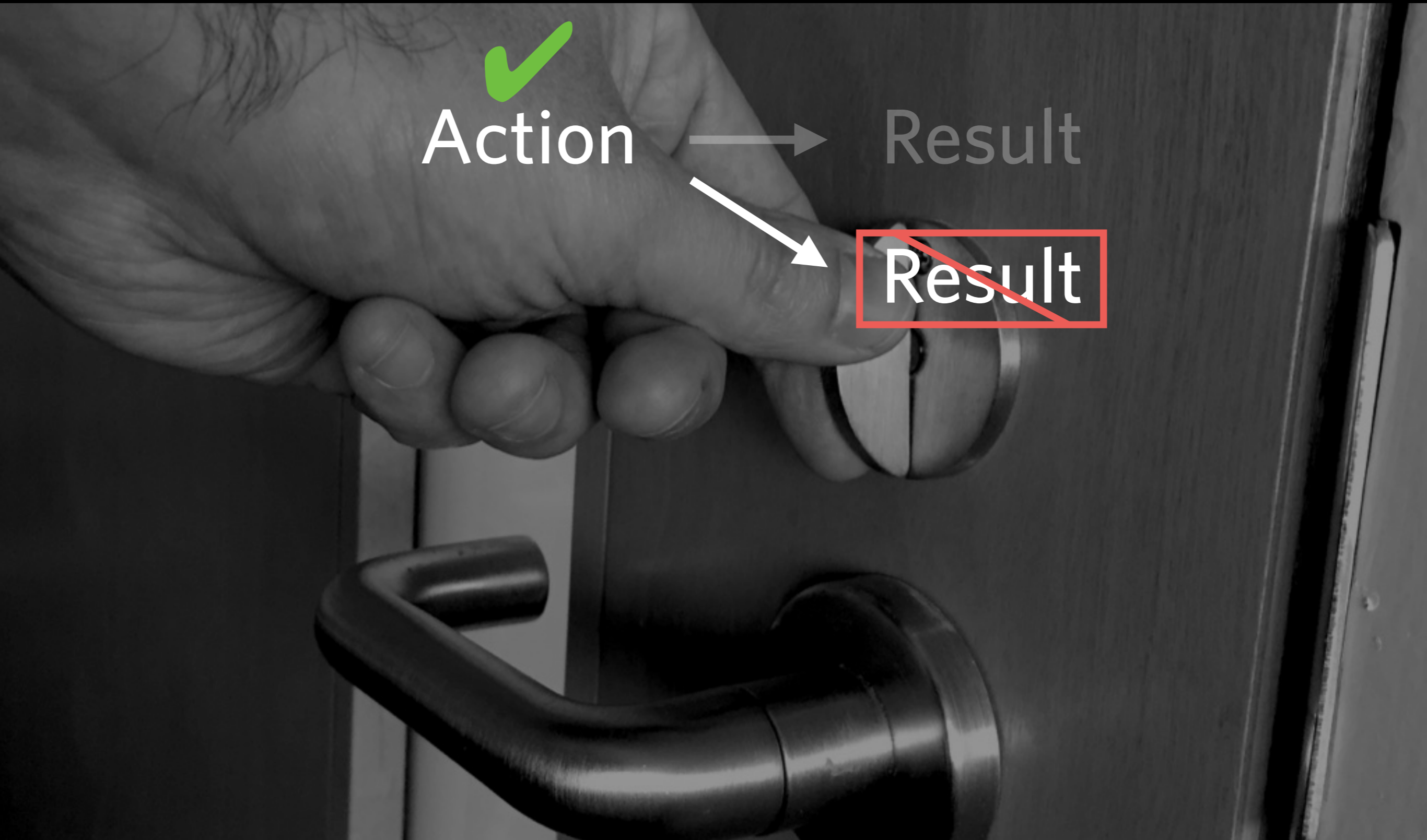
Action



Result



~~Result~~



HABITUAL



Action



Result



Result



HABITUAL



Action



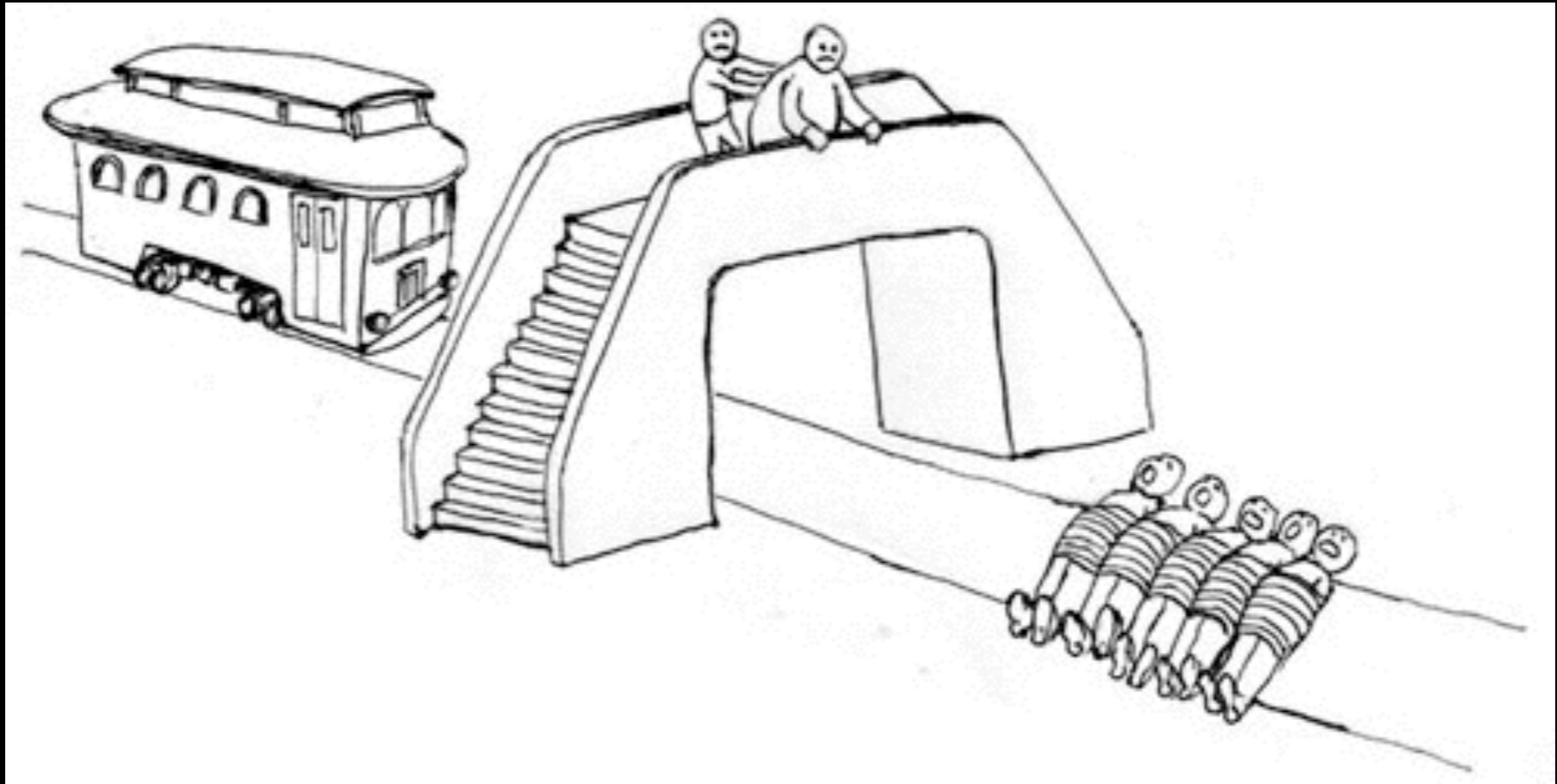
Result



~~Result~~

Tipping = Moral Habit









Total Peripheral
Resistance

100

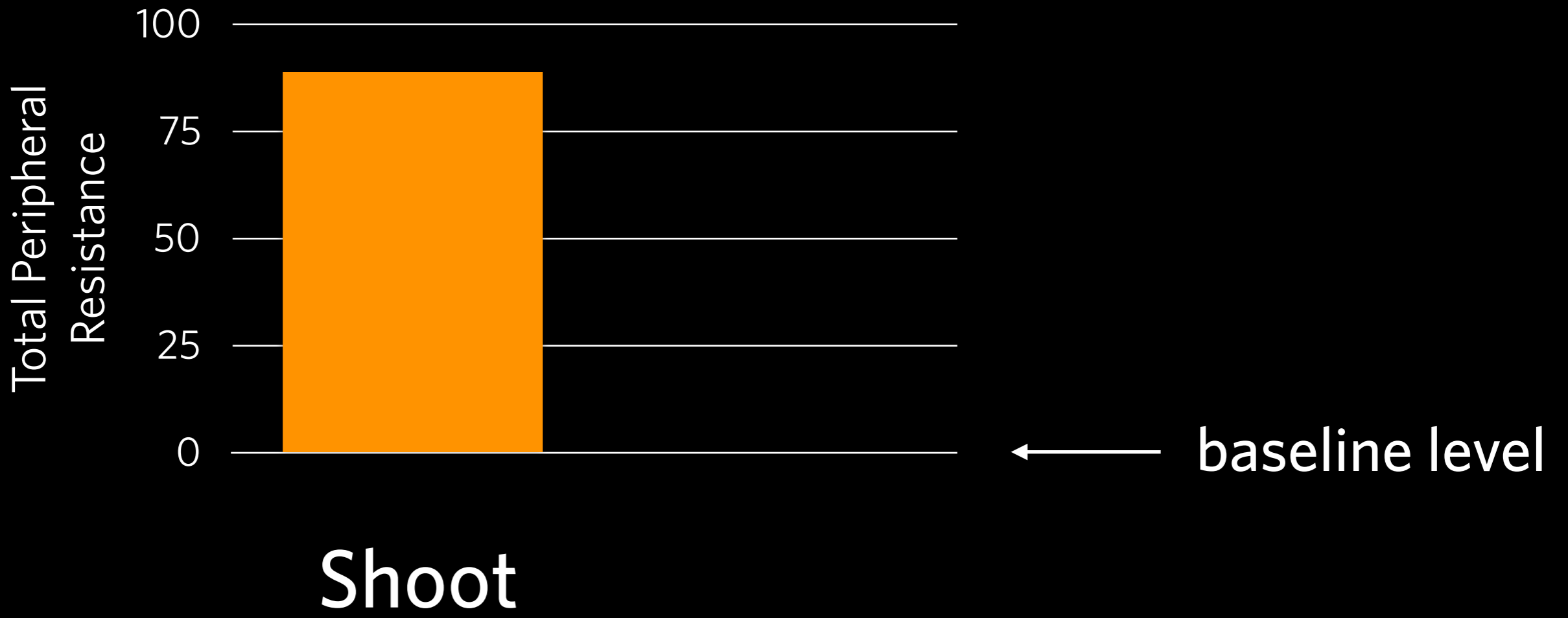
75

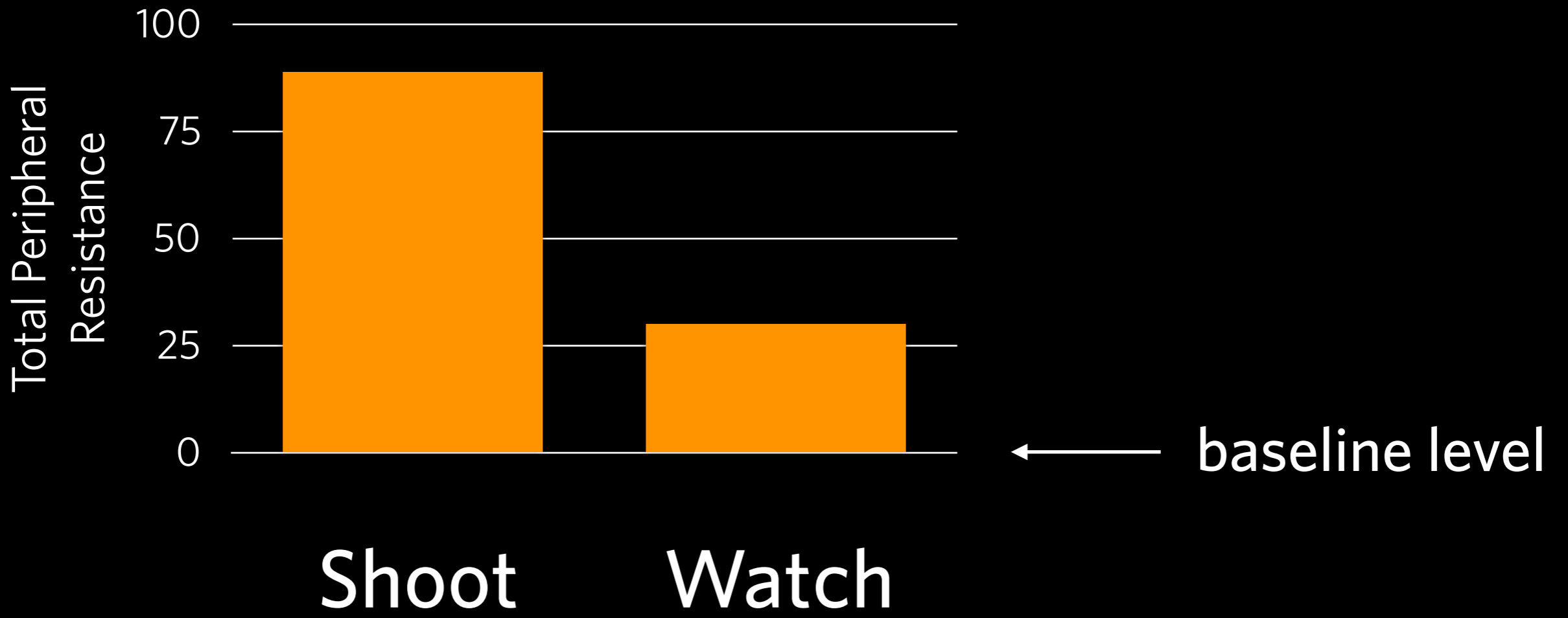
50

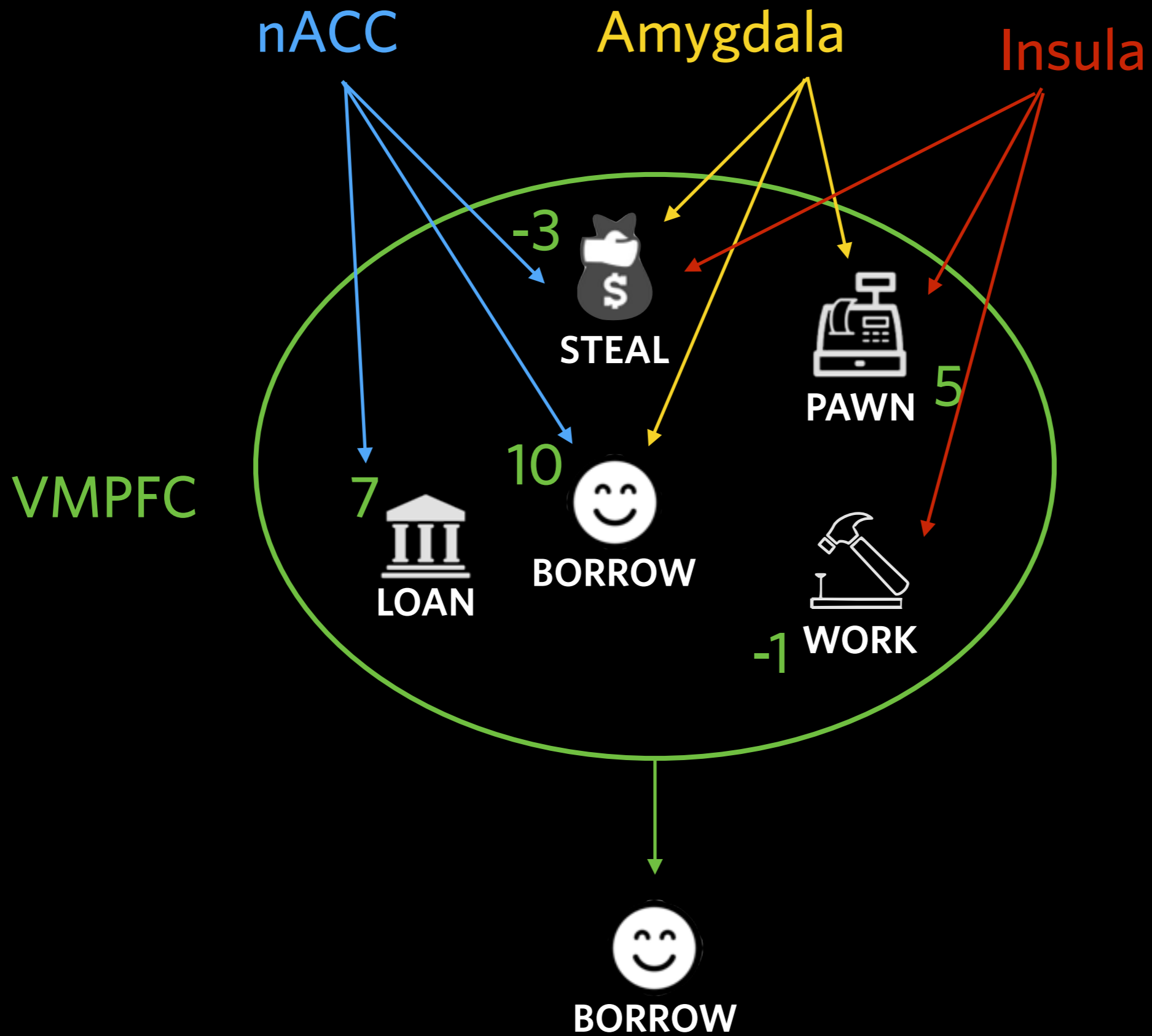
25

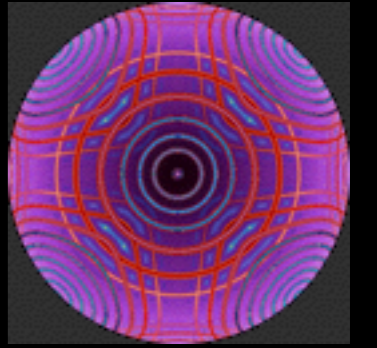
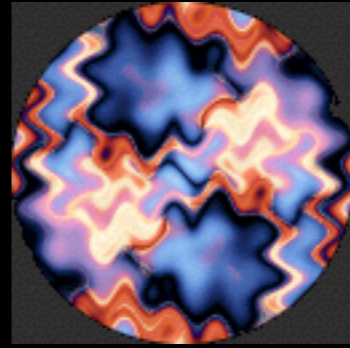
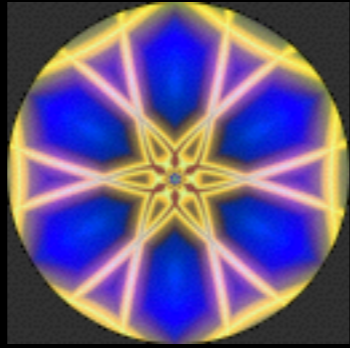
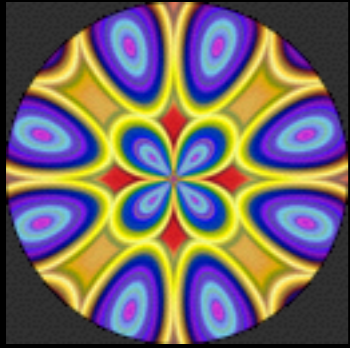
0

← baseline level



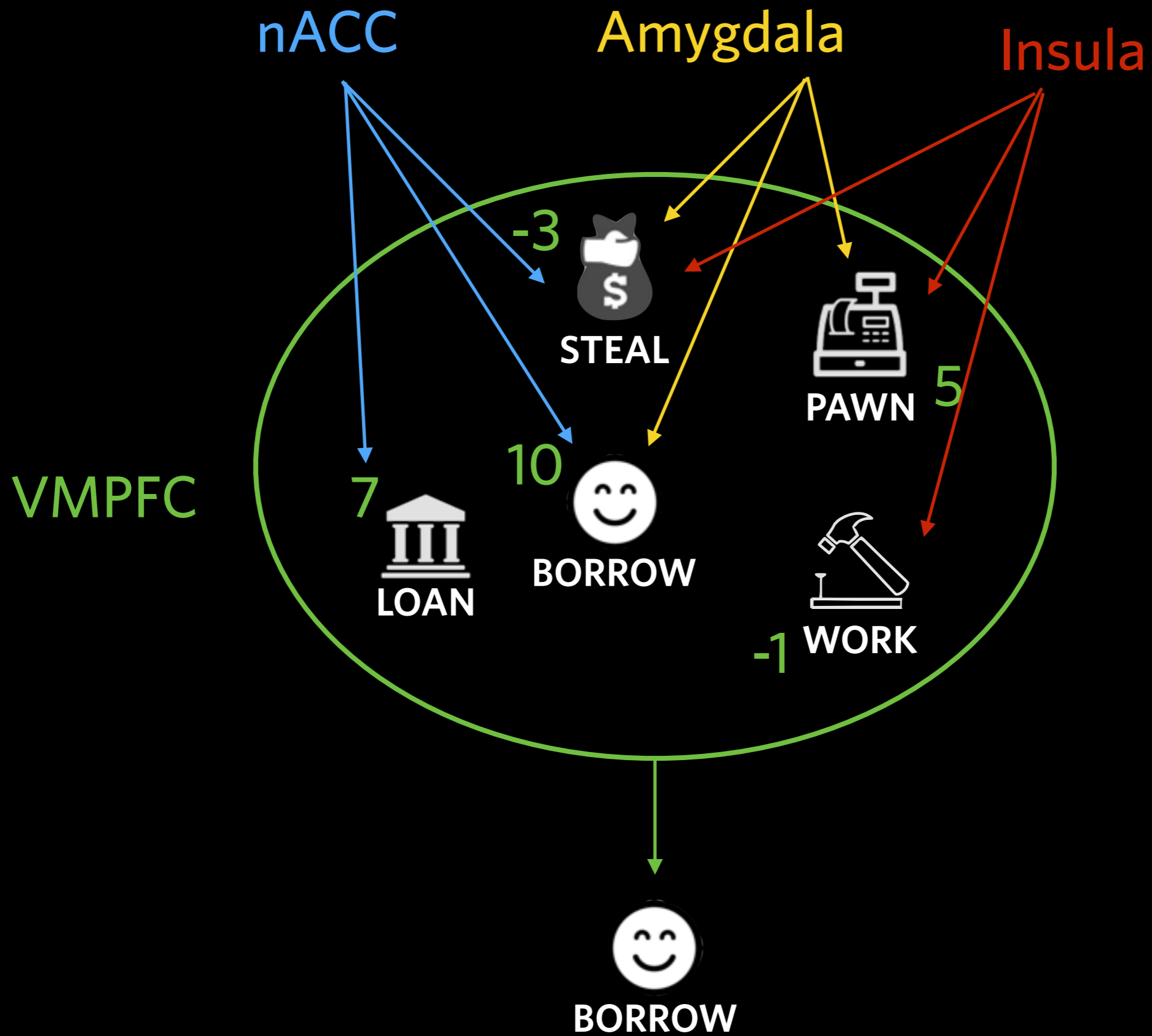


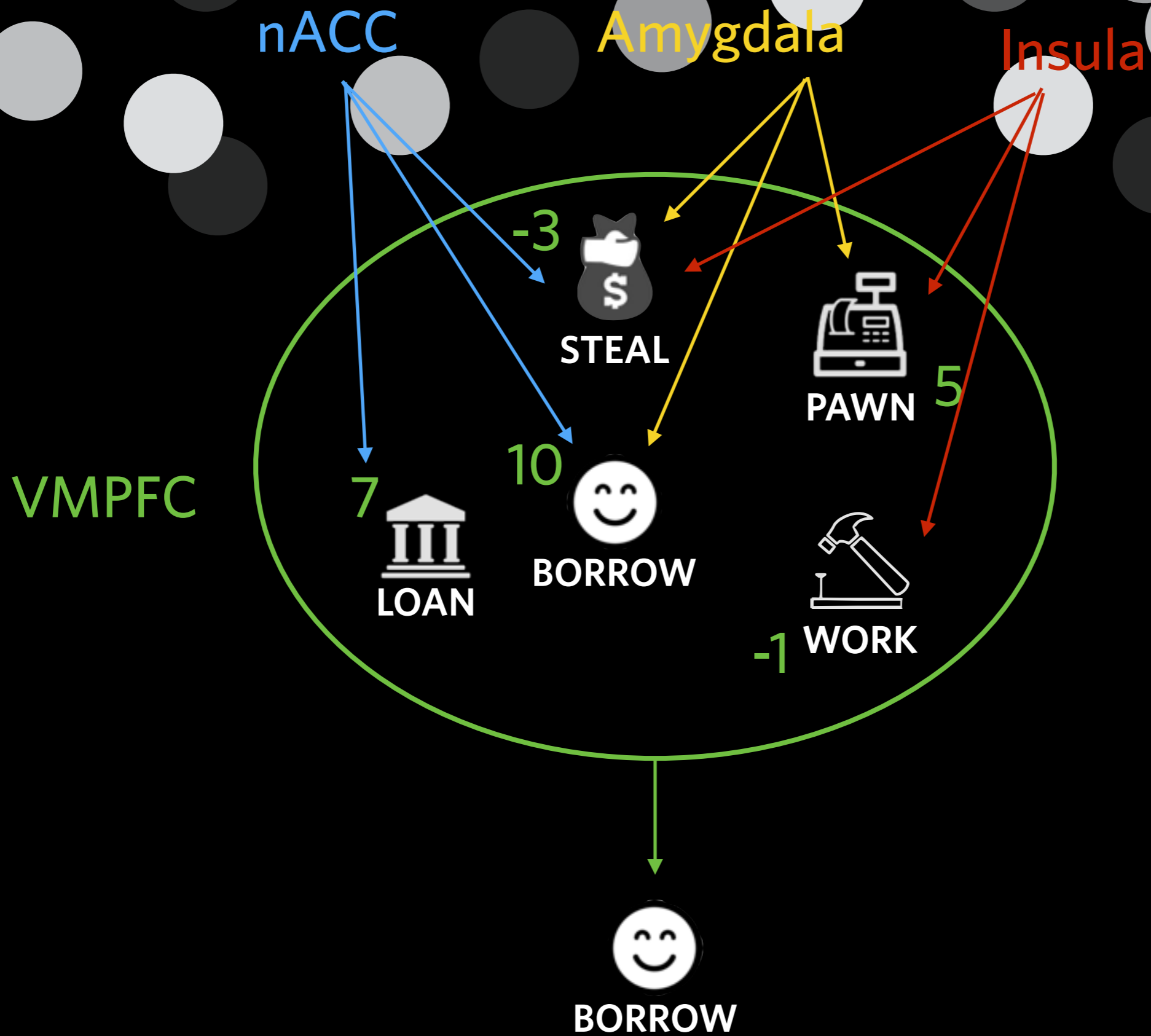


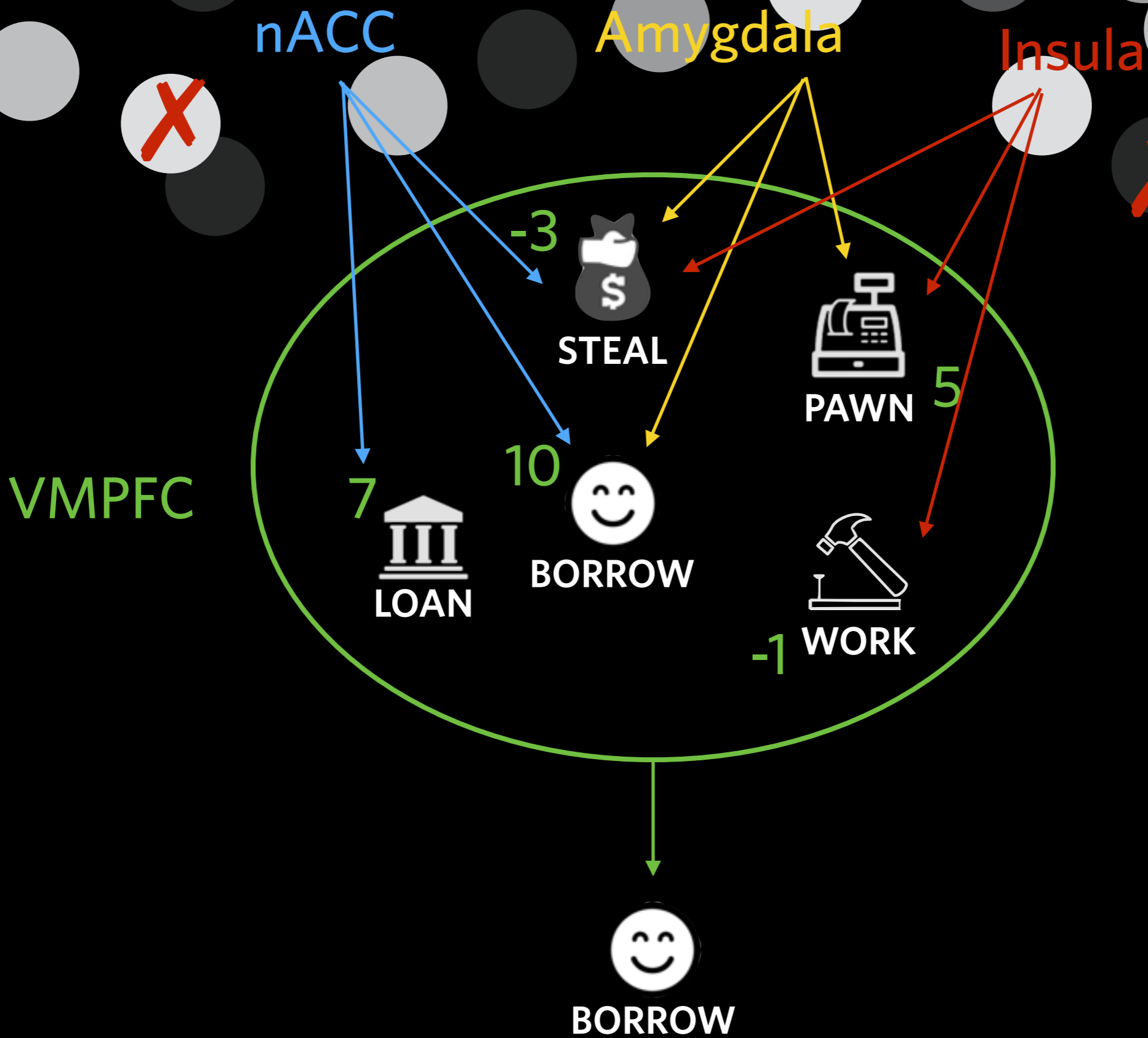


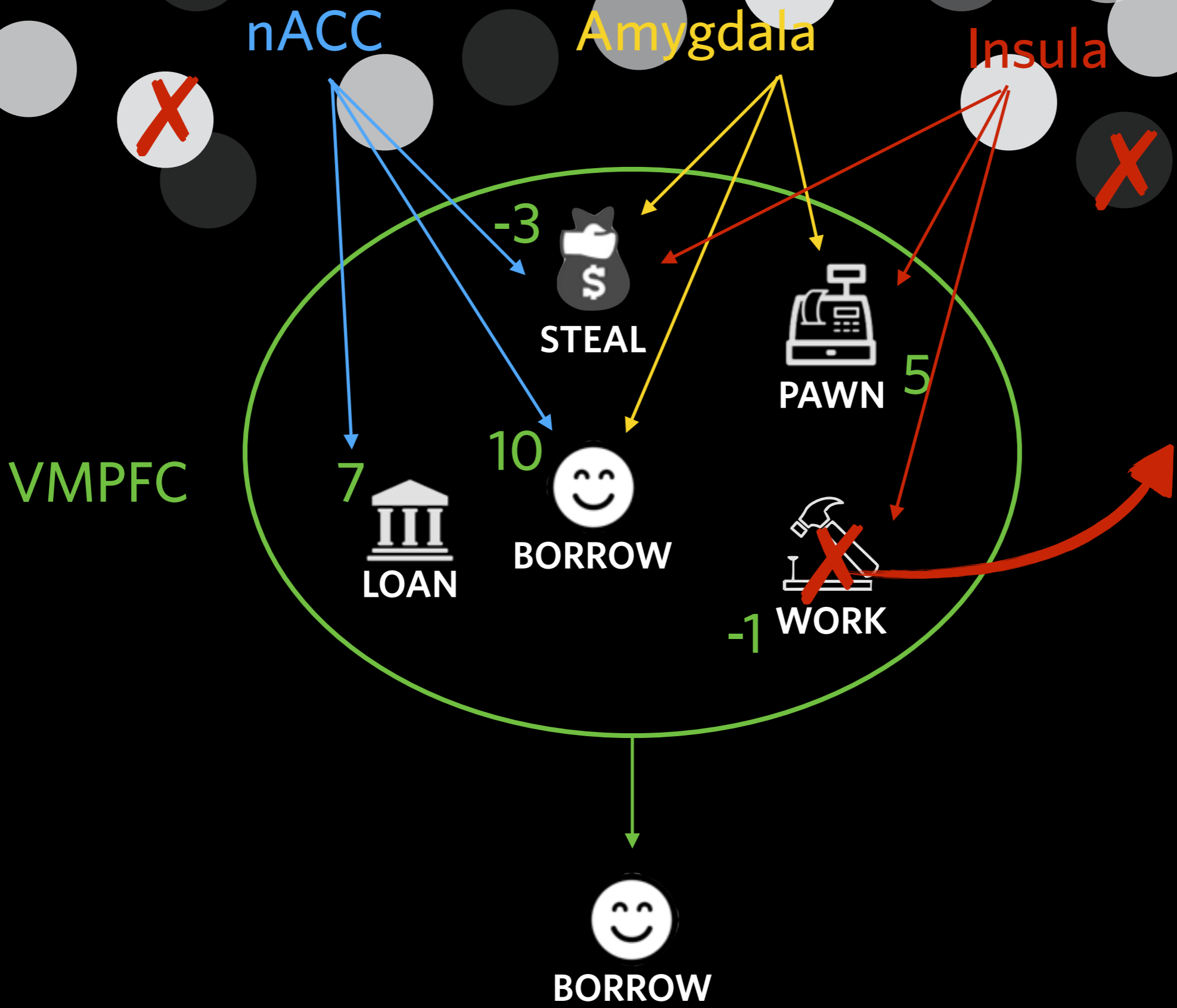








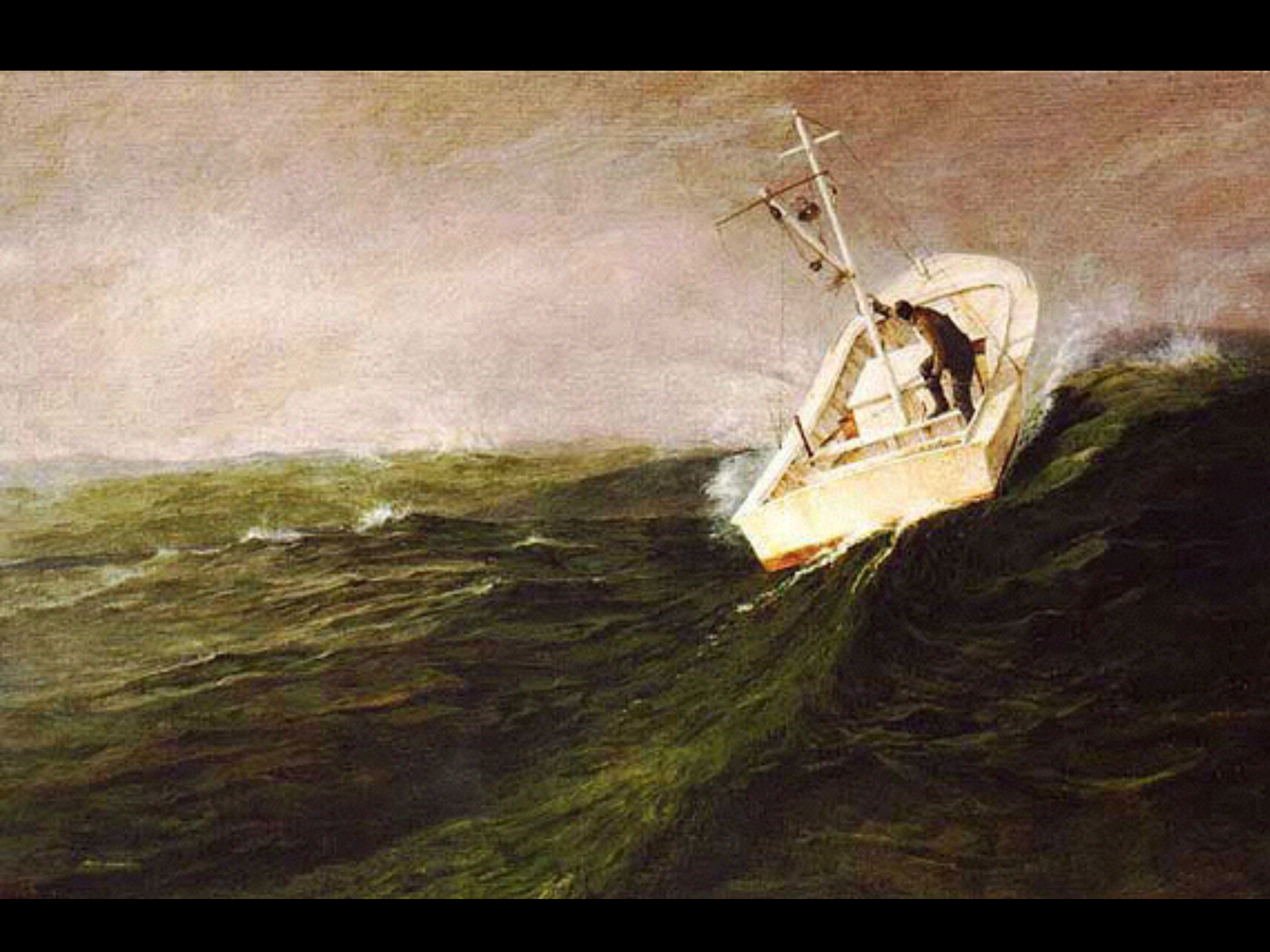




“Modality”

Representations of Possibility





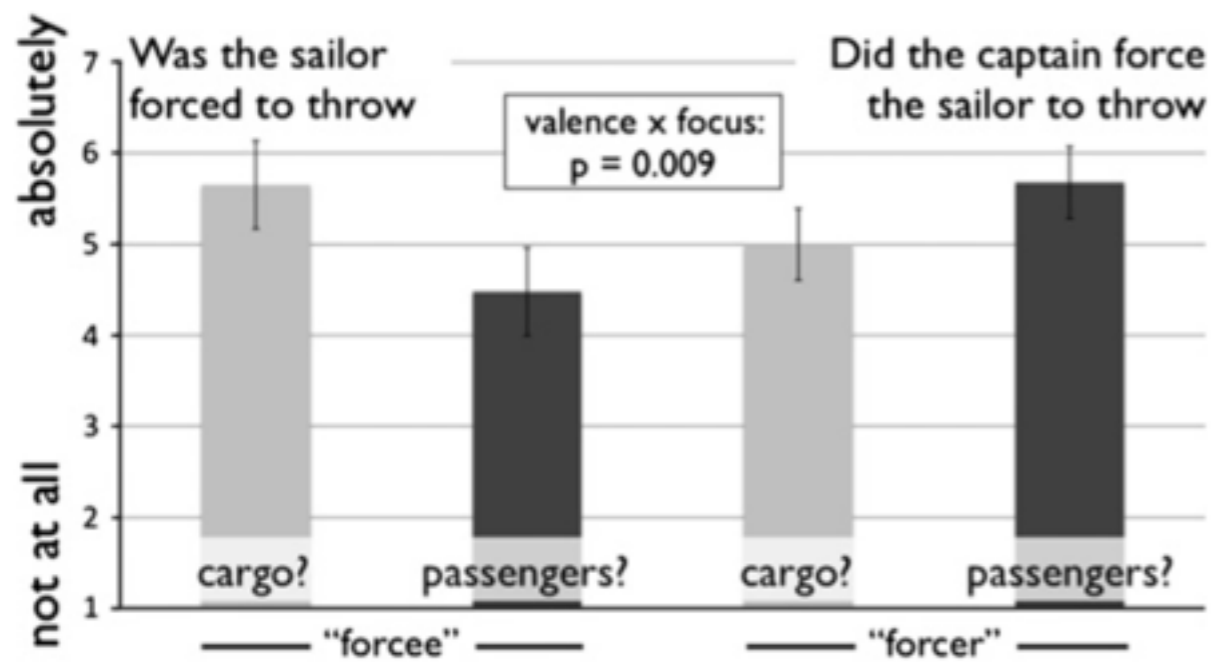


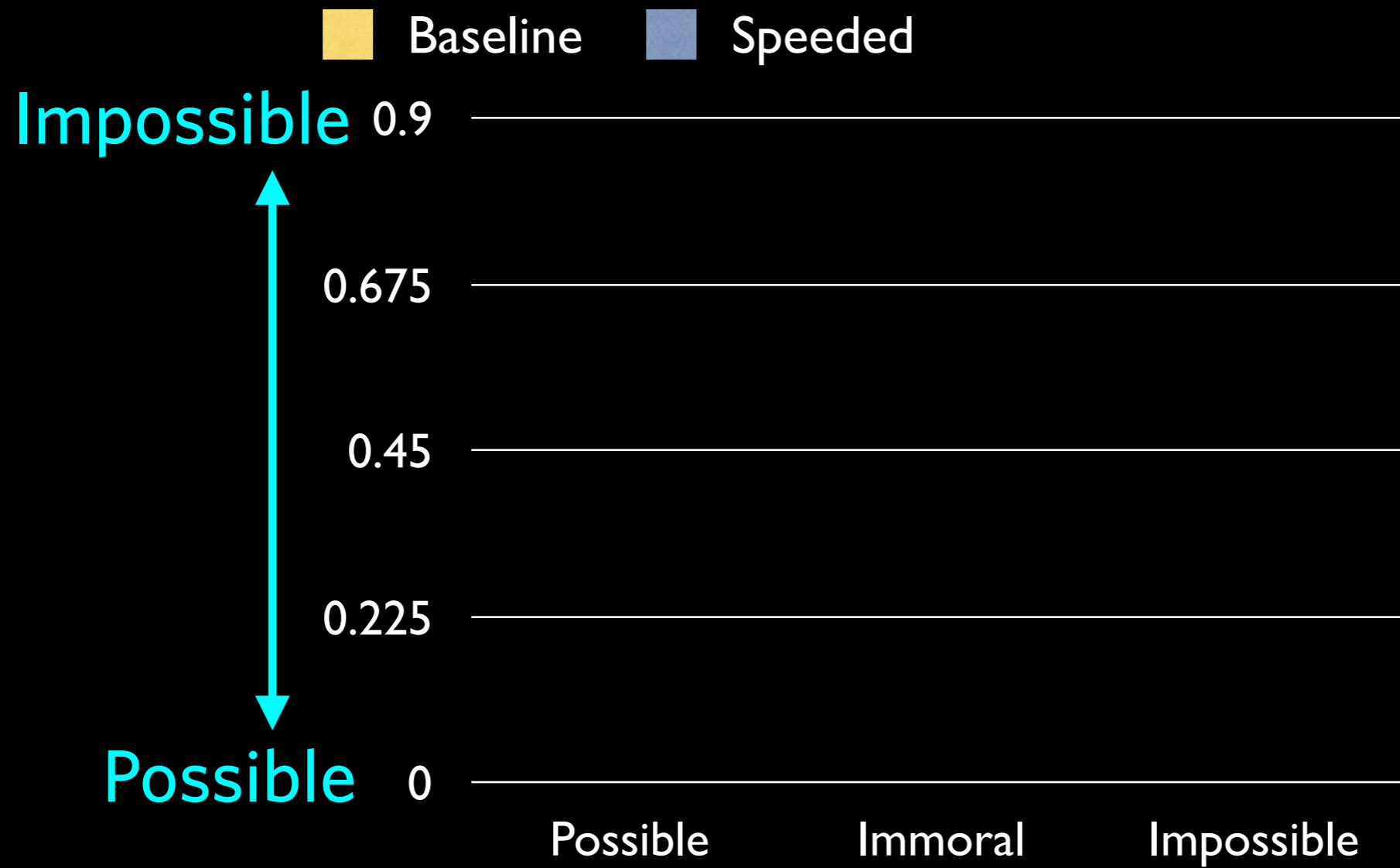
Fig. 1. Force judgments of the forcee (sailor, left) and the forcer (captain, right) for neutral actions (light bars) versus morally bad actions (dark bars). Error bars indicate standard error.



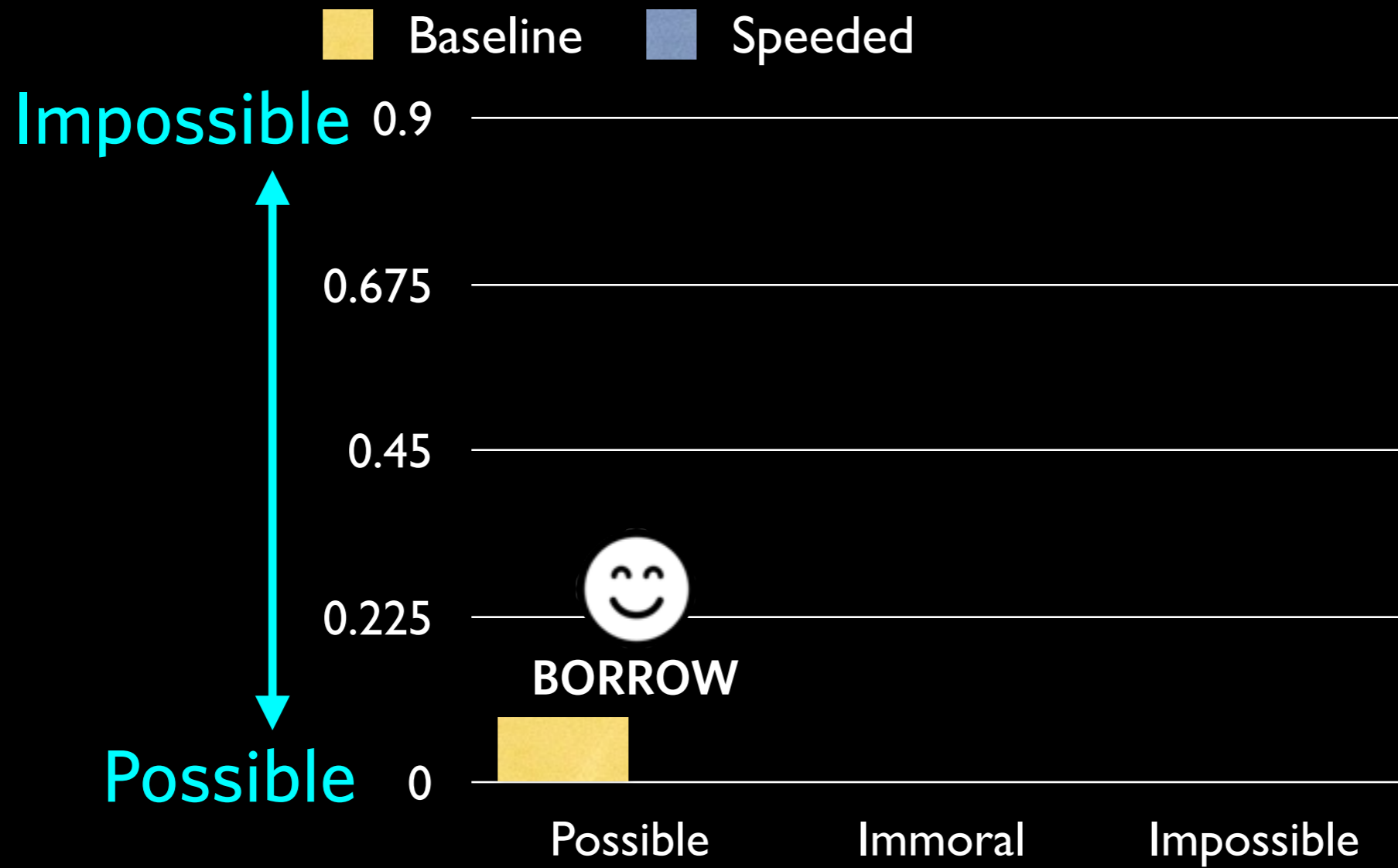


Practical Possibility

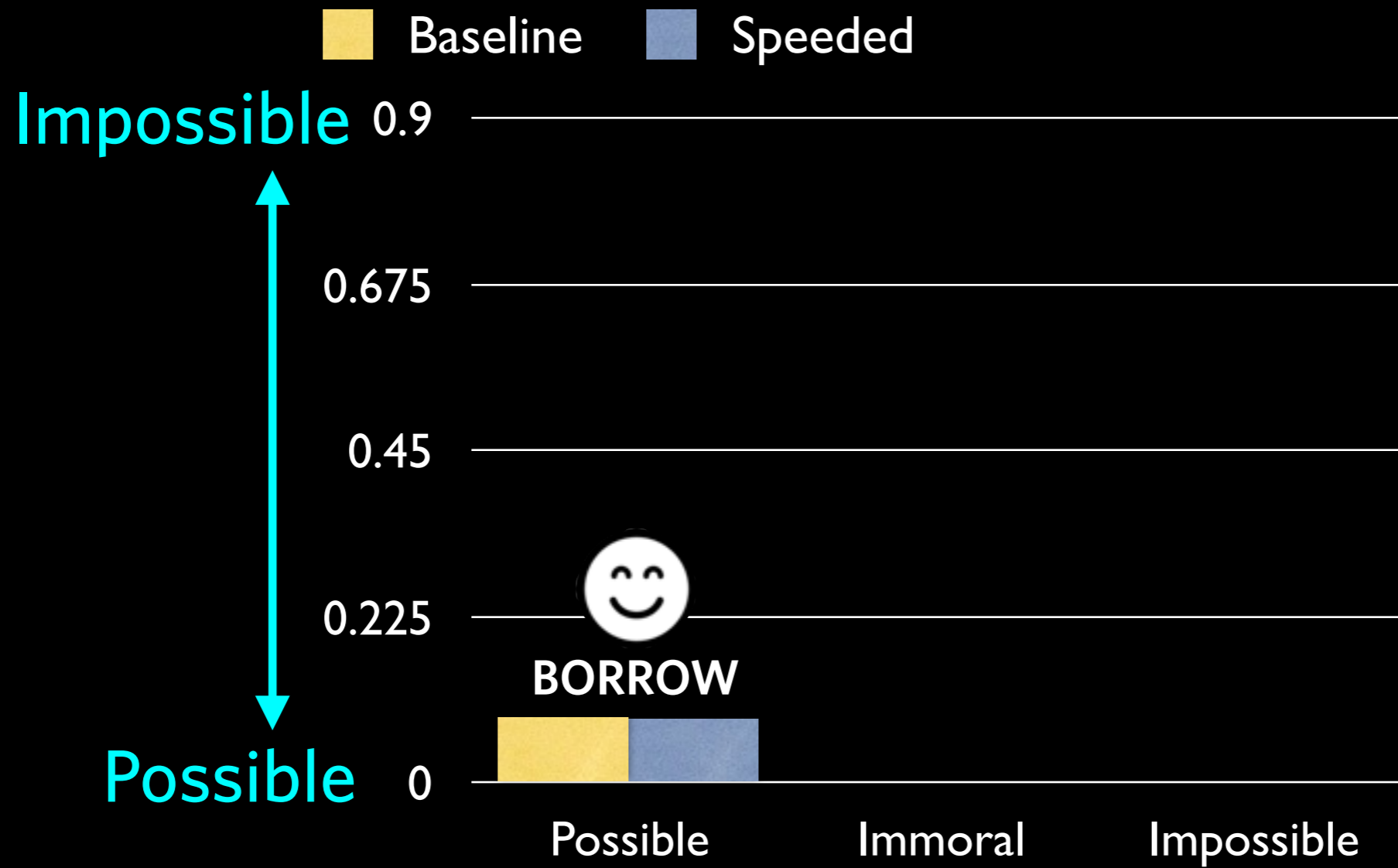
Practical Possibility



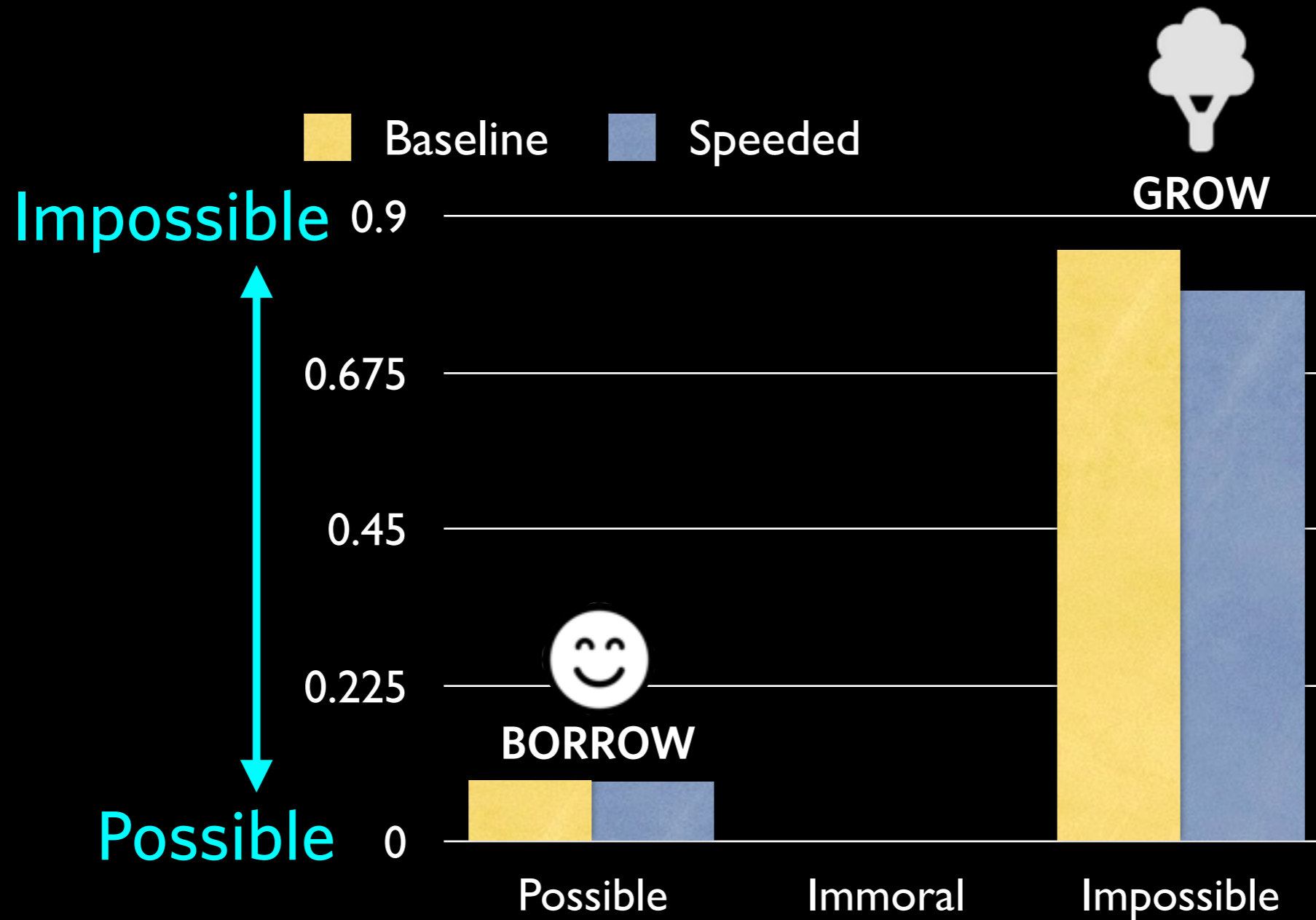
Practical Possibility



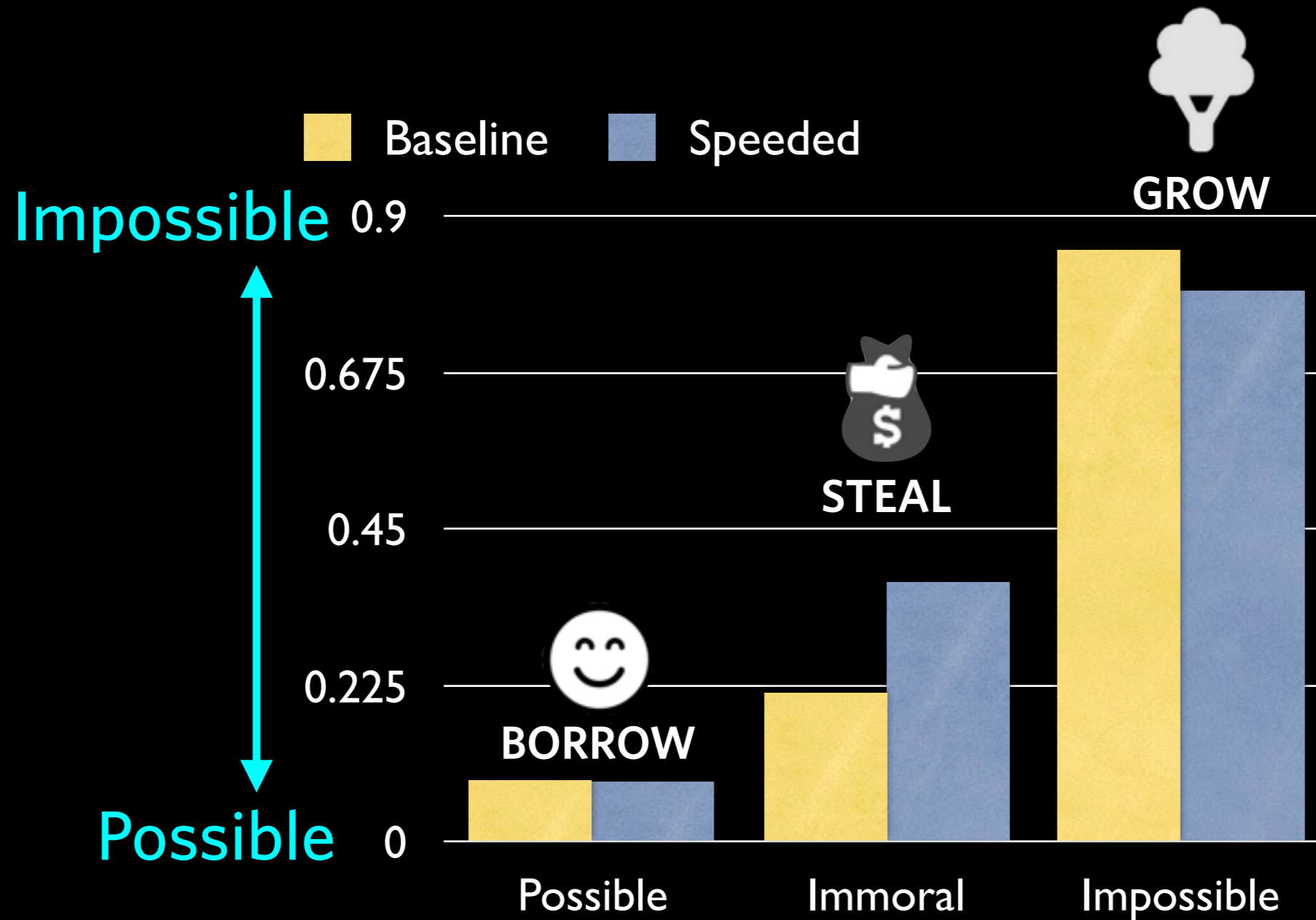
Practical Possibility

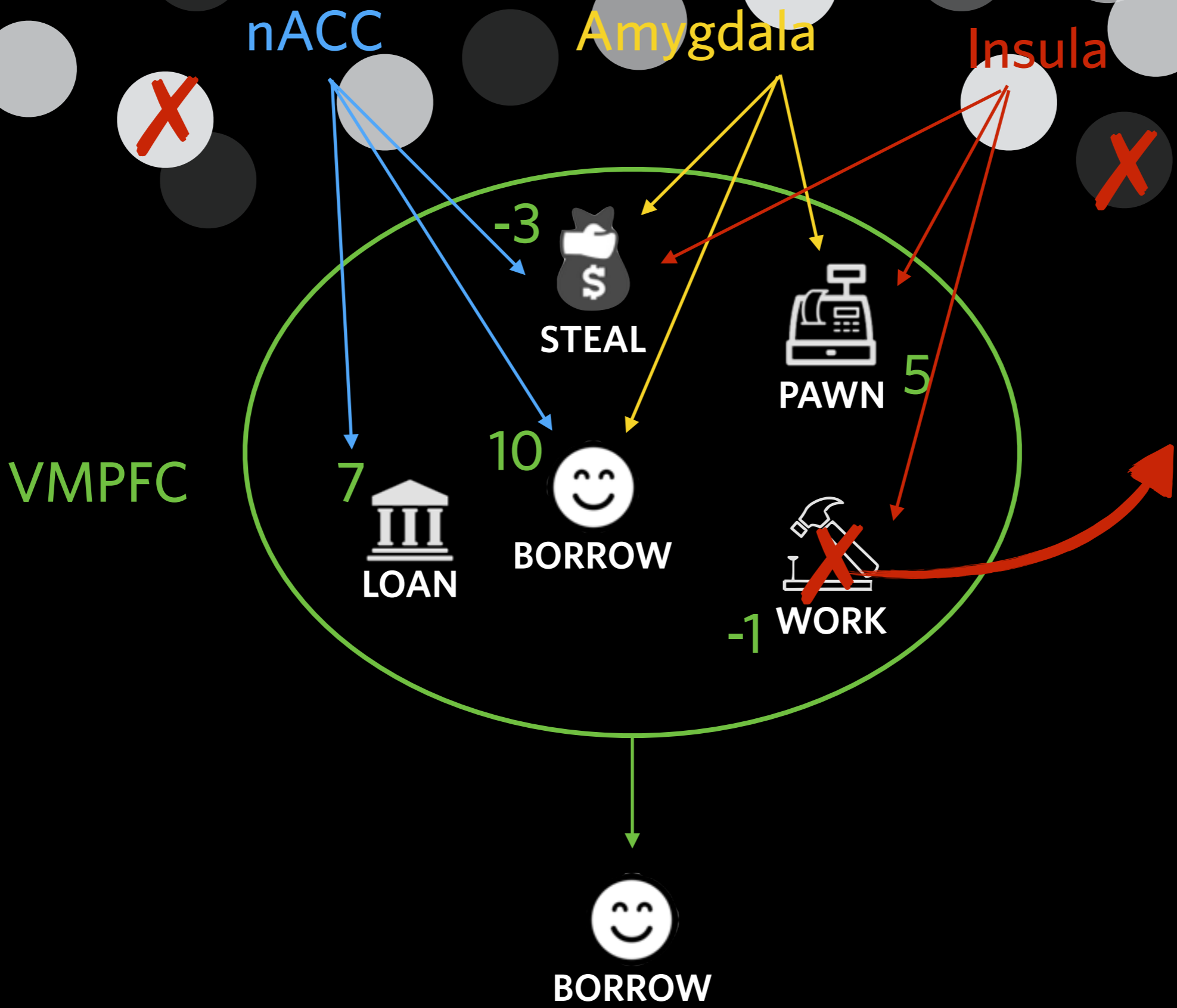


Practical Possibility



Practical Possibility





Neuroscience

Psychology

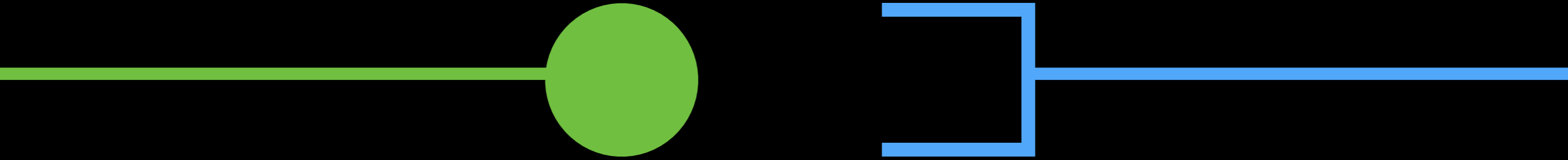


$$\text{\$} = \overset{\text{halo}}{\text{M}}$$

$$\text{\$} \neq \overset{\text{halo}}{\text{M}}$$

Neuroscience

Psychology



$$\text{\$} = \text{M}$$

$$\text{\$} \neq \text{M}$$

①

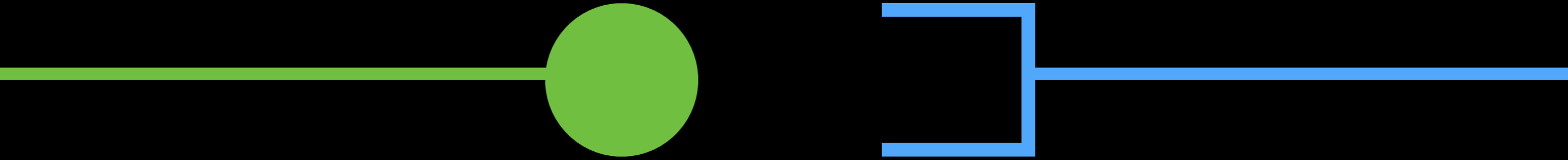
-300,000



STEAL

Neuroscience

Psychology



$$\text{\$} = \overset{\text{halo}}{\text{M}}$$

$$\text{\$} \neq \overset{\text{halo}}{\text{M}}$$

①

-300,000



STEAL

②

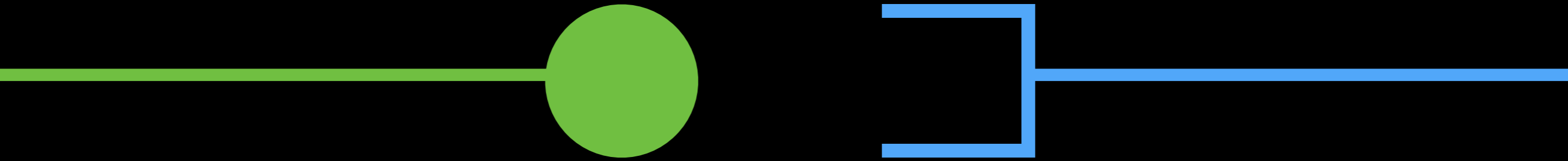
✓
Action



~~Result~~

Neuroscience

Psychology



$$\text{\$} = \overset{\text{halo}}{\text{M}}$$

$$\text{\$} \neq \overset{\text{halo}}{\text{M}}$$

①

-300,000



STEAL

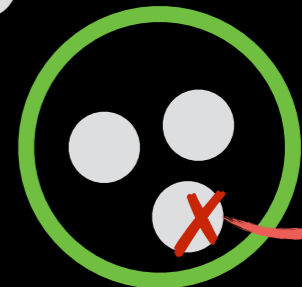
②

✓
Action



~~Result~~

③



Thanks

Ryan Miller
Justin Martin
Adam Morris
Jorie Koster-Hale

Jonathan Phillips
Wouter Kool
Josiah Nunziato
Matt Cashman

Keerthi Reddy
Morgan Henry
Michael Littman
James McGlashan

Joseph Austerweil
Mark Ho
Sam Gershman
Arunima Sarin



Moral Psychology Research Lab

Funding:



Best ideas due to:

Liane Young
Josh Greene
Joshua Knobe
Michael Littman
Sam Gershman
Dave Rand
& Many Others