# Lecture 4 - Dual Process

Jesse Hoey
School of Computer Science
University of Waterloo

September 28, 2021

Readings:

- Daniel Kahneman *Thinking, Fast and Slow*, 2011, chapters 1-3
- Simon, Strenstrom, and Read. The Coherence Effect: Blending Hot and Cold Cognitions, *Journal of Personality and Social Psychology*, Vol 109, number 3, p369-394, 2015.

De vulpe et catto.

**Figure 1**

$$17 \times 24 = ?$$

- **System I**: Operates automatically and quickly, without voluntary control
- **System II**: effortful mental activity, complex calculations
- **System II**: subjective feeling of agency, choice and concentration
- When system I "runs into trouble", it calls upon System II
- We identify with System II, however, as it is conscious



THINKING, FAST AND SLOW

DANIEL KAHNEMAN

WINNER OF THE NOBEL PRIZE IN ECONOMICS
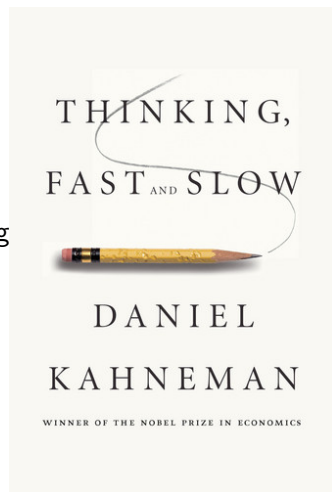
# System I and System II

## System I

- Orient to the source of a sound
- Complete the phrase "bread and ..."
- Detect hostility in a voice
- Answer $2+2=?$
- Determine if someone you see in a hospital is a doctor or a nurse
- Gorilla Experiment: `https://www.youtube.com/watch?v=vJG698U2Mvo`

## System II

- Focus attention on one person at a party
- maintain a faster walking speed than is natural
- count the number of "a"s in a page
- Answer $178+341=?$
- remember a new phone number
- construct a logical proof

- A bat and a ball cost $1.10
- The bat costs one dollar more than the ball
- How much does the ball cost?

# System I biases

- System I makes "mistakes"
- Biases perceptions and actions that are fast, but sometimes wrong
- can these be overcome? Do we want to overcome them?
- System II too slow and inefficient to be practical everyday
- Tasks that require a lot of "System II" thinking are effortful, tiring, and people will avoid them if they can.
- This has a profound influence on humans' ability to think "rationally"
- Kahneman and Behavioural Economists have noted that this also has a profound effect on economics.

- Division of labour: highly efficient
- minimizes effort and optimizes performance



Your first task is to go down both columns, calling out whether each word is printed in lowercase or in uppercase. When you are done with the first task, go down both columns again, saying whether each word is printed to the left or to the right of center by saying (or whispering to yourself) "LEFT" or. "RIGHT."

| | | | |
|---|---|---|---|
| LEFT | | upper | |
| | left | lower | |
| right | | | LOWER |
| RIGHT | | upper | |
| | RIGHT | UPPER | |
| | left | | lower |
| LEFT | | | LOWER |
| | right | | upper |

**Figure 2**

Figure 1. The Müller-Lyer illusion. The lines labeled "a" and "b" are the same length. Many subjects perceive line "b" as longer than line "a".
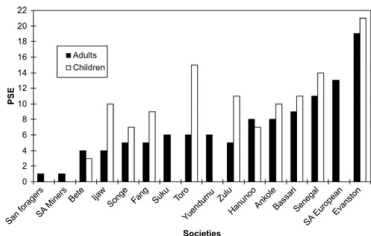


Figure 2. Müller-Lyer results for Segall et al.'s (1966) cross-cultural project. PSE (point of subjective equality) is the percentage that segment a must be longer than b before subjects perceived the segments as equal in length. Children were sampled in the 5-to-11 age range.

- People handle this illusion differently
- Americans are the most susceptible
- Some people don't perceive the illusion at all
- WEIRD people can't not perceive it (system I), even when they know (system II) that the lines are the same length
- Think about how people can't not perceive stereotypes

But it turns out Pupils are indications of <mark>Mental Effort</mark>, not necessarily attraction

# Flow

- Mihaly Csikszentmihalyi
- <mark>Flow</mark> 1990
- "optimal" balance between focussed attention and conscious control
- tasks are just slightly more challenging that capacity
- many aspects are "taken over" by System I
- more resources to handle smaller deviations



If you were red boots, would you pause to think it all through?

# Connectionist Dual Process models

- PCS: Parallel Constraint Satisfaction
- PCS: graph with nodes, weights on edges, and a learning rule
- PCS are initialized, then iterated weight updates until convergence
- converged solution may be consistent or inconsistent, coherent or incoherent
- $C \geq \theta$: more consistency than a threshold, use PCS "option"
- $C < \theta$: pass the buck to the DC
- DC: Deliberative constructions
- How is $\theta$ defined? How is the network initialized?
- Glöckner and Betsch. Modeling option and strategy choices with connectionist networks: `journal.sjdm.org/bn3.pdf`
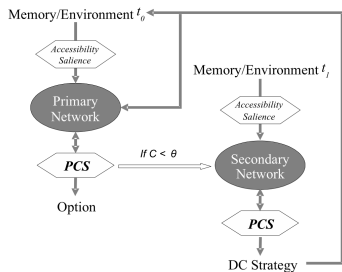
Figure 3: An integrative PCS model for the selection of options (primary network) and deliberate construction strategies (secondary network).

# Conectionist Dual Process Models

- Simon Stenstrom and Read. The Coherence Effect: Blending Hot and Cold Cognitions
  `http://dx.doi.org/10.1037/pspa0000029`
- The story of Debbie Miller
- participants played "Judicial Officer"
- Debbie is a student accused of cheating
- ambiguous evidence for and against her cheating
- sympathy aroused in half participants: death of younger brother
- finding: sympathy for Debbie decreases culpability (guilt)

# Guilt and Innocence "Facts"

- evidence from interviews with witnesses
- intricate and ambiguous
- facts "suggest" that she cheated (or not)
- guilty:
  - ▶ Debbie "looked like" she was hiding something
  - ▶ Proctor might have seen her stick something in her pocket
  - ▶ a student saw her copying
- not guilty
  - ▶ Debbie was a good student
  - ▶ Debbie committed to study
  - ▶ Proctor had a conflict?
  - ▶ Debbie subject of envy by students

# Results: human participants

Table 5
*Dependent Measures Grouped by Condition (Study 3)*

| | Control condition (n = 206) | | Tragic-information condition (n = 188) | | |
|---|---|---|---|---|---|
| | Frequency | | Frequency | $\chi^2$ | |
| Cold cognitions | | | | | |
| Decision (Debbie cheated) | 48.1% | | 29.8% | 13.75*** | |
| | Mean (*SD*) | | Mean (*SD*) | F | $\eta^2$ |
| Likelihood of cheating | 50.89 (31.57) | | 38.07 (29.10) | 17.45*** | .04 |
| Guilt facts | −1.74 (1.96) | | −2.23 (1.79) | 6.57* | .02 |
| Innocence facts | .06 (2.16) | | .84 (2.03) | 13.83*** | .03 |
| Hot cognitions | | | | | |
| Valence (Debbie | | | | | |
| disciplined) | 5.26 (2.32) | | 4.63 (2.26) | 7.37** | .02 |
| Motivation (university win) | 4.96 (2.77) | | 4.32 (2.65) | 7.05** | .02 |
| Liking Debbie | −.13 (.96) | | .14 (.96) | 7.85* | .02 |
| Anger toward Debbie | 2.51 (1.53) | | 2.07 (1.34) | 9.05** | .02 |
| Sympathy toward Debbie | 4.11 (1.70) | | 4.46 (1.60) | 4.43* | .01 |
| Fear of Debbie | 1.58 (1.10) | | 1.42 (.90) | 2.48 | .01 |
| Sadness toward Debbie | 3.26 (1.59) | | 3.38 (1.46) | .58 | .001 |

* $p < .05.$   ** $p < .01.$   *** $p < .001.$

# Debbie as a PCS

- **Valence** : **affect heuristic** e.g. "good people do good things"
- **sympathy** for Debbie is influential
- dashed arrows: **inihibitory**
- solid arrows: **excitatory**
- "brother killed" is evidence
- "special nodes" for **starting activations**
- numbers show **weights after convergence**
- weights correlate with human data **0.68**
- Sympathy, liking, guilt: **specific** emotions - can we be more general? Answer: Values
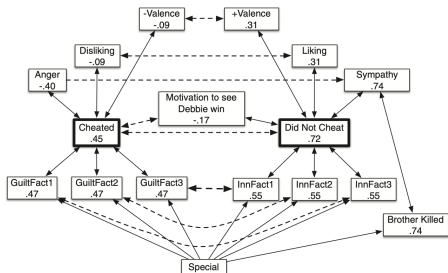
# What are values?

Definitions (Oxford Concise):

| definitions | worth , desirability , utility , standards | worth as estimated judgement of value / importance in life |
|---|---|---|
| usually | rational | arational |
| confusing labels | "cognitive" | "emotional" |
| values of the | individual | group |
| key method or algorithm | reinforcement learning | Affective Computing |
| implementation of values | normative behaviours , intrinsic reward symbolic | Dimensional emotional models subsymbolic |

Much of cognitive science and artificial intelligence makes two errors with respect to this definition

- assumes the group values can be encoded in individual values (e.g. through intrinsic rewards for "altruism")
- assumes that principles and standards encode group values, while emotions encode individual values (i.e. "passions").
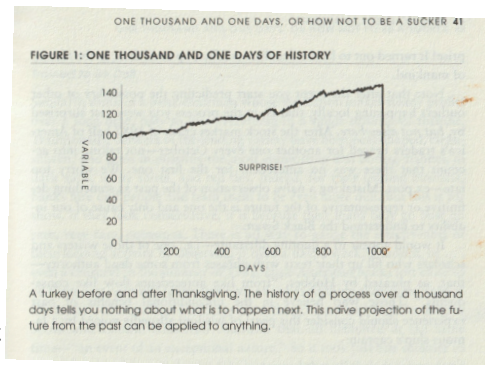
Let us consider the evidence

# Enigma of Reason

key ides

- In western philosophy since Descartes, reason is all that matters
- Reason comes first
- anything embodied or emotional is disruptive to pure reason
- clearly not how the human brain works
- emotional centers rule, make decisions, motivate action, and come first and quickly.
- reason is usually applied after the fact
- reason applied after a decision has been made or taken
- reason used in order to justify the action in the context.
- Humans fall prey to the narrative fallacy - stick with the story you know.

Hugo Mercier and Dan Sperber. The Enigma of Reason. 2017

# Narratives and Black Swans

- Thanksgiving turkey's last 1001 days
- last day is the day the turkey is killed
- from the Turkey's perspective , the narrative is positive
- Turkey's predictive model will accurately predict the events of day 1001
- but the predictions turn out to be very wrong
- humans will "reason" that it wasn't their fault
- the Bob Rubin story
- Black Swans can be postive (e.g. fame)
- Nassim Nicholas Taleb The Black Swan , 2007

ONE THOUSAND AND ONE DAYS, OR HOW NOT TO BE A SUCKER 41

FIGURE 1: ONE THOUSAND AND ONE DAYS OF HISTORY



A turkey before and after Thanksgiving. The history of a process over a thousand days tells you nothing about what is to happen next. This naïve projection of the future from the past can be applied to anything.

Much of cognitive science and artificial intelligence makes two errors with respect to this definition

- assumes the group values can be encoded in individual values (e.g. through intrinsic rewards for "altruism")
- assumes that principles and standards encode group values, while emotions encode individual values (i.e. "passions").

What did the evidence say?

# Values cause confusion

In fact, it is the opposite.

- principles and standards encode individual values

- emotions encode group values

- "Fast thinking" is what occurs quickly, subconsiously, and so-called "emotionally." Fast thinking occurs without explicit reason, and highly motivates action through emotion. Without cognitive checks, these actions may proceed, and are expected to encode group cooperative dynamics.

- "Slow thinking" is reasoned, thoughtful, rational, and only is called into play if necessary (i.e. if the fast system is apprised to be failing, or if explicitly called on to generate reason

- in the rest of this lecture, we focus on how to define and measure values

**FIGURE 1.2**

**A Sample SYMLOG Consulting Group Rating Form,**
**Individual and Organizational Values**

Reflect on the work-related experiences you have had with the person you are about to rate. Whether you have worked closely with this person or have had only limited contact, keep your impressions of him or her in mind as you answer the question below. Not all parts of a descriptive items may seem to go together. *If any part applies, then use it as your guide.*

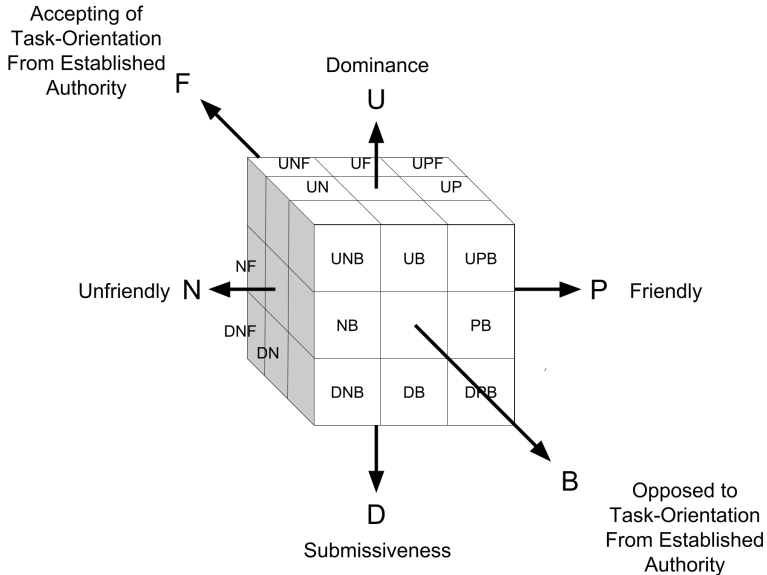In general, what kinds of values does this person show in his or her behavior?

**Descriptive Items—Individual and Organizational Values**

| | Rarely | Sometimes | Often |
|---|---|---|---|
| 1 Individual financial success, personal prominence and power . . | [ ] | [ ] | [ ] |
| 2 Popularity and social success, being liked and admired . . . . . . . | [ ] | [ ] | [ ] |
| 3 Active teamwork toward common goals, organizational unity . . | [ ] | [ ] | [ ] |
| 4 Efficiency, strong impartial management . . . . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 5 Active reinforcement of authority, rules, and regulations . . . . . . | [ ] | [ ] | [ ] |
| 6 Tough-minded, self-oriented assertiveness . . . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 7 Rugged, self-oriented individualism, resistance to authority . . . | [ ] | [ ] | [ ] |
| 8 Having a good time, releasing tension, relaxing control . . . . . . . | [ ] | [ ] | [ ] |
| 9 Protecting less able members, providing help when needed . . . . | [ ] | [ ] | [ ] |
| 10 Equality, democratic participation in decision making . . . . . . . . | [ ] | [ ] | [ ] |
| 11 Responsible idealism, collaborative work . . . . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 12 Conservative, established, "correct" ways of doing things . . . . . | [ ] | [ ] | [ ] |
| 13 Restraining individual desires for organizational goals . . . . . . . | [ ] | [ ] | [ ] |
| 14 Self-protection, self-interest first, self-sufficiency . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 15 Rejection of established procedures, rejection of conformity . . . | [ ] | [ ] | [ ] |
| 16 Change to new procedures, different values, creativity . . . . . . . | [ ] | [ ] | [ ] |
| 17 Friendship, mutual pleasure, recreation . . . . . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 18 Trust in the goodness of others . . . . . . . . . . . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 19 Dedication, faithfulness, loyalty to the organization . . . . . . . . . | [ ] | [ ] | [ ] |
| 20 Obedience to the chain of command, complying with authority | [ ] | [ ] | [ ] |
| 21 Self-sacrifice if necessary to reach organizational goals . . . . . . . | [ ] | [ ] | [ ] |
| 22 Passive rejection of popularity, going it alone . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 23 Admission of failure, withdrawal of effort . . . . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 24 Passive non-cooperation with authority . . . . . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 25 Quiet contentment, taking it easy . . . . . . . . . . . . . . . . . . . . . . . | [ ] | [ ] | [ ] |
| 26 Giving up personal needs and desires, passivity . . . . . . . . . . . . | [ ] | [ ] | [ ] |

*Source:* © 1997 SYMLOG Consulting Group, San Diego, CA. Reprinted by permission.

# Values Space (Bales)



Bales, R. F. (1999). Social interaction systems: theory and measurement. Transaction Publishers.

Used to extract values from behaviours

| IPA Category Number | IPA Category Name | Sample Behaviors | E | P | A |
|---|---|---|---|---|---|
| 1 | Shows solidarity | help, compliment, gratify | 1.78 | 1.29 | .21 |
| 2 | Shows tension release | josh, laugh with, cheer | 1.48 | .91 | 1.12 |
| 3 | Agrees | agree with, understand, accommodate | 1.60 | .78 | .01 |
| 4 | Gives suggestion | encourage, cue, coach | 1.28 | 1.18 | .25 |
| 5 | Gives opinion | evaluate, analyze, entreat | .16 | .59 | −.02 |
| 6 | Gives orientation | inform, educate, explain | 1.68 | 1.62 | −.14 |
| 7 | Asks for orientation | quiz, question, ask about | .50 | .62 | .45 |
| 8 | Asks for opinion | consult, prompt, query | .48 | .74 | .16 |
| 9 | Asks for suggestion | entreat, ask, beseech | .30 | .24 | .09 |
| 10 | Disagrees | disagree with, ignore, hinder | −1.00 | .35 | .45 |
| 11 | Shows tension | fear, cajole, evade | −.89 | −.16 | .35 |
| 12 | Shows antagonism | argue with, deride, defy | −.82 | .71 | 1.32 |

Bales, R. F. (1999). Social interaction systems: theory and measurement. Transaction Publishers.

Heise, D. R. (2013). Modeling interactions in small groups. Social Psychology Quarterly, 76(1), 52–72.

# Overlay



**FIGURE 1.1**
**An Example Field Diagram**

VALUES ON ACCEPTING TASK-ORIENTATION OF ESTABLISHED AUTHORITY
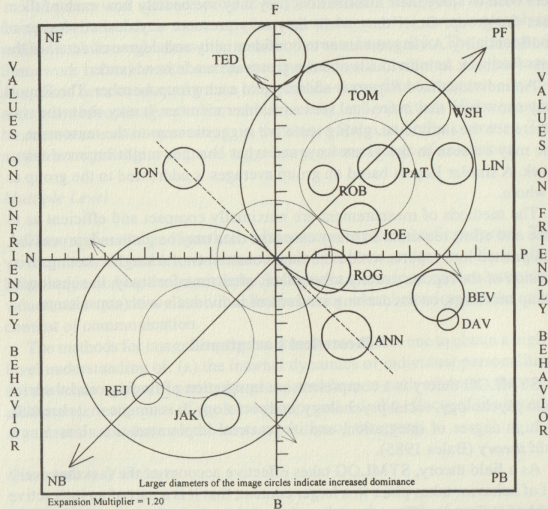
VALUES ON OPPOSING TASK-ORIENTATION OF ESTABLISHED AUTHORITY

*Source:* © 1997 SYMLOG Consulting Group, San Diego, CA. Reprinted by permission.
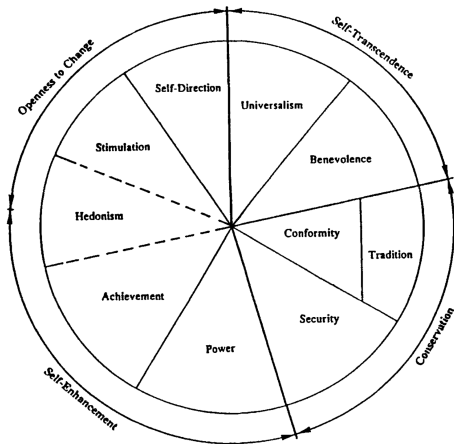
# Shalom Schwartz 1992/1994



**Fig. 1.** Theoretical model of relations among motivational types of values, higher order value types, and bipolar value dimensions (adapted, with permission, from Schwartz, 1992).

Reading: Shalom Schwartz "Are There Universal Aspects in the Srtucture and Contents of Human Values?" Journal of Social Issues, Vol 50, No. 4, 1994 pp. 19-45

Antonio Damasio *Descartes' Error* Chapter 3

- Elliott `Frontal damage` (tumor) - like Phineas Gauge
- Again, `could not make decisions` in real life
- Especially with respect to personal or social matters
- Felt `emotionless`, but `knew` he used to feel emotions

Antonio Damasio *Descartes' Error* Chapter 3

- Elliott Passed a massive battery of intellectual capacity tests: perception, memory, learning, language, math all intact
- Including ethical and moral decisions, standard "personality tests":
  - ▶ Generation of options for action
  - ▶ awareness of consequences
  - ▶ Means-ends problem solving for social goals
  - ▶ predict social consequences
  - ▶ standard issue moral judgment
- This caused problems for analysis as there seemed to be no standard tests to explain Elliott's problems

Antonio Damasio *Descartes' Error* Chapter 3

- Elliott's condition: *to know but not to feel*
- Lack of decision making ability was *not due to*:
  - ▶ lack of social knowledge
  - ▶ deficient access to knowledge
  - ▶ impairment of reasoning
  - ▶ defect in attention or working memory
- Defect in decision making ability happened *late*
- Elliott could generate choices, but could not choose one

# Emotions as Somatic Markers

- Animals are faced with a vast space of possible decisions
- Resources are bounded
- Time is limited
- Must somehow only evaluate the "good" actions
- Somatic markers indicate which ones these are.
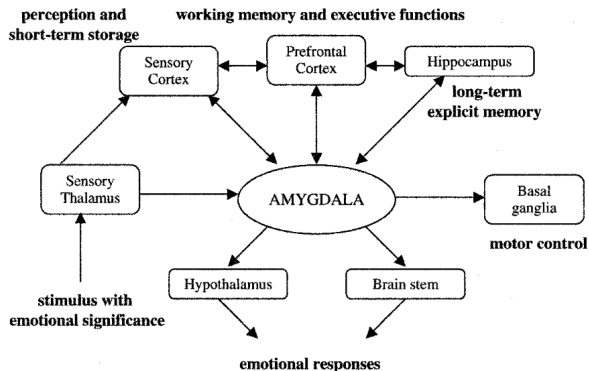- Like an "oracle", but a learned oracle (somehow)

Fig. 2. The amygdala: hub in the wheel of emotion (based on LeDoux, 1996, Chapters 6 and 9).

From Zhu & Thagard "Emotion and Action".
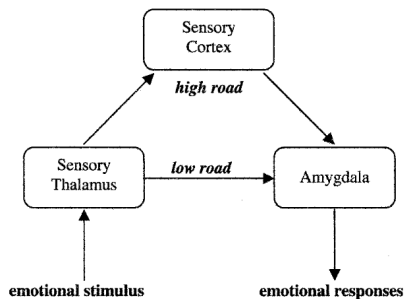Philosophical Psychology Vol 15 No 1, 2002.

Fig. 3. Two separate pathways from sensory stimulus to emotional responses (adapted from LeDoux 1996, p. 164).

From Zhu & Thagard "Emotion and Action".
Philosophical Psychology Vol 15 No 1, 2002.

# Cognitive Aspects of Emotions

| Damasio: | Kahneman: |
|---|---|
| <ul><li>Primary emotions: fast, pre-cognitive responses</li><li>Secondary emotions: slower, cognitive responses</li></ul> | <ul><li>System I: default, fast, heuristic reasoning</li><li>System II: rational, slow, thought</li></ul> |

- There is not a 1-1 correspondence between these, and the lines are blurry
- Primary/System I is often "wrong" but airs on the side of caution
- Risk-aversion is a unifying element
- Evolutionarily makes sense:
  - ▶ risks lead to death,
  - ▶ the rewards (for the survival of the species) outweigh the benefits for the individual

# Review of Dual Purpose models

Key ideas:

- System 1 (heuristic) processes cue
  <mark>default intuitive judgments</mark>
- System 2 is called upon to validate choices or make (bad) predictions
- Mercier and Sperber, Kahneman, etc – same argument
- Incorrect but often assumed: System 1 recognizes a context and retrieves a default program that was previously stored by system 2.
- This follows Simon labeling emotions as encoding rational heuristics, or being disruptive to rational thought
- Jonathan St. B. T. Evans Dual-Processing Accounts of Reasoning, Judgment and Social Cognition
- Andreas Glöckner and Cilia Witteman. Beyond dual-process models

# Next

- Shehroz Khan invited lecture (assistive technology and surveillance)
- Affective Computing