BCS153: Exam 2 Study Guide

Philosophy and Mind

- Video: "They're made out of meat"
  →we (meat) think a/b computations/machines (biological & mechanical)
- Mind/Body Problem
- →How do minds (consciousness, beliefs, etc.) arise out of a physical system?
- Terminology
  →Qualia: subjective experience of something
  ex. what qualia do you get from grapefruit?
  Ex. describing the way you experience a color
- How minds relate to Brains

Substance Dualism	*2 distinct things (mental/physical), soul/brain *Descartes (Cartesian Dualism) -no behaviors emerging from soul -says that bodies are machines, lots of biological factors take place (ex. cell division, digestion) but what a/b sensations, imagination? -said that soul interacted w/ body through pineal gland (was wrong lol)
Modern Dualism	*soul=hypothesis for resolving mind-body problem *most scientist/philosophers are NOT substance dualism

- Interaction Problem
  *How do physical/non-physical substances interact, and why?
  →How does a soul interact w/ our physical perception?
  *No matter how the system works, the physical/non-physical body must interact
- Physical Coupling Problem
  *Mental events really do seem to be linked to physical changes
  Ex. Clear degradation in behavior/thinking w/ physical changes (Alzheimer's, corpus callosum split, stroke=brain damage)
- Physical→Mental Link
  *Physical Causality→Penfield (neurosurgeon)
  -he stimulated the cortex electrically (access to stream of consciousness)
  -while patients were still awake, he stimulated regions of the brain and the patients reported different corresponding sensations
  →Strong link to soul b/c small physical change=large sensation/feeling (modern dual)
- Alien-Hand Syndrome
  *Alien-Hand Syndrome is seen in split-brain patients (split corpus callosum)
  -this is done to treat sever epilepsy from spreading to larger regions of the brain
  *Mental processes occurring at the same time that are uncoordinated

ex. one hand buttoning dress, other hand unbuttoning dress/one hand opening drawer, other hand closing drawer
→Challenge w. substance realism b/c soul is being manipulated by physical changes

- Property Dualism: AGAINST PHYSICALISM
*There is only one kind of thing (physical), but two kinds of properties (physical/mental)
-these properties are not the same thing, they are descriptions of physical matter
Ex. Jackson 1982
*Mary, a scientist, investigates the world form a black/white room
(She understands physical process of perceiving color)
→If you remove Mary from the black/white room and show her a red apple, she gains the qualia a/b the experience of seeing red that she did not have before in the black/white room
→The property (mental) from seeing red cannot be found in the physical
*so, she doesn't know the experience of seeing red until she actually sees red (because there's more to it than just understanding the neuronal explanation of perceiving red)
- Physicalism: (ex. behaviorism)
*all there is, is physical

Behaviorism	mental states= tendencies to behave certain ways
Identity Theory	Identity Theory: mental states= the same as physical states *values importance of physical device *mental state depends on the specific physical device (specific brain), mental event=physical event ex. Penfield stimulates area that gives you smell percept of cinnamon →Identity theory says that perceiving the smell of cinnamon= identical w/ activity in that brain area CRITIQUES: *multiple realizability: the same mental state can be realized in different systems *Neuro-prosthetics Ex 1: monkey hand- plays game with controller, mental process is recorded -controller is removed and the monkey still performs only through the use of his motor cortex Ex 2: Rat's Hippocampus -taught to remember to press either left or right lever -when hippocampus is simulated, they perform more accurately
Functionalism	Functionalism: mental states= characterized by the relationship to other mental states ex. parts of brain can be replaced, and this changes nothing a/b the mental process

- Philosophical Zombie: human w/o qualia (no individual experience of perception)
- Functionalism
  *make up does not matter
  -function/role is enough to understand
  →In favor: implementation
  →Against/Critique:
  Ex 1: Ned Block's "China Brain"
  *take all of your neurons, have people in China act as units that mimic neuronal functional connections, this changes process (subjectivity)
  Ex 2: Inverted Spectrum Argument
  *your red=my blue
  →Different qualia
  *Functional role/relationship isn't the whole story (subjectivity/perception)

Bayes

- Ways of Knowing:

Deduction	Induction**
*start from general facts and reason to something specific ex. all men have 2 legs, Bill is a man →Bill has 2 legs *Think Sherlock!	*reason from particulars to general/abstract ex. if 15/15 men have 2 legs, what's the probability that all men have 2 legs? →1! *But this can be disproven=defeasibility

- Bayes Thought Process
  Ex 1: box w/ mystery things in it
  →We have data/specific examples, and we reason through hypotheses toward an answer
  *we keep updating our belief as we reason through new data
  Ex 2: Coin Flipping
  *How many heads in a row would it take to convince you that the coin was a two-headed coin?
  H1: fair coin (t/h)
  H2: two-headed (h/h)
  H3: two-tailed (t/t)
  *Bayes rule: is all about how you update your prior belief based on new experiences/data
  -you start ranking your hypotheses as you observe more data, and start eliminating the ones that are ruled impossible (for example, in the coin flipping task, if you flip the coin and get T, you can rule out double-headed)
- Bayes Rule Formula
  *initial belief + new evidence= new/improved belief

$$P(H|D) = (P(H)*P(D|H))/(P(D))$$

  P(H|D)- posterior belief

P(H)- prior belief
P(D|H)- likelihood
P(D)- probability of data "evidence" = sum of all P(H)*P(D|H)
→4 Steps in Computing Bayes

1. Compute prior for each hypothesis: P(H) (normally given)
2. Compute likelihood of data under each hypothesis: P(D|H)
3. Multiply prior and likelihood: P(H)*P(D|H)
4. Re-normalize

Things to remember
*P(H) sums to 1
*P(H|D) sums to 1

Hypotheses	**P(H)**	**P(D|H)**	**P(H)*P(D|H)**	**P(H|D)**
H1	A	B	AB	(AB)/((AB)+(1-A)C)
H2	1-A	C	(1-A)C	((1-A)C)/((AB)+(1-A)C)

- Bayesian Techniques
*Provide the logic for describing what beliefs you SHOULD have, given the data that you see (relative assumptions)
→Bayes is VERY IMPORTANT to…
*Cognitive Science/Neuroscience
*Statistical Analysis
*Machine Learning
- Cognitive Science

Hypotheses (H)	*beliefs a/b the world
Probabilities	*strength of belief that each hypothesis is true
Priors P(H)	*assumptions a/b the world (before data)
Likelihoods P(D|H)	*your assumptions a/b how the data you see was generated by each hypothesis

- Tanenbaum: (2000)
→The Number Game
*Black box accepts certain numbers (for example, all even numbers)
-participants are shown set of some numbers the box accepts [2, 4, 10, 28], without being told the rule
-then they are asked if the box would accept any numbers in a certain set [6, 7]
→There is no right answer, but we still feel like some numbers are better to be accepted by the box then others (we form structure/rules to justify why one number would be accepted over another)
→Bayesian Inference in The Number Game

H1: Even #s= 1/50 (1-100)
H2: All numbers= 1/100 (1-100)
H3: Powers of 2= 1/9 (2-64)
*The Size Principal: preference for smaller sets

- Bayesian Occam's Razor
  *prefer simpler explanations (starting off with a high P(H)), strong prior
- The Power of Bayes
  →Comes from ability to build interesting structures into hypothesis space
  *structure, nodes, connections= relations b/w set of hypotheses
  →Structure Forms: (observing pair-wise interactions)
  -partition
  -chain
  -order
  -ring
  -hierarchy
  -tree
  -grid
  -cylinder
  *Inference: developmental models expand w/ matching features
- Base-Rate Neglect: not considering the priors
  *case where people don't perform the correct Bayesian statistics
- Bayes & Rational Analysis
  →Rational Analysis: explain cognitive system based on something that makes sense
  ex. Bayes is the right way to solve problems, so it is easy to compare processes to Bayesian Theory
  →Inference in Memory:
  *Do memory systems just encode/decode information, or is there inference involved?
  Ex. Remembering size of fruit, asked to recall
  -some participants might recall the size of the fruit based on what they were shown compared with their prior experience/belief of the size of each fruit= biased/inferenced answer
- Inference in Memory: Hemmer/Steyvers)
  *your strong prior belief moves your recall answer toward typical answers (in this case fruit sizes)
  →We're not just encoding/decoding information
- Ensemble Statistics in Working Memory: (Brady/Alvarez)
  *You set your biases according to the mean of the colored circles
  -so b/c you see big blue circles more than anything else, when you encode your estimate of red/blue circles, you are biased toward the average
- Power Laws in Memory
  →Hermann Ebbinghaus
  *Retention Effects: learn a list of syllables; recall & relearn later when needed
  *Practice Effects: # of trials it takes to learn

→Power Law related retention to time
R= (alpha)(time)^(-beta)

- Psychophysics
  *Variables relation to process/behavior
  →Keep track of memory encoding by our value of what we're learning, and how much we'll need the information encoded now vs later on
  *The strength of our memory activation depends on how likely we are to have to reactive this information later
- Terms

Rational Analysis	*approach that tries to understand under what assumptions the observed behavior would be a good solution to the problem
Bounded Rationally	*we are rational only relative to some resources bound's -finite time, limited memory, limited computation
Ideal Observer Model	*a formal/computational model of what an optimal observer would do -usually Bayesian: bias memory/respond one way or another

→Fruit Size recall task: rational analysis & ideal observer

Concepts 18-24

- What are Concepts:
  →Mental representations of class of objects
  *mental representations we can study, not tied to a single modality (meaning just more that an image/visual representation)
  Ex. Morning Star vs. Evening Star
  *people called Venus a star they saw in the morning b/c they didn't know it was Venus
  *The same thing happened when they saw Venus in the evening
  →in this case, it was the same thing (Venus) being called 2 different things
  *so the thing in the world that the concept refers to is NOT the same as the concept as a mental representation
  **we can have mental representations (like unicorns) that don't refer to anything in the real world
- 3 Incompatible Ways of Seeing Concepts
  *Classical View
  *Prototype/Exemplar View
  *Conceptual Role Theory
- **Classical View**
  →Concepts can be defined through necessary and sufficient features

ex. Spoon: size, used for edible liquids, depth, curvature
*Classical View= rule-based, definitional (strict, logical definitions)
→Compositionality: compositions of concepts can manipulate features
ex. RED spoon= same necessary/sufficient features that adopt RED as a physical feature
-context is important (b/c *tall* giraffe and *tall* grasshopper mean different things)

PROS:

*easy to make compositionality work (esp. in sentences)
ex. "All the rings that belong to your grandmother"

→Powerful Systematicity: concepts are inherently related to each other
ex. If there is a chair in the room…
-there is a physical object in the room
-there is a non-living thing in the room
-there is a piece of furniture in the room

→How do concepts COMPOSE?
*ex. Fast: fast car, fast driver, fast track, fast race, fast time

→For most concepts, there does not appear to be a definition:
ex. Bachelor
-unmarried male (what a/b the pope, widowers, engaged men)

- <u>Boolean Rule Induction</u>
→yes/no examples
→Phenomena in Boolean Rule:
*preference for simplicity
ex. prefer RED over BLUE
*preference for conjunction over disjunction
ex. BLUE *and* RED over BLUE *or* RED
*preference to re-use dimensions
ex. color or color over color and shape

CONS (Classical View):

→Kintsch
*manipulated semantic complexity and looked for processing effects
Ex. convince vs. believe
-convincing involves believing
-convince is more complex than believe (it requires making someone believe something)
**results showed that when participants were shown sentences w/ convince and believe, there was no difference in processing time between the two
→More Challenges:
*Graded Category Membership (2 is a better even number than 178)
*Classical Theory should either form a concept or not, not 50/50
*Typicality
-Robins: typical bird

-Penguins: atypical bird

- **Prototype/Exemplar**

Prototype	Exemplar
*store most typical example in the form of an average of all the examples	*store all of the examples that fit a category

CHALLENGES:

→What a/b things that do not have definitions?

*even #s have a rule, but we still have the intuition that some even numbers are better than others

*Grandmas are old, kind, and grey

-people can attain all these features and not be a grandma

OR

-a person can be a grandma w/o having these features

→How do concepts combine/compose?

Ex. pet fish

*these are two separate representations that we can easily accept as existing together

→Attempt to get compositionality w/ prototypes:

ex. prototype of an apple (w/ all of its features)

*Composition: RED apple; this up-weighs the red feature (we can still represent this even if it isn't the prototype we represent initially)

→We can represent/actively imagine things that don't exist

*create concepts "on the fly"

-these concepts are concrete, but there is no prototypical example for them (ad hoc)

ex. U.S. monarch, not a wolf

- **Conceptual Role Theory (CRT):** (kind of like the conceptual version of functionalism)

→Concepts can only be identified by the role they play in interconnected systems of knowledge

Ex. you can only consider something a horse if you understand genetics/other aspects that make a species individual

→conceptual representations depend on other knowledge

*switching b/w features in development

*characteristics-to-defining shift

-change which features you believe to be most relevant/important

-shift from superficial features (perceptual prototypes) to more deep/defining features

- Essentialism: Gelman/Markman

→idea that there are core/unobservable properties that determine categories (not just perceptual/observable features)

- Interconnection of Concepts, Categorization, and Casual Theories

→ex. someone fully clothed jumped into a pool; are they drunk?

*it's more likely than not that they are drunk (you believe this even if this isn't a stored prototype)

→concept of being drunk has to connect w/ the context of the situation

- Cause weighed over Perception
  Ex. Bill has virus x, which gives him a runny nose
  →Who is more similar to Bill, John or Phil?
  John: Virus x, which gives him a cough
  Phil: Virus y, which gives him a runny nose
  **John is more similar

  CHALLENGES (CRT):

  →Compositionality
  ex. How "pet fish" is built from the roles of "pet" and "fish"
  →Cases w/o any underlying theory
  ex. Discovered Higgs B particle
  -we have a concept of this even though we don't understand it's role/implication completely

Core Knowledge

Vallee-Tourangeau et al. (1998) ● Asked people to generate instances of a category (e.g. fruit) and answer how they came up with them. ● Analyze the responses: – Experiential – memories of specific personal experiences – Semantic – abstract conceptual characteristics – Unmediated – you just do it ● Experiential outnumbers semantic, 4-to-1, even for adhoc categories. → Evidence against abstraction in conceptual knowledge.

-Wu (1995) ● Subjects asked to describe concepts (e.g. watermelon, lawn, face, etc.) with either ● "construct an image for this concept and describe it..." ● "list the characteristics typically true of a concept..." → Similar features are described ● Asked subjects to describe features either for nouns (e.g. watermelon), or noun phrases (e.g. half watermelon) that make the internal features salient. – These noun phrases changed the listed features to highlight internally-available information (e.g. black seeds)

Solomon & Barsalou (1996) ● Feature-verification task: "Does a pony have a mane?" ● Prediction: ● Fast: mane-for-horse → mane-for-pony ● Slow: mane-for-lion → mane-for-pony ● Why? – Mane-for-horse activates the right perceptual category – pony and horse manes are similar – Mane-for-lion activates the right abstract feature, but the wrong perceptual category since lion and pony manes look different.
Battaglia, Hamrick, & Tenenbaum (2013)- intuitive physics engine (rocks on beach)
Lesion studies ● Lesions in sensory/motor areas can lead to apparent conceptual deficits (see Barsalou 2010). Area → Concept ● Visual area → LIVING THING ● Motor areas → MANIPULABLE ARTIFACT, TOOLS ● Color → Color knowledge ● Spatial processing → location knowledge
Smith, Battaglia, & Vul (2013) - S shaped curve for cutting task intuitive physics

Arguments against ● It seems like concepts can't be picture-like because that would require determinate visual features (e.g. the number of stripes on a tiger) (Dennett 1969) ● How do you represent the necessary abstractness? There is no perceptual analog of "all triangles have 3 sides" ● Don't you still need a language/formalism for expressing the simulations and relations between elements? ● Existence of high-level, amodal response cells.

Quiroga et al. (2009) ● Multimodal responses of neurons in medial temporal lobe.
Naive theories of motion (McCloskey 1983)- w/o physics class
The mental imagery debate ● Kosslyn – the brain works with picture-like representations ● Pylyshyn – propositional (logical) representations can get the same thing

*Liz Spelke/Susan Carey
-bootstrapping (learning # concepts)

- <u>Related to:</u>
  →Path Integration (system of mental representations)
  *ants determine direction
  *birds have innate mechanism of rotation of sky (used to determine direction of flight)
- <u>Core Knowledge is built in…</u>
  →early/innate systems of concepts/representations
  →Argument: we can look at early cognition and observe rich concepts/systems that are evolutionarily relevant
  EX. If we were to create a new species… we would build in:
  -concepts of objects
  -how objects interact w/ the world (this builds our expectations)
  ex. animacy/inanimacy
  *early cognitive concepts are good candidates for innate knowledge/concepts
- <u>Experiment: What is built in for humans? (Johnson)</u>
  →Face preferences
  *After being born, infants are shown 3 paddles
  -one blank (control)
  -one with a face
  -one with a scrambled face
  *this experiment tests how far infants are willing to move their heads in order to follow the paddles
  -if infants prefer faces innately, they should be willing to follow the face paddle over the other paddles
- <u>Core Knowledge Domains:</u>

Agents	*move on their own →look for complicated movements not expected, weird sounds, patterns *expectation violation

Objects	*don't move
Number	
Geometry	
Social Groups	*racial bias

- Helder/Simmel: Triangle Experiment
  →Social Cognition
  *triangle moves around trying to get into a box and keep other shapes out
  *very strong percept of their being an agent (even if it' just shapes w/ no human features)
- Gergery & Csibra
  *seeing if you attribute goals to behaviors/outcomes
  *using tasks w/ compatible/incompatible outcomes
  →we believe that the ball moving in a specific way is an agent w/ specific goal in mind and that its actions mimic intention
- Experiment: Social Reasoning (Face Paddles w/ eyes pointed in certain direction)
  →using faces as cues for which way to looks
  *infants looked more in accordance to the direction that the paddle's eyes indicated
- Core Properties of Objects (Kellman/Spelke)
  →What infants expect about objects in the world
  *objects moving side-to-side behind occlude
  -looking habituation paradigm
  *renewed interest when expectations were violated (ex. if infants expected the object behind the occlude to be solid but it is then revealed that the object is fragmented into two objects)
- Properties:

Cohesion	*solid object expected when occluded instead of fragmented
Continuity	*movement of object/occluder
Contact	*objects influence each other through contact (they don't float)

- Static Scene:
  →ex. parts move in separate direction
  *if they are not spatially connected, you perceive object boundaries
  -surprised if they move together
  -unsurprised if they move in different direction
  →ex. parts move together while spatially connected, you don't perceive object boundaries
  -unsurprised if they move together
  -surprised if they move separately
- Chick Studies (Lorenz)
  →Imprinting: from birth, chicks have short time to learn to follow parent/imprint on parent

*innate; learning mechanism (this is not conditioning as the chicks are not being rewarded)
→Chicks imprint on triangle/objects
*looked to see if chick would generalize to other stimuli
*Evidence that chicks perceive objects that are occluded as one object instead of segmented objects

- Core Knowledge of Social Group Formation
  →Evidence in adults for in-group preferences (biases)
  *evidence in infants; they prefer people who are similar to their parents
- Bar-Haim
  →Discrimination paradigm; looking preference for familiar stimuli
  *no evidence in new born infants for race-preference (no genetic influence)
  *by 3 months, there is a strong same-race preference (learning mechanism)
- Language Acquisition (Kinzler, Dupous, Spelke)
  →U.S. English-speaking infants presented two toys, one with English instructions, one with instructions in foreign language
  *they choose the English instructions way more often (prefer native language)
- Beyond Core Knowledge: Natural Language
  →Is language the "glue" that ties together core systems?
  *Does learning language change the types of things you can represent?
  *Language dictating the representations we can make