How We Learn by *Stanislas Dehaene* (2020)

Summary by Helen Reynolds https://itsalearningcurve.education/

Content	Questions/Thoughts
Introduction: Brain doesn't just draw knowledge from environment, possible to learn because brain is plastic. Learning afforded better chances of survival in an environment ∴ needs genes AND environment. Education accelerates learning, as does metacognition. Four pillars: attention, active engagement, error feedback, daily rehearsal/nightly consolidation. Humans' basic circuitry is the same (some variation, some extreme → strategies based on cogsci can help). Machines: 1. only mimic first few hundred milliseconds of brain activity 2. are less effective, but improving.	• Begin born prewired not an advantage • Classrooms exclusive to our species • Children have the same learning algorithm (but vary in what they know) • Fear free error feedback essential • Machine learning takes a lot of data
Part 1. What is Learning? Chapter 1: Seven Definitions of Learning: Brains acquire mental models of everything (physical world/language etc.) via learning, nested like Russian dolls, unconscious. Seven definitions – learning is: 1. Adjusting Parameters of Mental Model (constantly recalibrates from feedback) 2. Exploiting a Combinational Explosion (adding parameters greatly increases options, brain creates hierarchical/multi-level models) 3. Minimizing Errors (event→result→error →adjust = how brains and neural networks work, 'gradient descent algorithm (GDA)'→adjust to minimize future errors 4. Exploring the Space of Possibilities (GDA can get stuck in local minimum, inject randomness, see where it settles like evolution) 5. Optimizing a Reward Function (supervised →get correct answer vs. unsupervised →work towards a goal/reward, evaluate/correct, 'actor + critic' = success of AI, can use competing algorithms 6. Restricting Research Space (learning v. slow if you consider *all* data, need to generalize → useful, convolutional neural networks (CVV) apply learning more generally = efficient e.g. image/speech). 7. Projecting A Priori Hypotheses (need set of assumptions to use first like CVVs do, no learning in a vacuum like babies)	• Learning vs apprehending (Latin root) • We grasp fragments of reality to make mental models • We know around 50 000 words • Error backpropagation at heart of many computer applications • Can use data randomness like mutations • Brain like collection of experts that compete →metacognition helps • Need to reduce # parameter to be useful/generalize • Babies brains programmed by evolution
Chapter 2. Why Our Brain Learns Better Than Current Machines: Brains/machines both recognize/ categorize/assess meaning then brains consciously produce model of world/communicate with language. What babies can do but machines can't: learn abstract concepts/question our beliefs/correct first impressions (hence CAPTCHA), learn faster from less data, learn socially, learn from single trial, learn to generalize from a symbolic rule/conceive formulas or rules with a mental language/use complex & arborescent syntax = underlying tree structure, use higher level learning (1. inferring grammar of a domain: make or use meta level rules with intention/use exclusivity rules with language of thought 2. reasoning like a scientist: prediction/data/compare/judge/evaluate/ adjust = Bayesian, deal with uncertainties) everywhere with everything.	• Timescale: machines act like brains initially, not after time, we do more • Machines have no innate hierarchy • Real challenge to recognize the letter 'A' • Learning = inserting new knowledge into existing framework • We combine innate knowledge from evolution with experience → learning = Bayesian – update our priors
Part 2. How Our Brain Learns Chapter 3. Babies Invisible Knowledge: Learning needs immense set of models + efficient algorithms to adjust to reality: machine use adjustable connections/humans combine symbols. Babies born with concepts of: objects (incl. animals & people)/probabilities/ faces/ability to acquire language. Rapidly learn/develop: 1. how objects support each other/fall & how to prevent it 2. intentions/preferences of people/animals & animate vs. inanimate 3. accurate number sense incl. adding/subtracting 4. sense of which events are probable/which not → enables logic 5. facial recognition/preferring human faces to others 6. vowels in 6 months & consonants in 12/only related to what they hear/words that are common or don't appear, by 1yo they have rules of phonemes/prosody/lexicon/syntax, and can't not do this – it's innate.	Babies • Have brains far from 'blank slates' • have a lot of knowledge on which they rapidly build as infants • have intuitive logic & can do arithmetic! • learn faces quickly for attachment, & spoken language just as quickly • do this faster than any machine
Chapter 4. The Birth of a Brain: Imaging (MRI) enables investigation of structure of newborn brain. 1. Well organized: cortex is already subdivided (auditory/visual/tactile etc.) from birth, info channeled like adults (not understood though), circuits predetermined. 2. Language highways: cables of millions of neurons connecting brain areas already there. 3. Self-organization of cortex: folds start to develop 2nd trimester/almost adult like at birth, full of columns/stripes/borders, neuronal GPS/spatial awareness using grid cells. 4. Origins of individuality: cortical folding unique to individuals i.e. genetic, differences in degree not in kind & produce cognitive differences only at extremes of distribution (e.g. genes associated with dyslexia/dyscalculia, not 100% determined).	• Genes determine structure of the brain (not just the body) • Brain comes already wired up – happens in utero • Unlike machines fetus brain doesn't need data • Genes can tip the scales in terms of differences, environment matters
Chapter 5. Nurture's Share Brain structure: neuron (cell body/axon/dendrites), dendrites connect via synapses (where neurotransmitter (NT) released/absorbed, affects whether electrical signal is passed on). Synapses/dendrites change as we learn = synaptic plasticity. Modulated by NTs (mainly) acetocholine (strong emotions/events, links to amygdala), dopamine (reward/anticipation of reward), serotonin (longer term rewards). NTs determine what's important enough to remember, traumatic events link to amygdala (fight/flight/freeze/fawn response). Portrait of a memory: senses → neurons fire → changes to synapse/dendrites mean more likely to fire in future = memory. Not related to single region of brain, can involve many. 4 types of memory: working memory (WM) (lasts for a few seconds, acts like a buffer), episodic memory (unfolding life events, associated with hippocampus), semantic memory (brain replays memories from hippocampus during sleep, transfers them to other regions = long term memories, knowledge 'that'), procedural knowledge (doing things repeatedly makes memories, knowledge 'how'). Neuroplasticity = changes to brain in learning: larger synapses, more dendrites (so more synapses), more myelin on axon (faster signal), more blood vessels (supply oxygen,	• Main connections present at birth, modulated by environment (microscale) • Connections determined by repetition due to strengthening synapses • "Neurons that fire together wire together" (Hebbs, 1949) • Discovery of neurotransmitters 1921+, effect in brain 1950 onwards • Ideas/types of memory evolving • Brain changes to make memories via synapses, dendrites, axons • Sleep is essential for laying down long-term memories

remove CO_2), more glial cells (heal and feed neurons). Brain needs nutrients to change. Neuroplasticity both vast (sensory input not needed (like in utero)) and limited over space (how far neurons connect), over time (sensitive to change at particular times). Infants = overproduction of synapses followed by pruning/myelination related to environment (keep what's useful) in waves throughout brain by region: language/syntax (1yo) →sensory (visual 2yo)/ motor → higher cortical areas (pre frontal cortex (PFC) 5-10yo) → grammar (end of puberty). Sensitive period ends → synapses closing → balance of excitation(E)/inhibition (I). In children E >>I, I increases as they age. Evolutionarily advantageous to close window, we keep traces of childhood.	• Myelination = adding insulation to axons • A young child's brain uses about 50% of calories (adults more like 20-25%) • Peak sensitivity to learning = 1 – 2yo • Periods are sensitive not critical, can learn outside of them, just harder • Early education should be a priority • Higher cortical learning always possible
Chapter 6. Recycle Your Brain Paradox: brain physically limited/can learn indefinitely → neuronal recycling (NR) (author's hypothesis). Learning new cultural objects repurposes existing circuits, but need to be similar; regions impose constraints. 1. Maths: education molds innate number sense, links to attention circuits (number line ↔ looking left/right) expands *existing* circuits relating to quantity (sighted/unsighted → same brain region, visual cortex can be repurposed). 2. Reading: education molds innate sense of objects via vision, learning to read reorganizes visual cortex, competition with facial recognition (which relocates), visible in images of cortices of illiterate/ literate/those learning to read. 3. Music, math, & faces: music → visual cortex, equations/algebra compete with faces. Genes give structure & flood child's brain with synapses → door open for learning via adaptation to environment hence importance of schools/families in all areas.	• NR enables humans to learn a LOT more than evolution alone allows • Mathematicians & babies process math/ number in the *same* brain areas • Learning to read from 6-8yo dovetails with intense brain plasticity • Ditto learning to read music • A child's brain is structured AND plastic • Enriched environments (school and home) build better brains
Part 3. The Four Pillars of Learning: Evolutionarily optimize successful learning **Chapter 7. Attention:** = all the mechanisms involved in selecting/amplifying/channeling/deepening processing of info. Why? Triages input → avoids info overload → speeds learning. Conscious attention amplifies effects of sensory/conceptual neurons → to PFC where neurons fire for much longer than duration of image. 3 processes: 1. Alerting (tells us when to attend, level of vigilance, evolution – oldest = warning system for predators) 2. Orienting (signals what to attend to, filters and amplifies neurons firing, reduces attention on everything else = attentional blink, phonics = attending to link between sound & word activates reading circuits (whole words does not)) 3. Executive attention (EA) (= concentration/self-control, decides how to process the info from it, selects relevant processes, controls execution, all possible due to PFC BUT no multi-tasking possible just task switching/ slower). Learning to attend: 15-20y before PFC mature (so young children make silly mistakes), some evidence that EA can be improved (e.g. learn musical instrument). Social learning: I'll attend if you attend, learning from others = evolutionary adaptation, children attend to cues to attend. Teaching →attending others' knowledge/unique to humans as we have theory of mind. 2 modes: 1. Active, test hypotheses 2. Receptive, trust what others say without verifying it.	• Attention mechanisms evolutionarily ancient (and in all animals) • Teachers should pay a lot of attention to what students (Ss) are attending to • Can't learn what you don't attend to • 'Invisible Gorilla Illusion' good to demo 'biased competition' in attention • PFC of primates bigger/better connected with more neurons • EA like switchboard/controller • Improving EA can improve IQ (fluid I) • Your attitude/gaze mean everything to a child (from evolution) • 2nd mode susceptible to 'fake news'.
Chapter 8. Active Engagement: Passive ≠ learning. Active = 1. Form mental model 2. Project onto environment 3. Form prediction 4. Check prediction. More cognitive effort = deeper learning (Roediger). Discovery learning (sounds good but) fails → 'busy' active ≠ learning (but gives Ss illusion of mastery) = a myth like digital natives/learning styles. Make active by piquing curiosity/ drives exploration/releases dopamine/signifies reward/essential evolutionarily/cognitive curiosity = reward like food. Mirth triggered when assumption/prediction is wrong. Curiosity y-axis follows bell curve with familiarity x-axis. Babies do it – have metacognition/know what they don't know.	• Need attention + effort +deep reflection • Processing depth linked to effort hence lecturing (passive) ineffective • Beware zombie myths • Bear in mind bell curve: pitch stimulus • Reward (NEVER punish) curiosity • Be a curious teacher (and show it!)
Chapter 9. Error Feedback: Surprise drives learning → events violate expectations: prediction → error → adjust model according to size of error. Neural networks work same way/brain does more by using language of thought. Feedback matters not number of mistakes. Hierarchy in cortex tries to make sese of error → adapt model. Get reward/predict reward → dopamine. Grades = reward/ punishment AND not useful feedback, promotes fixed mindset for unsuccessful AND successful Ss. Better: self-testing/retrieval practice/spaced practice – because cramming ≠ novelty/uncertainty.	• Learning depends on degree of surprise • Brains don't absorb associations passively – need action • Needs more than 'you're wrong' • Low stakes testing better than grades • Overlearning capitalizes on uncertainty
Chapter 10. Consolidation: Automaticity is the goal: brain compiles routines outside awareness, saves time/energy/WM. Sleep essential: causes additional learning/previous day's experiences replayed & consolidated (role of stages of sleep not well understood)/can also produce insights. Hypothesis: during night brain works top-down, uses generative models to interrogate data from the day/update models. Dreams = enhanced training set of data. Puberty changes sleep cycle.	• Automaticity frees brain's resources • Sleep needed between learning episodes to enhance learning • Schools not helping adolescent sleep given change to sleep cycle in puberty
Conclusion: Reconciling Education with Neuroscience: 13 take-home messages: 1. Do not underestimate children 2. Take advantage of the sensitive periods 3. Enrich the environment 4. Rescind the idea that all children are different 5. Pay attention to attention 6. Keep children active, curious, engages, and autonomous 7. Make every school day enjoyable 8. Encourage efforts 9. Help students deepen their thinking 10. Set clear learning objectives 11. Accept and correct mistakes 12. Practice regularly 13. Let students sleep. Scientists should engage with teachers.	• Education needs to embrace cognitive neuroscience as medicine embraced biology • With freedom (to teach) comes the responsibility to understand how we learn.