

Week 4 — The Neuroscience of Value

A note on how these notes were made. They were drafted with the assistance of Claude Opus 5 (Anthropic), working from the lecture slides, the course materials and my own outline for the session. I reviewed and edited the result and I am responsible for the content — what is covered, how the studies are interpreted, and any errors.

Accompanying notes for the slides

How to use this. Open `DMUU_week_04_neuroscience_of_value.pdf` next to this document and read along. Slide numbers are PDF page numbers (61 slides). Some slides need a paragraph, some need one line.

Normally this is a 90-minute lecture. Give it about that.

One thing to know before you start. The first study in this deck is the Plassmann paper you just worked through in **Assignment 1**. These notes do not re-teach it from scratch. They recap it quickly, go a bit deeper in two places, and **answer several of the assignment questions directly**.

Where we are, and the question for today

Slide 1 — Title

Last week was about *methods* — how we measure anything at all. This week we point those methods at one thing: value.

Slide 2 — Recall week 2: "What do you value?"

This is the Padlet board from Week 2. You wrote: **family, health, time, love, self-realization, self-fulfillment, safety, fun, discipline, ecology, friendship, emotional stability, happiness, football, effort, friends, experience, work, freedom**. Family and health drew the most votes.

Now notice what is (almost) completely absent: money, food, drink, ...

Remember this, because almost every study you are about to see uses money, snacks or juice. By the end of the lecture the question is whether that gap matters — whether the lab studies something different from what you actually care about, or the same machinery with cheaper stimuli.

Slide 3 — Three questions

1. **What is value to a human brain?**
2. **How does the brain compute value?**
3. **Can we model and predict it, and how does that computation compare to the decision models used in economics?**

Question 3 is what connects this week to Weeks 2 and 3. Keep it in view.

Slide 4 — Remember: economic models of choice

Back to Week 2. Economics models choice with "**as if**" **models**: the decision-maker behaves *as if* they built an ordered list and picked the top one. That is a claim about behaviour, not about mechanism.

The third bullet is deliberately provocative: *whether or not there are neurophysiological correlates of utility is irrelevant*. For the economic model to do its job, it does not matter what the brain is doing.

That is not a weakness. It is a statement about what the model is for. But it means that when we go looking inside the brain, we are asking a **different question** — not a better version of the same one.

Slide 5 — Marr's levels of analysis

From previous weeks. The frame that stops the three disciplines talking past each other.

- **Computational** — what problem is being solved, and why this solution? (Economics.)
- **Algorithmic** — what representations and steps actually produce the choice? (Psychology.)
- **Implementational** — what physical hardware carries out the algorithm? (Neuroscience.)

This week is almost entirely implementational. Weeks 2 and 3 were mostly the other two.

Assignment 1, question 3.3. You were asked at which level Plassmann makes its main claim, and which level it leaves open. The claim is **implementational**: *this signal, in this tissue, tracks this quantity*. What it leaves almost untouched is the **algorithmic** level — the paper does not tell you *how* a willingness to pay gets computed, only that something in mOFC co-varies with the answer. That gap is not a flaw in the paper. It is the normal division of labour, and it is why you still need Weeks 2 and 5 to make sense of Week 4.

The caveat on the slide is real: Marr proposed this for a single system computing one function (vision). Use it, but don't over-trust it. Also remember that the mapping onto the three disciplines is an over-simplification.

The line from Week 3 worth carrying into today: *you cannot understand a brain signal until you know which quantity of the decision it is supposed to encode*. That becomes the whole point of the second half today.

Slide 6 — Principal brain structures for reward and decision making

Vocabulary, not content — and you located most of these yourself in **Assignment 1, question 2.5**: right OFC, left DLPFC, vmPFC, ventral striatum, dorsal striatum, and the bonus one, right anterior insula.

The two names that matter most today are **vmPFC** (ventromedial prefrontal cortex, which sits next to and overlaps with medial OFC) and **ventral striatum**. The **insula** — your bonus structure — comes back in Week 7.

Come back to this slide whenever a region name stops meaning anything.

Part 1 — What does the brain track?

Slides 7–19 — Plassmann, O'Doherty & Rangel (2007)

Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions

You just spent about three hours on this paper. So this section is a recap, not a first pass. Two things are worth doing properly: making sure the *logic* of the design is solid in your head, and seeing how it sets up everything that follows.

Slide 8 — Motivation. Every economic transaction contains a willingness-to-pay computation. We know a lot about what people *do*, very little about how the brain computes it. Hypothesis: medial OFC activity encodes WTP.

Slide 9 — Hypothesis based on... Prior work pointing at medial OFC. Note the first bullet cites monkey single-neuron recording (Wallis & Miller; Padoa-Schioppa & Assad). **We are not covering monkey electrophysiology** — treat that as background citation. (You will see one example of the method at the very end of the deck, as a historical note.)

Slide 10 — Experimental design. 50 familiar sweet and salty junk foods. Subjects **fast for four hours**, then bid for the right to eat one snack at the end. 100 bidding trials: each of the 50 items gets one free bid and one forced bid. One trial is picked at random and implemented; then they sit in the lab for 30 minutes and eat it if they won it.

Assignment 1, question 2.3 — the four-hour fast. The psychologist who would call this *the* key manipulation is **Maslow** (need-based value), and in a more mechanistic form **Hull** (drive). The point: the value of a Mars bar is not a property of the Mars bar. It depends on the state of the person looking at it.

That is a direct problem for the economic assumption of **stable preferences**. If four hours without food can move your willingness to pay, then "preferences" are not a fixed object the experimenter is reading off — they are something the brain assembles given its current state.

And the obvious follow-up — *what if the participants had just eaten?* — is not hypothetical. It is **slides 34–39 of this deck**. Someone ran exactly that experiment. Keep your answer in mind and check it against Gottfried later.

Slide 11 — The BDM auction. You bid b ; the computer draws a random number n (equal chance of \$0, \$1, \$2, \$3). If $b \geq n$ you get the item and pay n . If $b < n$ you get nothing and pay nothing.

Assignment 1, question 3.1 — why bidding your true WTP is optimal. Say an item is worth \$2 to you. Bid \$1 instead: whenever the draw comes up \$2, you lose an item you would happily have bought for \$2 — and bidding \$1 never lowers the price you pay, because the draw sets it. Bid \$3 instead: whenever the draw comes up \$3 you buy something worth \$2 to you for \$3. Shading in either direction can only hurt you, and never helps. So you bid \$2.

The assumption it rests on: your bid must not influence the price. The moment you believe it does — or that the experimenter will use today's bids to set tomorrow's prices, or that you can resell the item — the argument collapses and strategic shading becomes rational again. This is also why BDM works better in a lab than in a market.

Assignment 1, question 2.1 — where the randomness is. The authors chose highly familiar foods precisely to strip uncertainty *out* of the valuation itself: you know exactly what a Mars bar is like. Then the design deliberately puts randomness back in, in two places: **the random price draw n** , and **the random selection of which single trial gets implemented**.

Both are **risk** in Knight's sense — the distribution is known and stated — so on the Week 1 ladder the participant sits on the second rung, not in ambiguity and certainly not in Knightian uncertainty. And the design *needs* both: the random price is what makes truthful bidding optimal, and the random trial selection is what stops you hedging across 100 trials by treating them as a portfolio. The randomness is not noise. It is what makes the number mean anything.

Slide 12 — The novel approach. The most important idea in the lecture.

When you show someone a picture of a Mars bar, their brain does many things at once: recognises the object, generates hunger, plans a response, computes value. All of these co-vary with WTP. So if you simply look for regions whose activity tracks WTP, you will find regions doing *something correlated with WTP*. The slide says it plainly: *WTP erroneously attributed to wrong brain areas*.

The fix: two trial types.

- **Free-bid** — you choose the bid. You must compute your WTP.
- **Forced-bid** — the amount is given. Same picture, same hunger, same motor response, **no WTP computation**.

A region genuinely encoding WTP should track it in free trials and not in forced trials.

Assignment 1, question 3.2(b) — what the design does *not* rule out. Free and forced trials differ in more than the valuation. A free bid is a harder decision: more deliberation, more response conflict, more attention, longer to decide. So a region that lights up more in free trials and scales with the bid could be tracking **decision difficulty or effort** rather than value — and plausibly those two things are correlated, because deciding is hardest when an item is somewhere in the middle of your range.

If you wrote something along those lines, you were right. If you wrote "it doesn't prove causation" — true, but that is a different and weaker point, and it is slide 54.

Slide 13 — Preprocessed time course, and an example predictor. (*This slide was not in the paper's figures, so it may be new to you.*)

The top panel is **real fMRI data**: the signal from one region, across roughly 390 scan volumes. Two things to notice. The y-axis runs from about 770 to 860 in **arbitrary units** — there is no "amount of brain activity" scale, which is why every result in this field is reported as a *percent signal change* relative to some baseline. And the thing is visibly noisy: the trial-related signal is a small ripple on top of drift, heartbeat, breathing and head movement.

The bottom panel is the **predictor**. The grey bars are the events — when the food items appeared. The white line is what the BOLD signal *should* look like if that region responds to those events: each event convolved with the slow, lagged haemodynamic response, so the bumps are smeared out over several seconds and arrive late.

The analysis is then just this: how well does the white line explain the top line? Do it with the bid height as a weight on each event, and you are asking "does this region's signal rise more on trials where the bid was higher?" That is the entire logic of a parametric fMRI result.

Assignment 1, question 2.4(a) — why BOLD is indirect. This slide is the answer in picture form. You are not recording neurons. You are recording **blood oxygenation**, which follows neural activity by several seconds because the vasculature responds slowly. The signal you fit is a *model* of that lag, not a measurement of firing.

2.4(b) — EEG instead? You would gain time: millisecond resolution instead of a four-to-six-second smear, so you could see when in the trial a value signal appears. You would lose place: EEG is poor at localising sources, and worst of all for exactly the structures that matter here — mOFC and striatum sit deep, where EEG essentially cannot see.

Slide 14 — Behavioural results. The distribution of bids in free trials against the forced amounts. The point is only that the two are comparable in range, so no neural difference is driven by the numbers differing.

Slides 15–16 — Neuroimaging results.

Analysis 1 (free trials). Whole-brain search for regions scaling with WTP. Medial OFC, as predicted; also dorsal anterior cingulate and DLPFC.

Analysis 2 (forced trials). Same search. **No region correlates with WTP.** They also checked whether the brain tracks the forced bid amount instead — nothing.

Analysis 3 (the actual test). Find regions that both (a) increase with WTP in free trials and (b) are significantly more active in free than forced trials. Right mOFC passes, as predicted. Right DLPFC also passes, which was not.

Slide 17 — Neural correlates of WTP. Three time-course panels. In free trials the mOFC signal separates cleanly by WTP. In forced trials, no separation by WTP and none by the forced bid either.

Slide 18 — Summary. Hungry subjects, real bids, real snacks. Right mOFC and right DLPFC encode willingness to pay.

Assignment 1, question 3.3 — nineteen people. This is a group-level result from 19 participants. It tells you that *on average, across this sample*, mOFC activity tracked WTP. It tells you very little about any single one of them — some individuals will have shown the effect strongly, some weakly, some not at all, and the group map does not say which.

This is Week 1's **height analogy** again. "Dutch men are taller than Italian men" is a true and useful population claim that predicts almost nothing about the next Dutchman and Italian you meet. A group fMRI map is the same kind of statement. Keep it in mind every time you read a sentence of the form "the brain does X."

Slide 19 — Discussion, and the bridge. The authors raise the limitation themselves: everything so far is about **primary appetitive rewards** — food, for a hungry person. Does it generalise to books and DVDs, to money, to losses?

That question is what the next study answers.

Slides 20–24 — Chib, Rangel, Shimojo & O'Doherty (2009)

Does vmPFC encode willingness to pay "in general"?

Slide 20 — The question and design. Before scanning, subjects place BDM bids for an 80% chance of receiving items from three very different categories:

- **Money** (\$1, 2, 3 or 4)
- **Food** (sweet and salty snacks)
- **Non-food goods** (DVDs and Caltech memorabilia)

Then in the scanner they make buy/don't-buy decisions at a reference (median) price.

The test, in the slide's own words: does the **same** vmPFC region correlate with the WTP elicited before scanning, across all three categories? If yes, that is evidence for a **common neural signal encoding value for different goods**.

This is the common-currency hypothesis, as sharply as it can be put.

Slide 21 — Design. Pre-scan: free-response BDM auction (the Caltech hat, \$0–\$4). In-scanner: 2-second binary choice, then 1–10 seconds of fixation.

Slide 22 — Behaviour. Panel C: people bid differently across categories, as you'd expect. Panel B: but the probability of *selecting* an item rises with WTP in the same way for all three. The behaviour is consistent even though the items are not comparable in any obvious sense.

Slide 23 — Similar regions, regardless of category. Three separate analyses, one per category. Overlapping activation.

Slide 24 — The result. One vmPFC region whose signal correlates with WTP for *each* category, three regression lines sloping the same way.

Why this matters more than it looks. A common scale is not a curiosity. It is what makes choice between unlike things possible at all. You can compare a sandwich to a DVD to four dollars because somewhere the brain puts them on one axis. Without that, choosing between incommensurable goods would need a separate mechanism for every pair.

Slides 25–28 — Plassmann et al. (2008): the price of wine

Slides 25–26 — Setup. Subjects taste wine in the scanner and are told the price before tasting. Five price labels — \$5, \$45, \$10, \$90, \$35 — but only **three wines**. The \$5 and the \$45 are the same wine. The \$10 and the \$90 are the same wine.

Slide 27 — Behaviour. Liking ratings are significantly higher for the "expensive" presentation of an identical wine. **Intensity** ratings do not differ, so it isn't that price changed how strongly they tasted it. Tasted later without prices, the effect disappears.

Slide 28 — Neural results. The medial OFC signal follows the liking ratings. The same wine produces a larger mOFC response under the higher price tag.

The point. Nothing about the stimulus changed. The molecules were identical. What changed was a *belief* about the wine, and the value signal moved with the belief.

This is the first real crack in the idea that value is a property of the thing being valued.

Slides 29–31 — Hare, Camerer, Knoepfle & Rangel (2010): charitable donations

Slide 29 — Design. Subjects are endowed with \$100 and decide how much to give to each of 75 charities. One trial is picked at random and implemented, so the money is real. Before scanning they read a description of each charity and rate how *deserving* it is.

Then — and you have seen this trick before — two trial types: **forced donations** (a randomly determined amount) and **free donations** (you decide). Only in free trials must you compute a subjective value.

Slide 30 — Behaviour. Donations rise systematically with the deservingness rating.

Slide 31 — The neural result. The model on the slide, $U = \theta \cdot \ln(d) - d$, just says "the value of donating is the warm glow from the gift minus the money you gave up," with θ a subject-specific weight. You don't need the algebra. What matters: vmPFC activity correlates with that subjective value.

And the line on the right is the payoff: **this overlaps with the value signals for primary food rewards** in Plassmann (2007).

So: the same region, for snacks, for money, for DVDs, for wine, and for giving money away to strangers.

Remember slide 2. Nobody wrote "money." Several of you wrote things like *being generous, ecology, friendship*. This is the first result today that is actually about that kind of value — and it lands in the same place.

Slide 32 — Preliminary summary

vmPFC seems to encode subjective valuations for a wide variety of **nonsocial and social** goods.

(Two items on that list — delayed monetary rewards, and the fairness of offers — come from studies not shown here. They are on the slide because the claim rests on more evidence than fits in ninety minutes.)

That is the common-currency claim, about as strong as it gets. Now we complicate it.

Slide 33 — Three main functions of reward

A signpost. Reward does three jobs: it drives **approach behaviour and decision-making** (everything so far), it drives **learning** (the end of today, and Weeks 5–6), and it produces **pleasure** (which this course largely leaves alone).

Hold the second one.

Slides 34–39 — Gottfried, O'Doherty & Dolan (2003)

The complication: value is not a property of the thing

These six slides are one study, and they are the answer to your Assignment question 2.3. Read them as a block.

Slide 34 — The Pavlovian human. Classical conditioning run on people. Abstract fractal patterns (conditioned stimuli) are paired with food odours (unconditioned stimuli): one pattern predicts banana, another vanilla, a third nothing. The dog diagram is the reminder — Pavlov's experiment, different species, better equipment.

Slide 35 — Imaging conditioning. Trial structure in the scanner. Picture for 1000 ms; on some trials an odour; a sniff is cued. *Tgt* and *nTgt* are target and non-target — one of the two foods gets devalued later, and that is the target. *CS+p* and *CS+u* are paired and unpaired presentations; *CS-* predicts nothing.

Slide 36 — Neural responses to conditioned stimuli. After conditioning, the patterns themselves — coloured shapes — evoke responses in amygdala, piriform cortex, OFC and ventral midbrain. The cue has acquired value.

Now the manipulation: subjects leave the scanner and eat one of the two foods to satiety. In this case, the researchers gave them a lot of bananas and "encouraged" the subjects to eat a lot of them. Then they come back and do it again.

Slide 37 — Behaviour, before and after. Hunger drops sharply. The pleasantness of the food they ate drops sharply. And critically, panel C: **the pleasantness of the associated odour drops — but only for the food they ate.** The other odour is unchanged.

Slide 38 — Brain regions reflecting changes in preference. Amygdala, OFC, ventral striatum, insula, cingulate.

Slide 39 — The whole figure assembled. Behavioural data (A–C) with the neural results beneath. Orange bars are the devalued food's cue, blue bars the other one. The cue's value signal falls for the food you just ate and holds for the one you didn't.

What this means. The fractal pattern did not change. The odour did not change. The only thing that changed is the state of the person looking at it. And the value signal moved.

Value is not a stable property sitting in the object. It is computed, on the spot, given what you currently need.

Back to Assignment 1. Question 2.3 asked what you would expect if the participants had just eaten. Here is the experiment. Compare it with what you wrote.

It also settles **Part 4, headline C — "Hungry brains value snacks differently."** Plassmann cannot support that headline: every participant was hungry, so hunger was held constant and never varied. The study has no evidence either way. **Gottfried is the study that earns the headline** — and notice it took a within-subject manipulation, scanning the same people twice, to do it. That is the difference between controlling a variable and testing it.

Be careful which kind of change this is. This is value shifting with your *current internal state* — hungry versus full. It is **not** value being updated by learning across many trials. That second sense is Weeks 5–6, and the last slides of today sit exactly on the boundary. Keep them separate.

The turn

Slide 40 — Two different points of view

The most important slide in the deck, and there is almost nothing on it. Read it slowly.

Economics looks at **choice**. A choice happens at a moment: you picked A over B. Utility describes which choice you will make. That is a complete and useful account, and it was Week 2.

Psychology and neuroscience look at **decision-making**. That is not a moment; it is a process unfolding over seconds. Something represents the options, something values them, something picks, something evaluates what happened.

Brain activity is continuous. There is no single instant at which "the value" exists to be measured. So "where is value in the brain?" is **incomplete**. You have to say *when*.

Everything in Part 1 was one quantity, measured at roughly one moment in a trial — and we got a beautifully consistent answer: vmPFC. The rest of the lecture asks what happens when you stop doing that.

This is also the Week 3 line coming due: *you cannot understand a brain signal until you know which quantity of the decision it is supposed to encode*. Rangel's framework is how we name the quantities.

Part 2 — When does the brain compute value?

Slide 41 — Value-based decisions: a framework (Rangel et al., Nat Rev Neurosci, 2008)

Representation → Valuation → Action selection → Outcome evaluation, with **Learning** feeding back into the first three.

Do not memorise this as a taxonomy. It is the answer to a measurement problem: if you are going to put a number on "value," this tells you which number and at what point in the trial it exists.

Slides 42–46 — The five stages

Each slide highlights one box and lists open questions in that part of the field. You don't need to answer them. You need to see that each stage raises *different* questions.

Slide 42 — Representation. How does the brain decide which actions are worth considering at all? Is there a limit? How are internal states (hunger) and external states computed and passed on?

Slide 43 — Valuation. *This is where today sits.* What are values? Are there multiple types of value, or multiple valuation systems? Where are they represented? How does the brain handle delayed rewards, and indirect rewards like money?

The second question is exactly what Part 1 left hanging.

Slide 44 — Action selection. Does the brain compare the value of outcomes, of the actions needed to get them, or of action–outcome pairs? How are costs and benefits integrated?

Slide 45 — Outcome evaluation. What networks compute positive and negative outcome values, and do they overlap with valuation? Do outcome values reflect hedonic enjoyment, long-term goals, social norms? How are delayed benefits credited back to the decision that caused them?

Slide 46 — Learning. How are the units and rates of learning determined? Is learning passive or active? What are the neural mechanisms? Can value systems learn by observation rather than direct experience?

This is next week's slide. Note that it feeds back into all three earlier stages — which is why learning is not a separate topic bolted on at the end.

Slide 47 — Multiple valuation systems

Three decision systems from the behavioural literature:

- **Pavlovian** — hard-wired responses to predictive cues. Eat all the food on the plate.
- **Habitual** — learned, automatic, stimulus-driven. The morning coffee; the usual route to work.
- **Goal-directed** — flexible, outcome-sensitive. Choosing a film; going for a run.

Everything in Part 1 was goal-directed, which is why it is highlighted. The other two exist, behave differently, and may use different value signals in different circuits.

Slide 48 — Even more values...

The slide that makes "value" visibly plural. Five distinct quantities, all called "value" in ordinary speech:

- **Goal value** — what the outcome is worth to you.
- **Stimulus value** — expected subjective value of consuming it; accounts for risk, delay and price.
- **Action cost** — the subjective cost of what you must *do* to get it.
- **Action (net) value** — stimulus value minus action cost.
- **Decision value** — the *relative* value of two options.

These are not synonyms. The same cake has a high stimulus value and, if it is in a shop across town in the rain, a low action net value. A study measuring one is not measuring the others.

Go back to slide 32 with this in hand and ask which of these Plassmann measured, and which Chib did. The answer is not obvious, and that is the point.

Slides 49–50 — Boorman et al. (2009): tracking value across the trial

Slide 49 — Design. Subjects choose between two options with different reward probabilities. They are *not told* the probabilities — those are inferred from choices using a Bayesian reinforcement-learning model. The timeline is spelled out in seconds: trial at 0s, response allowed from ~6s, response ~7.5s, outcome ~13.5s, outcome removed ~16.5s.

The long, deliberately spaced timeline is what makes the next slide possible: with events this far apart, the sluggish BOLD signal can be pulled apart into stages.

Slide 50 — The result. Sit with this one.

Top: vmPFC tracks the **relative chosen value** — how much better the thing you picked is than the thing you didn't — peaking around the response.

Bottom: the same data split into **chosen value** (red) and **unchosen value** (blue), plotted across the whole trial. **They have opposite time courses.**

Three things fall out of one figure:

1. Chosen and unchosen value are **different quantities**, not one quantity measured twice.
2. Each has its **own time course**. Neither is "the" value signal.
3. The answer to "does vmPFC encode value?" depends entirely on **which second of the trial you ask about**.

That is slide 40 demonstrated rather than asserted. Part 1's answer was clean partly because every study looked at roughly the same moment.

Slide 51 — Subjective value in the human brain is...

The widest version of the claim. Ventral striatum responds to: music you like, cuddly animals, Facebook likes, winning virtual points, being right, being generous, beauty, erotic stimuli, curiosity and novelty, opiate drugs, exercise-induced pain, being in love, **meaningful hugs and honest smiles**.

Look back at slide 2 once more. *Friendship. Love. Happiness. Fun.* The lab studies snacks and dollars because snacks and dollars can be delivered in a scanner — not because the machinery is different.

The box along the bottom is worth more than the list above it:

- **Rewards are interchangeable?** — a question, because it is not settled.
- **Uncertain rewards produce higher activation** — a preview of Weeks 7–8.
- **Individual differences** — the same signal has different gain in different people. The recurring theme, and the same point as the 19 participants in Plassmann.
- **Depends on memory; adapts to counterfactuals and comparison** — value is relative to what else was possible and what came before.
- **Reflects future prediction (optimism/pessimism).**
- **This computation, and the resulting feeling, happen nonconsciously.** You have no introspective access to any of it.

Slide 52 — Components of the value network

A meta-analysis across many studies (Bartra, McGuire & Kable, 2013). "ALE" is activation likelihood estimation — a way of asking which coordinates appear reliably across a whole literature rather than trusting any single experiment.

The answer: **vmPFC, ventral striatum (VSTR)**, and posterior cingulate (vPCC, dPCC).

Note the plural in the title. Not a value *region*. A value **network**.

Closing

Slide 53 — Summary: what the brain tracks

Five claims, each earned by something you have now seen. Nothing new here — but if any line doesn't connect to a specific study in your head, go back and find it.

Slide 54 — ...and what it does not show

Correlation, not causation. This is the single most important sentence of the week, and you have already met it twice: once in the design of Plassmann's forced-bid trials, once in Assignment 1.

Every result today is a signal that **tracks** a quantity. None of it shows that a region *computes* value, or is *necessary* for it. The method family that could show that is **interventions** — lesions, stimulation, pharmacology — which you met as the fourth column in Week 3, and which we return to later in the course.

Assignment 1, Part 4 — the three headlines.

A. "Scientists find the brain's price tag." *Partly supported.* There is a signal that scales with willingness to pay, so the headline is not invented. But "the price tag" overstates it in two ways: the result is correlational, and mOFC was not the only region that passed the test — right DLPFC did too, unexpectedly.

B. "Your OFC decides what you buy." *Not supported.* "Decides" is a causal claim, and nothing in the design can support it. Plassmann observed; it did not intervene. This is the headline that the whole free/forced apparatus still cannot earn.

C. "Hungry brains value snacks differently." *Not supported by this study* — everyone was hungry, so there was nothing to compare against. Supported by **Gottfried**, which you met at slides 34–39.

If you rated all three the same, that is the thing to look at. They fail in three different ways: A overstates, B misattributes causation, C asks a question this design cannot answer.

Back to week 2. You said family, health, time, friendship, freedom. The lab says snacks and dollars. The same circuitry prices a Mars bar, a charitable donation and an honest smile — the gap is one of convenience, not of mechanism. Whether that is reassuring or unsettling is a reasonable thing to have an opinion about.

Preview: where this goes next

Slide 55 — Divider

Everything from here is a **preview of Weeks 5 and 6**. You are not expected to work with it yet. Read it once, let it be strange, and come back after next week.

Slides 56–57 — Where this came from: recording from single dopamine neurons

A historical note, not examinable. We have not covered monkey electrophysiology and we won't. But every computational account of reward in the brain — including next week's — traces back to these recordings, made in awake monkeys from the late 1980s onward, so it is worth knowing what they actually looked like.

Slide 56, panel A. A monkey sits with its hand on a resting key and reaches into a covered food box. The raster (each row one trial, each tick one spike) and the histogram above it are aligned to movement onset. The dopamine neuron fires a burst around the moment the hand touches the food. Nothing warned the animal — the food was simply there.

Panel B. The control that makes A mean something. Top: touching a piece of apple. Bottom: touching a bare wire. The neuron responds to the food and not to the wire. So this is not a response to reaching, touching or moving. It is specific to reward.

Slide 57 adds panels C and D.

Panel C is the heart of it — three conditions, labelled on the figure as **+error**, **0 error**, **–error**:

- *No prediction, reward occurs.* Burst at the reward. **Positive error.**
- *Predictive stimulus, then reward.* The burst has **moved to the stimulus**. At the reward itself, nothing. **Zero error.** The reward has not become less rewarding; it has become expected.
- *Predictive stimulus, no reward.* Burst at the stimulus, then at the exact moment the reward was due, firing **drops below baseline**. **Negative error.**

Panel D shows two further paradigms from animal learning theory — blocking and conditioned inhibition. These are the controls that separate "this neuron tracks prediction error" from the weaker claim "this neuron responds to things that have been paired with reward." The response follows what the animal has learned to *predict*, not its pairing history.

Slide 58 — Neural responses to reward prediction errors (Schultz et al., 1997)

The same result, redrawn cleanly, with the dopamine pathways on the right and Pavlov's dog underneath as the reminder of where the logic comes from.

Three conditions, stacked: response to an unexpected reward; response to the predictive cue once it has been learned; response to an omitted reward.

Put them together. **The neuron is not responding to reward. It is responding to the difference between what was expected and what arrived.** Better than expected: up. Exactly as expected: nothing. Worse than expected: down.

That is a **prediction error**, and it is the signal that drives learning.

Slide 59 — Learning

Rangel's framework again, with the Learning box highlighted. This is deliberate: the last four slides are not a new topic. They are **the one box we skipped** when we walked through the framework at slides 41–46.

Learning feeds back into representation, valuation and action selection. That is why it cannot be left to the end as an afterthought — it changes all three.

Slide 60 — Reward-based learning

The loop in words: you hold a **reward prediction**. An **outcome** happens. The gap between them is the **reward prediction error**. That error, scaled by a **learning rate**, updates the prediction. Then round again.

Two things to notice, because they are what Week 5 is actually about.

If the prediction error is zero, nothing changes. A perfectly predicted outcome teaches you nothing — which is exactly what the middle panel of slide 58 showed. Learning is driven by surprise, not by reward.

The learning rate decides how much you move. Set it high and you lurch after every surprise, treating noise as news. Set it low and you are stable but slow to notice when the world has genuinely changed. There is no universally right setting — it depends on how volatile the environment is, and it differs between people. That is the individual-differences thread again, now as a parameter you could in principle measure.

Slide 61 — Reinforcement learning (preview)

Don't panic at the algebra. The structure is the same as slide 60, written precisely.

The **value function** $V(s_t)$ is the sum of all future rewards from here on, with later rewards discounted by γ . By definition that can be rewritten recursively: the value of where you are now is the immediate reward plus the discounted value of where you will be next. If your current estimate does not satisfy that identity, the mismatch is the **temporal-difference error** δ_t — and that mismatch is what you learn from.

δ_t is the formal version of what the dopamine neuron was doing on slide 58. Same quantity, two languages.

This is Week 5. You are not expected to use these equations. You are expected to recognise that the dopamine response and the TD error are two descriptions of one thing — and that is the single most important bridge in this course.

What to do with this week

- **Read the self-study document on the history of value and reward in psychology.** It covers the traditions behind these ideas — reinforcement, expectancy-value, needs — and goes further into conditioning and reinforcement learning than the slides do. It is the bridge into Week 5.
- **Compare your Assignment 1 answers against the answer guide on Moodle** and against the notes above. Several of the questions are answered here in more depth than the guide gives them, particularly 2.1, 2.3, 3.1, 3.2 and Part 4.
- **Exercise.** The self-study document distinguishes six kinds of value — normative, subjective, behavioural reward, motivational incentive, need-based, and experienced. Take three studies from today — **Plassmann (2007)**, **Chib (2009)** and **Gottfried (2003)** — and for each say which of the six it actually measured, and how you can tell from the method. They do not all give the same answer, and one of them is clearly different from the other two.
- **Then the reverse:** which of the six did *nothing* today measure? Noticing that gap is worth more than any of the answers.