Hypothesis Generation

Companion to the Week 4 episode. A hypothesis is a predicted answer to your question — a genuine bet, placed before the race.

1. The one idea

A good hypothesis is a **falsifiable, directional prediction, tied to a theory, specific about how big an effect would matter** — and written **before** you see the data.

2. Four properties of a good hypothesis

Property	Test	Example
Falsifiable	Can you state now the result that would make you say "I was wrong"?	"Creatine will *improve* sprint time" (slower/no-change = wrong). Not "creatine will *affect* performance".
Tied to theory (Task 1)	Derived from a mechanism?	Creatine → ↑phosphocreatine → faster ATP resynthesis in short maximal efforts.
Directional	Does theory predict which way?	"faster" (risky) beats "will differ" (confirmed either way) — *if* predicted in advance.
Magnitude (Task 2)	How big an effect would *matter*?	A pre-set smallest effect of interest — e.g., a competitively meaningful sprint gain (→ Week 7 + SESOI toolkit).

3. The null hypothesis is a tool, not a belief

The **null** (usually "no effect") is a *reference strawman*: the test asks "if nothing were going on, how surprising would our data be?" No one is emotionally invested in it. Understanding it as a *tool* (not "the thing I'm rooting against") is the first step to using statistics thoughtfully (Week 8). Caveat: an *exactly-zero* effect is rarely literally true — which is why **magnitude** matters.

4. The habit that wrecks it — HARKing (Kerr, 1998)

HARKing = *Hypothesizing After the Results are Known*: presenting a hypothesis you formed *after* seeing the data as if you'd predicted it all along.

Honest (confirmatory)
1. Predict "HRV → performance" → 2. Collect data →

HARKed
1. Collect 15 variables → 2. Correlate all vs

3. Test that one prediction, once → 4. Report result at face value.

performance → 3. Find the one chance "hit" → 4. Write intro as if *that* was the prediction. **Hides the 14 you ignored.**

Why it's damaging: the statistics assume **one** prediction tested **once**. HARKing makes a lucky pattern look like a survived prediction — it looks far more trustworthy than it is, and doesn't replicate. Kerr (1998): widely practised, widely seen as improper. **Fix:** predictions *before* data; discoveries *labelled exploratory* (→ Week 6, preregistration).

5. Name the alternatives (Task 3)

If sprinters improve, could it be…	Ruled out by…
A training effect over 4 weeks	A comparison group doing the same training
Placebo / expectation	Blinded placebo
Regression to the mean (picked slow starters)	Randomisation; not selecting on baseline

6. Do this — write your hypothesis

Hypothesis specification

1. Falsifiable + directional: _____

2. Tied to a mechanism (from your Week-2 theory): _____

3. Magnitude — smallest effect that would matter: _____

4. Main alternatives + how you'd rule each out: _____

Write it so specific that a stranger could see exactly what would prove you wrong.

References

Kerr, N. L. (1998). HARKing: hypothesizing after the results are known. *Personality and Social Psychology Review, 2*(3), 196–217.

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*(6), 632–638.

Lakens, D. (2022). Sample size justification. *Collabra: Psychology, 8*(1), 33267. [vol/page to confirm]