A. Evaluating Existing Research

• Be skeptical but not cynical when reading published research. Poor study design and selective reporting are common features that, along with inadequate statistical power and many forms of bias, decrease the reliability and replicability of much published research. • Be wary of studies that (i) are published in predatory journals; (ii) emphasize statistical over biological significance, or *P*-values over effect sizes; (iii) report results from analyses that were unplanned or not reproducible; (iv) draw causal inferences from non-experimental research without sufficient evidence; (v) reference primarily supportive works while downplaying/ignoring contrary findings; or (vi) fail to discuss study limitations or alternative interpretations of their findings. These points should also inform **B-D.** • Distinguish exploratory (hypothesis generating) from confirmatory (hypothesis testing) research. Given its crucial role, exploratory research (ER) should be common and discoverable, but research culture favours confirmatory research (CR), causing much ER to be disguised as CR, raising the risks of hindsight bias and Hypothesizing After Results are Known (HARKing), and decreasing research reliability. • Pilot studies often blend ER with CR, and commonly deviate from their intended use, reducing their reliability. Interpret them carefully. • P-values are most meaningful in pre-registered CR addressing plausible hypotheses, and least meaningful in ER disguised as CR. • Consider conducting a systematic review of published research, including critically evaluating risk of bias following standardized protocols.

B. Planning Your Research

• Consider research partnerships from the outset Ideally, those who could benefit from or be impacted by the research would be directly involved through respectful collaboration. • Complete ethics approvals and initiate a data management plan (DMP) prior to any data collection, ensuring ongoing respect for and adherence to community-specific engagement protocols, data sovereignty policies, and data management procedures, as per FAIR and CARE guidelines. • Consider using version-controlled reporting systems. • Keep track of all contributors and their contributions from the outset; consider using Dragon Kill Points, or the MeRIT or CRediT frameworks for reporting contributions to project outputs. Encourage all contributors to acquire an ORCID to facilitate discoverability and accurate attribution. • Specify whether the research is ER or CR (or includes both), and provide rationale(s). • Clearly define questions and/or hypotheses, and devise an appropriately designed observational or experimental study; unless subjects/units are randomly assigned to treatments, the study is observational or quasi-experimental. • Ethical and/or logistical constraints often necessitate observational or quasi-experimental designs, in which case causal inference is more challenging (compared to experiments) but possible, requiring careful pre-planning, and meeting myriad criteria and assumptions. • For experimental studies, ensure proper controls and randomization, and implement blinding where possible. • For all studies, define the target population, scope of inference, dependent and independent variables, units of observation, and replication (sample sizes). • Choose statistical methods best suited to addressing the question(s) / hypotheses; new methods are rarely required. • Multiple alternative data pre-processing and/or analysis methods may legitimately be suitable, so consider evaluating the robustness of the findings to these varied protocols (i.e. a form of "multiverse" or sensitivity analysis). • Plan to accommodate non-Gaussian error distributions, multivariate responses and/or hierarchical/non-independent sampling designs where appropriate. • Specify how model assumptions will be checked and violations dealt with. • For CR, clearly define what would constitute a meaningful effect size, and specify what will constitute evidence (i) consistent with and (ii) contrary to each hypothesis.• For CR, consider prospective power analyses that employ the planned statistical methods and plausible effect sizes. • Especially when power is limited, acknowledge uncertainty and focus on sound study design and data useability. • For CR, plan to adjust *P*-values for multiple testing when appropriate. • For ER, justify sample sizes based on feasibility, resource constraints, or desired precision. • Especially for CR, strongly consider detailing your study design and analysis plan in a pre-registration (PR), which can help improve research reliability. PRs shift effort appropriately towards the planning stage, where it has the greatest benefit. Deviations from PRs are absolutely okay if transparently reported (see section C). • Have your research plan reviewed by both subject matter experts and others (possibly as a registered report). • Foster a peer community that values providing and receiving constructive criticism.

C. Conducting Your Research

• Careful planning (part B) helps make the "conducting research" stage smoother. • Document procedures transparently (ideally with versioning), detailing all deviations from planned protocols. Data quality control and updates to data management plans are ongoing responsibilities.

D. Reporting, Interpreting, and Communicating Your Research Findings

• Transparently report all evaluations of model assumptions, and remedies to violations, as per analysis plans. • Prioritize reporting and interpreting effect sizes, their uncertainty (e.g. confidence intervals), and their real-world relevance. • When appropriate, report effect sizes alongside *P*-values. • Absence of evidence (i.e. a statistically non-significant result) is not necessarily evidence of absence. • Statistical significance does not necessarily imply biological relevance. • Do not imply causation with words like "cause" or "affect" unless appropriately justified (e.g. from experimental results); consider whether "is associated with" is more suitable. • Use effective visualizations: reveal, don't conceal data. • Clearly report sample sizes for all treatments, each analysis, and in each relevant figure and table. • Maximize the accessibility and inclusivity of outputs. • There should be no surprise analyses: each analysis should be linked to a pre-specified question/hypothesis. • Adding unplanned analyses is sometimes warranted (e.g. as a follow-up to planned), but these tend to yield inflated false positive rates (e.g. due to researcher degrees of freedom), so label them clearly as unplanned; only planned analyses retain the intended statistical relevance. • Interpret findings objectively and discuss alternative interpretations where appropriate. • Avoid using persuasive language, overselling, or over-generalizing your findings. • Do not extrapolate findings beyond the target population without sufficient justification. • Speculation can be valuable when labeled clearly as such. • Evaluate your draft manuscript against points in section "**A**" above.

E. Enhancing the Replicability, Discoverability, Accessibility, and Reach of Your Research

• Ensure data and executable analysis code are as open as possible and as closed as necessary according to FAIR and CARE guidelines, archived with appropriate metadata, and shared to facilitate independent replication. • Publish Open Access and/or a preprint, and follow best practices to maximize discoverability when writing your title, abstract, and keywords.

These are guidelines not requirements: every step taken towards more reliable, reproducible research is valuable and should be celebrated.