EGAP Learning Days Research Design Form

## Section 1: Introduction	
1. **Researcher name**	
2. **Research project title**	Short title
3. **One sentence summary of research question**	Jargon free
4. **Substantive motivation: [half page]**	Why should anyone care about the results of this research? Does your topic address specific policy concerns?
5. **Theoretical motivation [half page]**	What theoretical questions can this research shed light on?
6. **Key literatures**	List 3 or 4 readings that this work will speak to
7. **Primary hypothesis**	List 2 or 3 hypothesis
8. **Primary causal effect**	What is the primary causal effect of theoretical interest? (e.g., could be a common estimand like the average treatment effect, or something else)

## Section 2: Sample	
9. **Where**	Where and when will your study take place?
10. **Units**	Who/what are the units of measurement in your study? How many such units are in your study?
11. **Sample selection**	How is this sample selected?
12. **Ethics**	Are there any ethical concerns about who is in the sample or who may not be able to consent to being in the study (e.g., prisoner populations)?
13. **Subgroups**	Do you expect the intervention to work differently for certain subgroups?

## Section 3: Intervention	
14. **Interventions**	Describe your intervention(s) Single or multiple intervention arms?
15. **Control condition**	What will your control condition be? Will it be a pure control or placebo?
16. **Ethics**	Are there any ethical concerns with the intervention?

17. Level	At what level will you randomize the intervention?

## Section 3: Outcome	
18. Outcomes	What is your primary outcome?
19. Data	How will you measure it? What data do you need? (e.g., administrative data, surveys, other)
20. Levels	At what level is or will be the measure available? (e.g., individuals, villages, schools)
21. Priors	What are your priors about the outcome (mean, SD)? This may come from previous studies or educated guesses.
22. Rounds	How many rounds of data collection will you conduct?
23. Attrition	How will you minimize attrition?
24. Measurement	How will you minimize mismeasurement and untruthful reporting?

## Section 4: Identification Strategy / Causal Inference Strategy	
25. Identification	What type of identification strategy will you use? (ex. What type of random assignment such as simple; complete; blocked; cluster; factorial two level; phase-in; waiting-list?)
26. Blocks	How many blocks and what are they (if any)?
27. Clusters	How many clusters and what are they (if any)?
28. Interference	Is interference a possible concern? If so, what is your plan for minimizing or managing or learning about interference?

## Section 5: Implementation	
29. Randomization	If applicable: How will you do the actual randomization? In public, drawing from a bowl, on a computer?
30. Implementation	Who will implement the intervention?

31. Quality	How will you track the quality of implementation?
32. Compliance	How will you track compliance with the intervention? How will you minimize non-compliance with the intervention (if applicable)
33. Data management	How will data be anonymized and securely stored (if applicable)?

## Section 6: Power	
34. Effect Size	What is your expected effect size? What effect size do you want your experiment to show? What effect sizes have similar studies found? This might be from a previous study or a target size below which one would not be interested in future interventions.
35. Intra-Cluster Correlation (ICC)	If you have clusters, what is the intra cluster correlation? This requires insights from previous studies or representative data
36. Power Calculation	What is your power? If you want to calculate your sample size, given expected effect: use STATA/R. If you want to calculate effect size, given a maximum sample: use STATA/R. Take into account type outcome (binary, continuous) and clustering (ICC) if applicable.

## Section 7: Analysis & Threats	
37. Analysis strategy	What is your estimator? (e.g., difference in means, OLS with block weights, any clustering). If you plan to report confidence intervals, what kind of standard errors will you calculate? (e.g. HC2/Neyman SEs? CR2 cluster robust SEs? Etc.) If you plan to report a p-value, what kind of test will you use? Note that this should be closely linked to your design.
38. Interpretation strategy	If you find that your results are consistent with your hypothesis, what alternative explanations might there be? What data would help you distinguish between your explanation and alternative ones? Make sure you have a plan to collect this data.

	If you find that your results are not consistent with your hypothesis, what data will help you figure out what might have happened? Make sure you have a plan to collect this data.