# Pre-analysis plan: Design of the YourSuper comparison tool

*Note: The original version was registered on 8 January 2021. This version was submitted for pre‑registration on 13 January, before any survey data was viewed. This version amends the original in three ways:*

1. *Respondents who complete the survey in under 2 minutes and 30 seconds will be dropped from the sample prior to analysis. (A robustness check that retains these records will also be performed.)*
2. *The analysis of a second experiment within the survey was detailed.*
3. *The analysis model was revised to include an interaction term between the treatment arm and the covariate (level of education).*

## Policy problem, trial aims and research question

As part of the *Your Super, Your Future* superannuation reforms outlined in the 2020 Budget, the Australian Government has committed to the creation and implementation of an online superannuation comparison tool. The purpose of this tool is to make it easier for consumers to choose a well-performing MySuper product that suits their needs. The aim of this trial is to investigate how the display of superannuation information (specifically, information about performance and risk) influences comprehension and engagement with the tool.

This study utilises a survey to administer two experiments presenting hypothetical scenarios. Respondents will be asked to:

1. Choose the most appropriate fund for a person in a presented vignette. The list of funds will be presented to mimic the online tool, with differences in the presentation of available superannuation funds. We will test a list of 10 fictitious superannuation funds with different net return levels.
2. Select the most appropriate investment risk option for a young person starting off in their career, with differences in the description of risk options. Three investment risk options (covering low, medium and high risk and reward settings) will be presented with different label sets.

## Outcome measures

**Experiment 1**

Primary outcome 1 – *Choice of fund*

This will be an ordinal variable based on the answer provided when choosing a super fund from the list. Each fund will be given a ranking ranging from 1-10, with 1 being the worst performing fund (lowest net return), and 10 being the best performing fund (highest net return). For the analysis this will be treated as a continuous variable.

Primary outcome 2 – *Time to complete*

This will be a continuous variable based on how long they spent on the experiment page (measured in seconds). There is a chance that a participant could be interrupted and take very long to complete this section of the survey. In order to ensure this doesn’t impact our results, we will remove anybody who is more than 3 standard deviations above the mean.

Secondary outcome 1 – *Choice of ‘Top’ fund*

This will be a binary variable indicating if they chose one of the three funds with a performance rating of ‘Top’.

Secondary outcome 2 – *Choice of any fund that is not ‘poor’ performing*

This will be a binary variable indicating if they chose a fund with either a ‘Top’, ‘Fair’ or ‘Good’ performance rating.

Secondary outcome 3 – *Ease Rating*

This will be an ordinal variable measured through a 1-5 Likert scale rating of how easy it was to complete the task. For the analysis this will be treated as a continuous variable.

Secondary outcome 4 – *Intention to switch*

This will be an ordinal variable measured through a 1-5 Likert scale rating of their likelihood to switch funds if faced with that same scenario. For the analysis this will be treated as a continuous variable.

**Experiment 2**

Primary outcome – *Selection of correct investment risk answer*

This will be a binary variable indicating if the respondent selects the most appropriate investment risk option for a young person starting their career (i.e. the higher risk, higher reward option).

## Interventions and hypotheses

**Experiment 1**

Experiment 1 will test two primary conditions: the number of categories the participants see and the sort order of the presented funds. It will be a three-arm trial with differences in the number of categories and sort order. It will include the following groups:

* 1. Group A: two categories (Good and Poor performance), randomly sorted
	2. Group B: three categories (Top, Fair and Poor performance), randomly sorted
	3. Group C: three categories, sorted by performance category (then randomly sorted within each category)

Hypothesis 1.1: On average, respondents in group B will choose a higher rated fund as compared with Group A (Group A < Group B).

Hypothesis 1.2: On average, respondents in group C will choose a higher rated fund as compared with Group B (Group B < Group C).

Hypothesis 1.3: On average, respondents in group C will choose a higher rated fund as compared with Group A (Group A < Group C).

Hypothesis 1.4: On average, respondents in Group B will take less time to make a decision than Group A (Group A > Group B).

Hypothesis 1.5: On average, respondents in Group C will take less time to make a decision than Group B (Group B > Group C).

Hypothesis 1.6: On average, respondents in Group C will take less time to make a decision than Group A (Group A > Group C).

Our interpretation of Hypothesis 4, 5 and 6 will be informed by the secondary outcome regarding ease of decision making. This is because of the potential for our objective measure of timing to be confounded by measurement error. We will refer to the subjective measure of self-reported difficulty making a decision to draw conclusions on the findings.

**Experiment 2**

Experiment 2 will test whether the selection of the most appropriate investment risk option varies by the terms used to describe the investment risk profile. It will include the following groups:

1. Group A: Risk level noted, with a statement on the expected frequency of negative returns over a 20 year period (e.g. ‘Low risk – expect a negative return 1 out of every 20 years’)
2. Group B: Investment strategy noted using common terms in the superannuation sector (‘Conservative’, ‘Balanced’, ‘Growth’)
3. Group C: Risk and expected return levels noted (e.g. ‘Low risk & low expected return’)

Hypothesis 2.1: On average, respondents in group B will select the most appropriate response at a different rate to those in Group A (Group A ≠ Group B).

Hypothesis 2.2: On average, respondents in group C will select the most appropriate response at a different rate to those in Group B (Group B ≠ Group C).

Hypothesis 2.3: On average, respondents in group C will select the most appropriate response at a different rate to those in Group A (Group A ≠ Group C).

## Sample size and power calculations

*Note: calculations were performed to assess only the minimum detectable effect (MDE) of Experiment 1.*

For the power calculations we set alpha at 0.1 as we are less concerned with a false positive. The sample size is capped at approximately 2,000 participants. We conducted sensitivity analysis to find the MDE. All analyses were conducted in RStudio (R version 4.0.2).

For all hypotheses we will conduct one-sided tests. It is unlikely that changing categories or sorting will result in participants selecting a worse performing fund or taking longer to make a decision. Our calculations suggest that with N = 2000, we will have 95% power to detect an effect size of *d* = 0.16. For fund choice, this equates to approximately 0.3 rank change assuming a pooled standard deviation of around 3 rank scores. For time to complete, this is a difference of less than a second assuming a pooled standard deviation of around 15 seconds.

## Trial design

*Note: trial design considerations are the same across Experiments 1 and 2, unless otherwise noted.*

**Sample**

We will recruit approximately 2,000 people from the survey panel provider Dynata. Participants will be eligible if they are Australian residents with a superannuation account. They must also be between the ages of 18 and 64. The sample will also have age, gender and metro-regional splits that are nationally representative. We will exclude from analysis any responses that have taken less than 2 minutes and 30 seconds to complete the survey. This is a conservative threshold below which we do not believe people can have actually read and engaged with the survey. By excluding these responses, it means that we will not be doing an intent to treat analysis. We will not be implementing any other exclusion criteria.

**Randomisation**

For each experiment, participants will be randomised to a trial-arm within the survey platform, Qualtrics, with roughly equal allocation to each condition. Note that assignment to a given group in Experiment 1 will not carry through to Experiment 2 as a second randomisation will occur. We will conduct a joint orthogonality test as a balance check for each experiment. If there are differences in demographic characteristics between groups we will check that the randomisation was conducted appropriately.

**Trial threats**

There is a chance that some participants will fail to complete the survey and we will therefore have missing data for the primary outcomes. In that case we will exclude missing cases to conduct a complete case analysis.

This trial is a survey experiment and will therefore have limited generalisability. While the results will be used to inform policy, the interventions may not have the same effect in real-world settings as any found in this project.

**Consort flow diagram**

Note: the Consort flow diagram is only provided for Experiment 1. The same sequence and observation numbers apply to Experiment 2.

Analysis (n=)

Analysis (n=)

Analysis (n=)

Lost to follow-up (give reasons) (n=)

Discontinued intervention (give reasons) (n=)

Lost to follow-up (give reasons) (n=)

Discontinued intervention (give reasons) (n=)

Lost to follow-up (give reasons) (n=)

Discontinued intervention (give reasons) (n=)

Group B: Three categories, randomly sorted (n~667)

Group C: Three categories, sorted by performance category (n~667)

Dynata to assess for eligibility (n=2,000)

Group A: Two categories, randomly sorted (n~667)

Enrolment

Randomisation (allocation 1:1:1)

Follow-up (finishing survey)

Analysis

## Method of analysis

**Experiment 1**

*Primary Analysis*

The primary analysis will use ordinary least squares (OLS) regressions to estimate the effects of our intervention. The following model will be used:

$$Y= α+ βT\_{b,c}+ δX+ΩT\_{b,c}X+ ε$$

Where $Y$ is the primary outcome variable, $T\_{b,c}$ is a binary variable indicating if the individual was in treatment arm *B* or *C*, $X$ is a binary covariate for education (set to 1 if they have university training or higher, and 0 if they have below university training) and $T\_{b,c}X$ is an interaction between treatment arm and education binary.

* For hypotheses 1.1 & 1.4, the base group will be arm A, and the treatment group will be arm B.
* For hypotheses 1.2 & 1.5, the base group will be arm B, and the treatment group will be arm C.
* For hypotheses 1.3 & 1.6, the base group will be arm A, and the treatment group will be arm C

There may be exploratory analysis related to H4, H5 and H6, which restricts the analysis to certain cohorts of respondents.

The primary analysis will exclude those that complete the survey in under 2 minutes and 30 seconds. A robustness check will be performed to test the effect of their inclusion.

*Secondary Analysis*

The secondary analysis will use the same model as above with Y being the secondary outcome measure of interest.

In addition to the robustness check noted in the primary analysis, we will also conduct robustness checks using a logit model for secondary outcomes 1 and 2.

**Experiment 2**

*Primary Analysis*

The primary analysis will use ordinary least squares (OLS) regressions to estimate the effects of our intervention. The following model will be used:

$$Y= α+ βT\_{b,c}+ δX+ΩT\_{b,c}X+ ε$$

Where $Y$ is the primary outcome variable, $T\_{b,c}$ is a binary variable indicating if the individual was in treatment arm *B* or *C*, $X$ is a binary covariate for education (set to 1 if they have university training or higher, and 0 if they have below university training) and $T\_{b,c}X$ is a two-way interaction between treatment arm and education binary.

* For hypothesis 2.1, the base group will be arm A, and the treatment group will be arm B.
* For hypothesis 2.2, the base group will be arm B, and the treatment group will be arm C.
* For hypothesis 2.3, the base group will be arm A, and the treatment group will be arm C.

Primary analysis will exclude those that complete the survey in under 2 minutes and 30 seconds. A robustness check will be performed with these observations included. An additional robustness check will be performed with a logit regression model used in the place of OLS.

**Sub-Group and Exploratory Analysis**

We will also be conducting sub-group and exploratory analysis. The exploratory analysis will involve looking at responses to questions motivating levels of different campaign slogans and interaction with superannuation comparison tools. We will also explore whether people are choosing the fund that appears first on the list over other better funds.

We may also conduct subgroup analysis on a range of dimensions such as:

* Age
* Gender
* Superannuation balance
* Financial literacy
* Financial risk preference
* Location
* SEIFA index for disadvantage

There will be no set model for this exploratory analysis and we will not be using it to determine causal inference.

## Pre-analysis plan commitments

No analysis has been undertaken prior to the completion of this pre-analysis plan.

We will be transparent about, and provide justification for, any deviations (additions or omissions) from this plan.