

Easy as pie: encouraging greater survey completion – Pre-analysis plan

1. Introduction

We pre-registered this trial on the AEA Social Science Registry on 2 May 2017 and this pre-registration included many details of how we planned to conduct our analysis. While our pre-registration occurred after the launch of the trial on 13 March 2017, it was before we had received any data or commenced any analysis. We subsequently prepared this, more detailed analysis plan, which was finalised on 22 January 2018. This occurred after we had received some trial data and conducted some preliminary analysis. We discuss this in the Technical Appendix to the report. Our trial pre-registration can be found here:

<https://www.socialscienceregistry.org/trials/2110/history/17220>.

2. Policy problem, trial aims and research question

Data Exchange is the new Department of Social Services (DSS) approach to collecting and storing programme performance data. It is intended to shift the focus of performance measurement from outputs to more meaningful information about service delivery outcomes. A Data Exchange survey is offered to clients both at the beginning and end of their service. It is important to collect complete feedback from clients to inform continual program and service delivery improvement, as there are a diverse range of programs covered by the Data Exchange survey that are accessed by many Australians. This study is intended to help solicit client feedback using two behaviourally informed approaches. BETA and DSS inserted behavioural insights into the Data Exchange survey to: (1) increase initial survey completion rate, using pie charts to denote progress, and (2) increase second survey participation rate, using a pre-commitment device in the first survey.

The aim of the study was to examine whether using behavioural insights techniques can improve the data collection rate for the Data Exchange surveys, which in turn will assist DSS with evaluating service delivery outcomes.

Our original preregistered research questions included:

- Does a pseudo-set design encourage survey completion or completion of more questions over control?
- Does a pre-commitment device encourage commencement of a follow-up survey over control?

This detailed analysis plan was not included with the original pre-registration and is intended to supersede the original analysis plan. This plan was written and finalised after receiving preliminary data which resulted in a number of changes to the original plan due to data quality issues. This plan was written before analysis of the final dataset commenced.

3. Trial objectives and outcome measures

3.1 Outcome measures

Our pre-registered primary outcomes were i) proportion of subjects who completed the initial survey, ii) average percentage of initial survey questions completed by subjects, and iii) proportion of subjects who commenced the follow-up survey. Organisational- and individual-level baseline data was collected on: organisation ID, organisation type, organisation outlet ID, organisation outlet postcode, age, sex, main language spoken at home, number of jobs, highest education obtained, household income, and health status.

However, due to data quality issues and a smaller than anticipated sample size we will treat outcome (i) as the primary outcome and all other outcomes as secondary.

4. Hypotheses and interventions

In our original pre-registration plan, we hypothesised:

- H1. Pseudo-set design will perform better than control on the primary outcome measure: proportion of subjects who complete the initial survey.
- H2. Pseudo-set design will perform better than control on the primary outcome measure: average percentage of initial questions completed by subjects
- H3. Pre-commitment will perform better than control on the primary outcome: measure proportion of subjects who commence the follow-up survey.

Because of data quality issues and a smaller than anticipated final sample, we will only assess H1 in our primary analysis and consider H3 in secondary analysis. Incommensurate survey lengths rendered H2 difficult to meaningfully test, and thus we discarded it.

BETA assisted the Department of Social Services in designing the Data Exchange survey as to: (1) increase initial survey completion rate, using pseudo-set framing, and (2) increase second survey participation rate, using a pre-commitment device in the first survey. There were three experimental arms:

- **Control** – Survey alone
 - **Pseudo-set design** – Pie charts that indicate progress in completing survey.
 - **Pre-commitment** – Final survey page incorporating pre-commitment device signal willingness to complete follow-up survey.
-

5. Trial design

5.1 Population and sampling frame

Our study examined the behaviour of clients from participating DSS partner organizations. These clients represented the population of interest, from which a sampling frame was drawn from participating organisations over a 7-month period (211 days, from 13 Mar 2017 to 10 Oct 2017).

These 21 organisations work in the 8 areas of –

- **Housing assistance or homelessness support**
 - Assistance with care and housing
 - Reconnect
- **Carer supports**
 - Care relationships and carer support
 - National carer counselling programme
 - Mental health respite: carer support
- **Children and families**
 - Children and parent support services
 - Children's contact services
 - Family and relationship services
 - FaRS – specialised family violence
 - Communities for children – facilitating partners
- **Financial wellbeing and capability**
 - Commonwealth financial counselling and financial capability
 - Financial counselling, capability and resilience IM hubs
 - Financial crisis and material aid – emergency relief
 - Problem gambling
- **Family law services**
 - Family dispute resolution
 - Family law counselling
 - Family relationship centres
 - Parenting orders programme
 - Regional family dispute resolution
- **Mental health**
 - Personal helpers and mentors
- **Home support and care**
 - Community and home support
- **Settlement services**

5.2 Power

Previous trials of pseudo-sets have reported large increases in completion rates, therefore this study was powered to detect a 10 percentage point increase in survey completion.

Prior to the trial commencing, it was calculated that a total sample of 423 (three groups, 141 units per group) would provide 80% power at a 5% significance level to detect a change in survey completion from 85% to 95% (this reflects a Cohen's d of 0.35, a small to medium effect size).

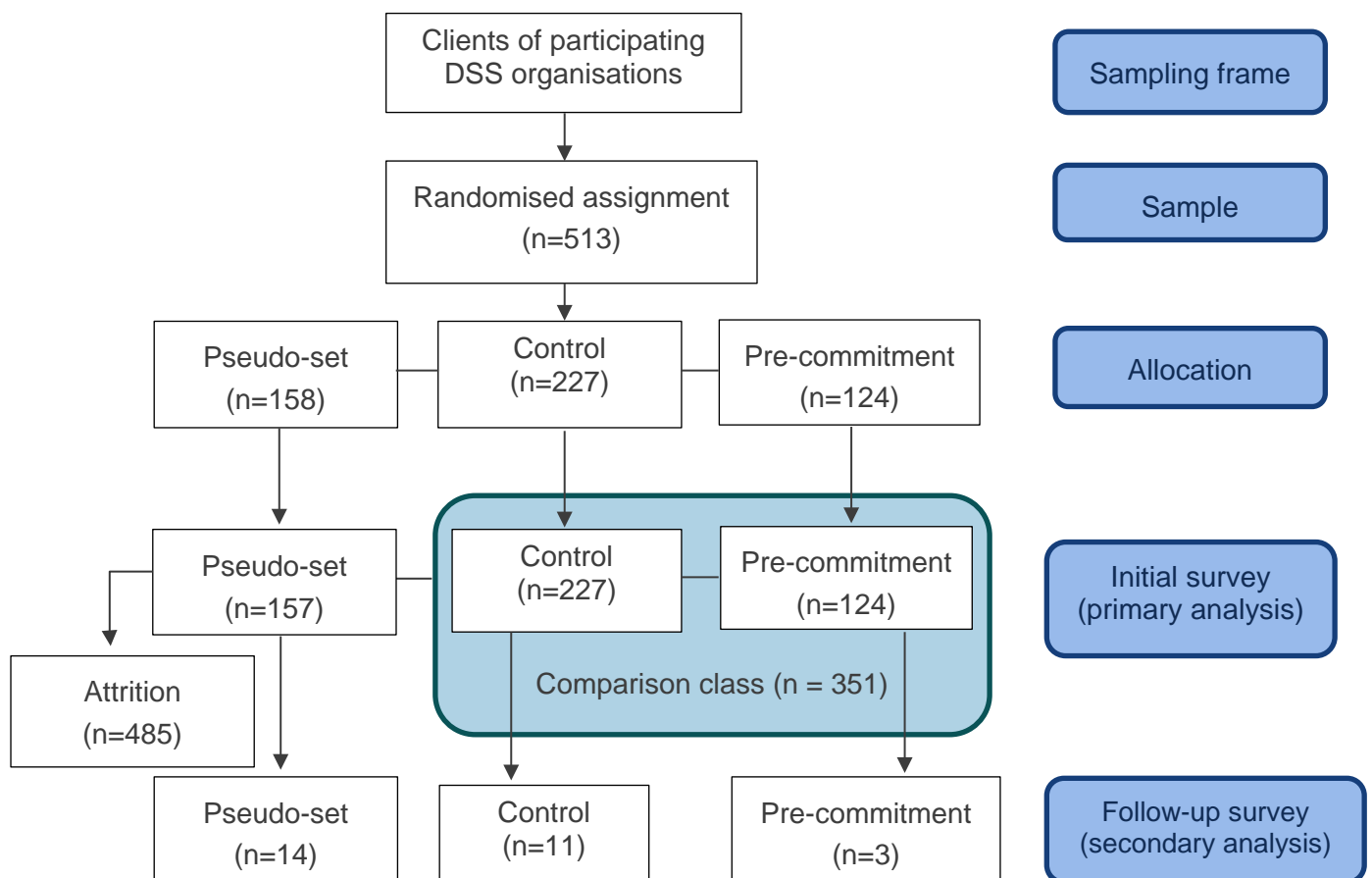
As we specified in our preregistered analysis plan, because our sample size was less than anticipated, we will combine participants assigned to the pre-commitment treatment group into the control group. This will be compared to the pseudo-set group. This is a valid comparison because

the pre-commitment intervention is delivered at the end of the survey after the pseudo-set treatment has been delivered and the relevant outcomes measured.

5.3 Randomisation

Randomisation followed a DSS-developed computer algorithm. The three trial arms (control group and two treatment groups) were intended to be randomised at the individual level using simple randomisation, however, an ordered assignment algorithm at the individual-level was implemented. This allocated individuals to each arm in order of their arrival. The algorithm operated independently within each organisation (see next section).

Figure 1: Trial flow diagram



5.4 Trial threats and design details

Implementation fidelity provided some limitations. These included:

- Randomisation was not truly random but rather ordered (the first person in an organization to take a survey received the control, second person received treatment 1, third person received treatment 2, fourth person received the control, and so on).
- Only a handful of trial participants were observed for several organizations and outlets.
- Ultimately, the trial was stopped and data collection halted indefinitely because three organisations were not complying with the trial protocol.

Three organisations were non-compliant. One organisation was provided a settlement service. The organisation was given permission to sit with clients and translate the questions, as 90% of clients do not speak English. The client survey team was interested in gaining input from a settlement service in the knowledge that only a small amount of surveys would be completed (bilingual worker involvement is resource intensive). IT literacy was also identified as an issue by the organisation, so it is likely that staff also inputted the answers.

A second organisation was involved in community aged-care service. Staff were very supportive of the survey, and managed IT literacy issues of their clients by either inputting answers directly onto devices (while in clients' homes) or phoning clients and completing the survey while asking the client each of the questions.

A third organisation was involved in financial crisis and material aid – emergency relief. It printed out the survey questions and took them into the field (where internet access is not available). Personnel entered response options onto a spread sheet (and were advised by DSS to destroy survey responses once it had become aware of this practice).

6. Qualitative evaluation

No qualitative evaluation was conducted.

7. Analysis

We will use hypothesis testing to elicit evidence of a statistically significant difference between experimental conditions for our one primary outcome measure (completion rates of the initial survey). Namely, we will conduct a logistic regression of the treatment condition (categorical variable) on the primary outcome variable (binary variable).

The model will also be adjusted to include potential covariates. We will *a priori* select covariates in partnership with DSS and use pre-trial baseline data on 116 individuals who completed the survey to confirm that selection. Potential covariates include: main language spoken at home, highest level of education obtained, and health status. Because data for education and health status was collected late in the survey (Q58 and Q41), there is a strong likelihood such missing data for these two measures is correlated with treatment assignment. We will test for such a relationship, and if it exists for either measure, one or both will not be selected as covariates but rather will be included in subgroup analysis. The latter will also focus on age, sex, household income, employment status, and organisation type. All data processing and analysis steps will be performed using STATA script and will involve manual checks at each stage to ensure there are no errors introduced.

We did some preliminary analysis mid-trial, but our decision to stop was based on a power calculation to achieve a pre-determined designated sample size. Consequently, we will not make a correction for optional stopping.

7.1 Balance checks

We will perform balance checks on the pseudo-set group and pooled control/pre-commitment group to judge whether observed covariate imbalances are larger than would normally be expected from chance alone. This will entail regression of the binary treatment indicator on all covariates, using an

F-statistic to gauge if any coefficient is different from zero. The results of the balance checks will be duly reported. However, given the abovementioned limitations of the trial no correction for potential imbalances can be made.

7.2 Hypothesis testing and primary analysis

We will make one comparisons: first testing pie chart against a comparison class (control and pre-commitment) on the first primary outcome and second testing pre-commitment to control on the second primary outcome. We will conduct a logistic regression of experimental condition on the response variable for the primary outcome measure, adjusted for potential covariates. We will also run an unadjusted model as well as a robustness check.

7.3 Secondary analysis

We will test pseudo-set to precommitment and control on a secondary outcome: proportion of subjects who commenced the follow-up survey. This will entail a logistic regression of experimental condition on the response variable for the outcome measure, adjusted for potential covariates.. We will also run an unadjusted model as well as a robustness check.

7.4 Subgroups

A number of secondary subgroup analyses using logistic regressions (comparing treatment effects and interaction effects) will also be performed and will be considered exploratory. We will also do subgroup analysis on sex and age (in which age will be aggregated into 15-year age groups).

7.5 Interpretation of results

In interpreting our results, we will accommodate the principles of the *ASA Statement on Statistical Significance and P-Values* to the extent possible while still reporting our results against the conventional threshold for statistical significance ($p < 0.05$). For example, we are mindful that statistical significance is different from practical significance (ASA Statement, Principle 5). Our estimates for the effects of treatment on, for example, survey completion may be large but not statistically significant. If these estimates reflect a true effect, they are highly material for government policy and so should not be overlooked. But nor will we overlook statistical significance in reporting such results with extreme caution.

7.6 Robustness of analysis

Most data was reported by DSS partner organisations, indicating that missing data was possible for six baseline variables (age, sex, employment status, household income, health status, main language spoken at home, highest education obtained). We will test the overall rates of missing data as well as for any patterns in missingness across experimental conditions. Thus we will conduct an intention-to-treat analysis on those variables for which it is possible and duly report all robustness checks and limitations.

With regard to exclusion of collected data from non-compliant organisations, we will run the analysis without any exclusions followed by a secondary analysis with exclusions. We will report the number

of affected clients and test for any skewness of sample size and other data across experimental conditions that can be attributed to excluded clients.

7.7 Outcome tables

Baseline characteristics and balance

		Control (n =)	Pseudo-set (n =)	Pre-commitment (n =)
Age (years)				
Sex (female)				
Employment status (employed)				
Main language at home (English / not answered)				
Org. type	Housing assistance Carer support Children/family Financial wellbeing Family law Mental health Home support Settlement			
Qualification	Adv. Diploma Bachelor deg. Cert. I/II Cert III/IV. Diploma Grad. Diploma Other non-school Postgrad. Deg. Not answered			

		Control (n =)	Pseudo-set (n =)	Pre-commitment (n =)
	N/A			
Household income	Negative or \$0 \$1-9,999 \$10K-\$19,999 \$20K-29,999 \$30K-\$39,999 \$40K-\$49,999 \$50K-\$59,999 \$60K-\$69,999 \$70K-\$79,999 \$80K-\$99,999 \$100K-\$124,999 \$125K-\$149,999 \$150K-\$199,999 \$200K+ Not answered Not presented q			
Health status	Excellent Very good Good Fair Poor Not answered Not presented q			

Main effects analysis

Outcome	Experimental condition	
	Intervention class – pseudo-set group (n =)	Comparison class – control + pre-commitment groups (n =)
Overall number of surveys completed	Raw percentages Odds ratio (CI, and p-values) Predicted probabilities (with CIs)	Raw percentages Odds ratio (CI, and p-values) Predicted probabilities (with CIs)

Secondary analysis

Outcome	Experimental condition	
	Intervention – pre-commitment device (n =)	Control (n =)
Number of commenced “Survey 2”	Raw percentages Odds ratio (CI, and p-values) Predicted probabilities (with CIs)	Raw percentages Odds ratio (CI, and p-values) Predicted probabilities (with CIs)

8. Pre-analysis plan commitments

If our final report contains analyses that deviate from this plan we will make it clear that these analyses not pre-specified and provide justification for them. Conversely, if we omit pre-specified analyses we will make these available as supplementary material. We will be transparent about, and provide justification for, any deviations (additions or omissions) from this plan.