Pre-analysis plan: Web content advice survey experiment

This pre-analysis plan was finalised and pre-registered on 06 April 2020, after the trial was launched on 04 March 2020 but before the receipt of any data.

Policy problem, trial aims and research question

In partnership with the Australian Cyber Security Centre (ACSC), the Behavioural Economics Team of the Australian Government (BETA) is conducting research to improve cyber security advice for individuals in their personal lives.

Individuals can protect themselves and reduce their risk of becoming victims of cyber attacks by implementing certain behaviours. In particular, using strong and different passwords across important accounts, and regularly updating their devices' software.

The study involves a survey and, embedded within that, a survey experiment. The survey itself aims to gain deeper insight into the attitudes towards, awareness of, and current practices in cyber security of Australians. The embedded survey experiment aims to examine which ways of presenting information might result in the most significant change to behavioural intentions, or actual behaviours.

A follow-up survey will go to all participants and assess actual behavioural change for passwords and updates. Both the initial and follow-up surveys will be conducted through an online survey platform with Australian Survey Research.

Interventions

This study effectively involves two separate components however the interventions for each component have a similar design. Participants will be exposed to advice relating to (a) improving their password security and (b) software updates. There are six variations to the advice, following a factorial design:

- Messenger: They may see this advice in the form of an attention control (no messenger), via a 'peer' messenger, or via an 'expert' messenger.
- Consequences (financial/non-financial): They may also see the advice framed around financial gains and losses, or around the impacts that their suboptimal behaviour may have on non-financial aspects of their lives.

Outcome measures

We have four primary outcomes for each experiment:

1. knowledge (at the time of exposure),
2. knowledge (two weeks later),
3. self-reported behavioural intentions (at the time of exposure),
4. self-reported behaviours (two weeks later).

The details of how each outcome will be measured for each experiment are set out in the tables below.

Some outcomes will be treated as continuous for the purpose of analysis even though they are measured on a scale. In these cases, we will provide context by also reporting descriptive cell percentages.

Table 1. Outcome measures: password security

Outcome measure & source	Question	Response options
Knowledge at exposure: main survey (We will only use the answer for the second password (the passphrase). It will be treated as binary, with 'Very strong' or 'Strong' coded as 'correct'.)	How do you rate the strength of these passwords? (password1, **fieldhayfaretoss**, wjh63m&92mk11gr9)	1. Very strong 1. Strong 0. Weak 0. Very weak
Knowledge 2 weeks later: follow-up survey (We will only use the answer for the second password (the passphrase). It will be treated as binary, with 'Very strong' or 'Strong' coded as 'correct'.)	How do you rate the strength of these passwords? (Tuesday25, **trendagepairdeer**, n8j2n3wzhz3edygs)	1. Very strong 1. Strong 0. Weak 0. Very weak
Self-reported behavioural intentions: main survey (Indexed from summing the responses to the two questions, each on a 0-4 scale, giving a maximum possible score of 8.	How likely are you to create **strong** passwords for your important accounts (such as your online banking, email, and social media accounts)? How likely are you to create **different** passwords for your important accounts (such as your online banking, email, and social media accounts)?	4. Extremely likely 3. Very likely 2. Moderately likely 1. Somewhat likely 0. Not at all likely

Outcome measure & source	Question	Response options
Self-reported behaviours: follow-up survey (Indexed from summing the responses to the two questions, each on a 0-3 scale, giving a maximum possible score of 6.)	In the last two weeks, did you create **strong** passwords across your important accounts (such as your online banking, email, and social media accounts)? In the last two weeks, did you create **different** passwords across your important accounts (such as your online banking, email, and social media accounts)?	3. Yes for ALL of my important accounts 2. Yes for MOST of my important accounts 1. Yes for SOME of my important accounts 0. No

Table 2. Outcome measures: software updates

Outcome measure & source	Question	Response options
Knowledge at exposure: main survey (This will be treated as binary, with the final option coded as 'correct'.)	When you receive a notification to update software on your personal device, does it matter how soon you update it? Select the best answer:	0. No, as long as you update eventually 0. No, as long as you update within a week 0. Yes, you need to update it within 24 hours 1. Yes, the longer you wait the more vulnerable you are
Knowledge 2 weeks later: follow-up survey	When you receive a notification to update software on your personal device, does it matter how soon you update it? Select the best answer:	0. No, as long as you update eventually 0. No, as long as you update within a week 0. Yes, you need to update it within 24 hours 1. Yes, the longer you wait the more vulnerable you are
Self-reported behavioural intentions: main survey (We will treat this variable as continuous.)	When prompted on a personal device, how likely are you to update the software immediately?	4. Extremely likely 3. Very likely 2. Moderately likely 1. Somewhat likely 0. Not at all likely

Outcome measure & source	Question	Response options
Self-reported behaviours: follow-up survey (We will treat this variable as continuous. People who have not received a notification in the last two weeks will be coded as 0, the same as 'Haven't done the update yet')	How long after you got the update notification did you do the update? (If you got more than one update, think of the last one you received.) Note: The preceding question is: 'In the last two weeks, have you received a notification to update your software or your computer, laptop, tablet or mobile phone?' Respondents are shown the next question if they respond 'yes'	4. Immediately 3. Within 2 days 2. Within 7 days 1. More than 7 days later 0. Haven't done the update yet

Hypotheses

Since the interventions for both experiments (password security and software updates) have the same structure, we have the same hypotheses for both. As indicated below, directional hypotheses will be tested using a one-sided test; non-directional hypotheses will be tested using a two-sided test.

Messenger conditions

H1a-H1d: The four outcomes will be higher among respondents exposed to **any** messenger condition (pooled) compared to the attention control condition (one-sided test).

H2a-H2d: The four outcomes will be higher among respondents exposed to **each** messenger condition compared to the attention control condition (one-sided test).

H3a-H3d: The four outcomes will be different among respondents exposed to the peer messenger condition compared to the expert messenger condition (two-sided test).

Financial consequences condition

H4a-H4d: The four outcomes will be different among respondents exposed to the financial consequences condition compared to the non-financial consequences condition (two-sided test).

Trial design and randomisation

This experiment has a factorial design. Participants saw advice on two cyber security behaviours (two separate experiments). They saw:

1. password security advice; then
2. software update advice.

In each experiment, all participants were randomised into one one of six possible cells based on a combination of either financial or non-financial consequences, and one of the three messenger arms.

Participants were initially randomised at an individual level for allocation to the password security experiment (A1 through A6). All participants were then re-randomised at an individual level for allocation to the software update experiment (B1 through B6). In both cases, the allocation ratio will give an equal number in all six cells. Randomisation into B1 through B6 was blocked on randomisation to A1 through A6 (see Table 3).

Note that randomisation took place in advance using a larger sample frame of 20,000 participants, but that data collection ended once 4,500 responses were collected. We don't have control which 4,500 people respond, so numbers presented below for each group are approximate rather than precise.

Table 3. Randomisation

(N= 4500, deterministic, participants randomised at an individual level and sorted into one of 36 possible pathways)				
Passwords	**Financial** (N=2,250)	**A1** Attn. Control N=750	**A2** Expert N=750	**A3** Peer N=750
	Non-Financial (N=2,250)	**A4** Attn. Control N=750	**A5** Expert N=750	**A6** Peer N=750
(Interstitial questions relating to physical security habits)				
Updates	**Financial** (N=2,250)	**B1** Attn. Control N=750	**B2** Expert N=750	**B3** Peer N=750
	Non-Financial (N=2,250)	**B4** Attn. Control N=750	**B5** Expert N=750	**B6** Peer N=750
Main survey: knowledge & intention outcomes				
Follow-up survey (2 weeks later): knowledge & behaviour outcomes				

Sample selection and exclusion criteria

Participants will be sourced through a recruitment panel, which will aim to recruit a sample that is representative of the general population in Australia. To ensure balance across age, gender, and location, potential respondents must declare their gender, age bracket, and state, before they are permitted into the survey. If the quota for their age+gender+state has been filled, they will not be able to proceed with the survey. For those who do complete the survey, this initial demographic information will be included in their response data.

Sample size and power calculations

Our sample will be 4,500 participants, with the sample size dictated by budget considerations. Participants were recruited through *Australian Survey Research*, who endeavoured to ensure that the sample was representative of the larger Australian population.

We adapted baseline prevalence levels for behavioural intentions from the control group of a related RCT conducted on small-to-medium business operators.

Although there will be attrition between the main survey and the follow-up survey, we believe we will still have sufficient sample size to make meaningful inferences about behaviour based on follow-up responses.

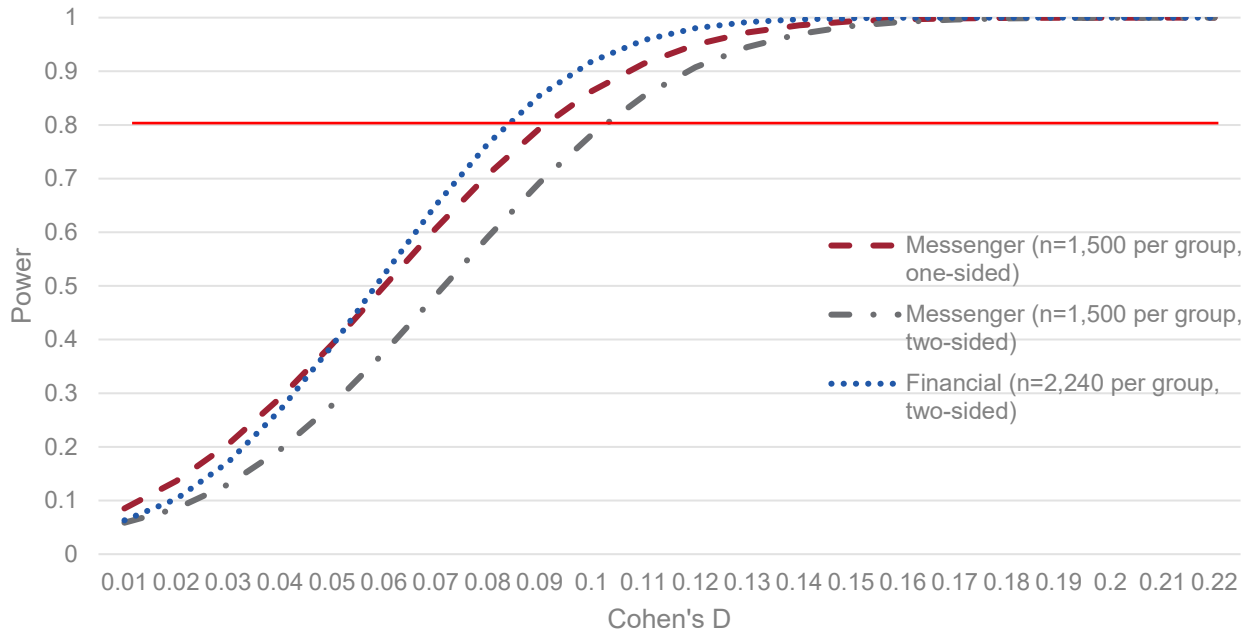

Figure 1: Minimum detectable effect size power curves

The following table indicates minimum detectable effect sizes for each outcome measure, based on an alpha of 0.05 and 80% power. These power calculations also assume we have a full sample of 4,500 for the initial measures.

Table 4. Minimum Detectable Effect, Power= 0.8, Alpha=0.05

Messenger (H1a-b & H2a-b, one-sided) N=1,500 per group	Messenger (H3a-b, two-sided) N=1,500 per group	Financial (H4a-b, two-sided) N=2,250 per group
0.09	0.10	0.08

Threats to the trial

Missing outcome data

We do not expect missing outcome data from the main survey as the responses to outcome questions will be mandatory. If the mandatory questions are not completed, that survey will be discarded for the purposes of the survey experiment (though their responses to the survey will be kept) and another respondent recruited.

We will have missing data for the follow-up survey. Although respondents are compensated for their time, we expect an attrition rate between the main survey and follow-up survey of around 40 per cent. However, we do not believe that the form of treatment delivered in the main survey could have any impact on respondents' subsequent decisions about whether to complete the follow up survey. Consequently, we will undertake complete case analysis (ie, drop the records with missing outcomes) and proceed on the assumption that the dropped records are missing independent of potential outcomes (MIPO).

Spillovers

We expect no risk of spillovers for the main survey, and very low risk of spillovers for the follow-up survey

Blinding

Participants will be aware they are taking in part in a study on cyber security but not aware that the survey contains an experimental component.

Method of analysis

The principal analysis of the effect of the intervention will be an adjusted comparison of each our primary outcomes. These estimates, confidence intervals (CI) and p-values will be derived from a linear regression model with the following specification:

$$Y = a + b1T1 + b2T2a + b3T2b + b4X + B5XT1 + b6XT2a + b7XT2b + e$$

Where Y is one of our primary outcomes, T1 is a dummy variable for financial consequences, T2a is a dummy variable for the peer messenger, T2b is a dummy variable for the expert messenger, and X is a vector of mean-centred covariates, which are interacted with each of the treatment dummies.

For binary outcomes, we will conduct a robustness check by running a logistic regression and then calculating average marginal effects.

Exact p-values and confidence intervals will be reported. Our primary analysis will not adjust for multiple comparisons.

Covariates

The table below shows the covariates that will be included in all estimation equations.

Table 5. Covariates

Covariate	**Response format**
Reported frequency of installing software updates on the day they are released	4. Every time 3. Most of the time 2. Sometimes 1. Rarely 0. Never 0. Don't know
Reported frequency of using a **different** password for important accounts	
Reported frequency of using a **strong** password for important accounts	

Pre-analysis plan commitments

No trial data have been collected and 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.