Australian Government
The Treasury

Australian Centre for Evaluation

Pre-analysis Plan: The effectiveness of different online training formats

Contents

Summary table

Project title	The effectiveness of different online training formats
Evaluator (Institution)	Australian Centre for Evaluation (ACE)
Principal investigator(s)	Peter Bowers, Ethan Slaven
Trial design (including number of arms)	Three-arm randomised controlled trial
Unit of randomisation	Individuals
Target group	Subset of APS based on readily available email distribution lists
Anticipated number of participants	600
Primary outcome measure	Training Effectiveness
Secondary outcome measures	Engagement rating, Net Promoter Score, Training Completion, Learning comprehension

Pre-analysis plan commitments

This pre-analysis plan was pre-registered on 14 February 2025. This was after the trial launched in January 2025 but before the trial data was downloaded from Qualtrics or analysed. The data will be downloaded on 17 February 2025. Any deviations from this pre-analysis plan will be documented and justified in the final report.

Policy context

The Australian Public Service (APS) is currently offering numerous online learning opportunities ranging in content and form. These learning opportunities are widely used and an important tool for APS officers' professional development and learning.

Despite the investment in these online learning opportunities across the APS and their importance to professional development, there have been few evaluations of their effectiveness. There has been no formal research or comprehensive assessment to measure the impact of these different online learning methods on engagement, knowledge retention, skill development, and overall professional awareness.

This trial will provide insights into the effectiveness of 3 common online learning methods, a traditional e-learning module, a micro-learning video and a podcast. The results of the trial will guide the development of future training initiatives, ensuring resources to develop training are allocated effectively and the overall quality and engagement with online learning is enhanced.

Trial aim

This trial aims to test an approach to improve the engagement of APS staff with training materials. Specifically, the trial aims to test whether online trainings in the format of micro-learning videos or podcasts can improve training effectiveness compared to the traditional online click-through training modules. Online click-through training modules are currently the most common mode of online training across the APS.

Intervention Type

The research will consist of a 3-arm randomised controlled trial with the following groups:

- Control Group: Click-through training module
- Treatment Group 1: Podcast
- Treatment Group 2: Micro-learning video

Randomisation will occur at the level of the individual, with participants being randomly assigned to one of the 3 arms after they commence the survey.

Each of the training formats will teach participants about the same topic: How randomised controlled trials (RCTs) work, and what they are useful for.

Ethics and consent

Ethics approval for this study has been granted by Macquarie University's Human Ethics Committee (Reference No: 520251868360834).

Participants will be fully informed about what the study is about and what it involves before they participate.

Participants will be invited to take part by email which includes a link to the survey and training material. When participants click the link, they will be sent to the questionnaire that is included in Appendix A. The first page of this is a detailed Participant Information Statement and Privacy Notice that explains the benefits and risks of participating, and what the study is about. The Qualtrics survey platform will ask for consent before proceeding.

Outcome measure(s)

Table 1: Description of primary and secondary outcome variables

Primary outcome

Variable	Training Effectiveness Index – we create an index of training effectiveness that measures both learning and engagement.
Measure	• We create the Training Effectiveness index by first converting the following questions from the survey to a binary variable [0,1]: engagement rating, net promoter score, training completion, and learning comprehension question 1, learning comprehension question 2. • We then sum the binary scores together to create a score out of 5.

Secondary outcomes

Variable	Engagement rating – measures the extent to which participants agree that they found the training material engaging
Measure	• To what extent do you agree with the following statement: "The training content was engaging" (Answer: Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree)
Variable	Net Promoter Score – measures the likelihood of an APS employee recommending the training
Measure	• How likely would you be to recommend this training to a friend/colleague? (Answer: 10-point scale)
Variable	Training Completion – measures how much of the training a participant completed
Measure	• How much of the training module/video/podcast did you complete? (Answer: Little or none, some, most, all of it)
Variable	Learning comprehension – measures participants knowledge of RCTs with 2 learning outcomes
Measure	Learning outcome 1: Consider this scenario: The Australian Government is studying a new approach to language learning through flash cards for primary school children. The government wants evidence to decide whether this is a good approach. Children from the same suburb, with similar NAPLAN scores, and demographics are randomly assigned to 2 groups. Group A receives the flash cards, while group B doesn't. In this scenario, what is the control or comparison group? a. NAPLAN b. Flash cards c. Group A d. Group B
Measure	Learning outcome 2: Which **ONE** of the following correctly completes the sentence below? "Randomising between the treatment and control groups helps find the true causal effect of a program because it ensures... a. Participants in the trial are representative of the overall population of interest. b. People in the treatment group are not systematically different to people in the control group c. The sample size of the study is large enough.
Variable	Preference for alternative format – records whether participants think an alternate training format would be better
Measure	What format do you think is best to deliver this training? (Answer: Short explainer video, online click-through module or podcast format)

Population of interest and sample collection

Our population of interest is all Australian Public Servant (APS) employees. We will email the APSC's MyAcademy email distribution list (totalling approximately 6000 recipients) to sample APS employees willing to undertake the training. Our sample will exclude APS employees who are not currently on the email list as we are relying on a convenient sample through existing email distribution lists.

If there is appropriate time, we will also investigate whether there is the potential to share the survey on other APS email distribution lists that are similar to the APSC newsletter to expand our sample size. However, we will assume we only have the APSC newsletter for our sample size calculations.

Power and sample size calculations

In this section, we discuss the expected sample size and calculate whether we expect to have sufficient power given this sample size.

Sample size

For those who receive the email, we expect completion rates to be around 10% (based on APSC's prior experience). This gives an estimated sample size of 600 participants.

It is of course possible that the response rate could be greater or less than 10%. Included in the table below power is calculated with response rates of 5% and 1% as well.

Power calculations for exploratory analysis

In this section we calculate that the likely sample size of 600 described above would result in sufficient power for our exploratory analysis (200 in control, 200 in treatment 1, 200 in treatment 2).

Given the consequence of a false positive is not severe and since we would like to avoid a large type II error rate, we have opted for an alpha level of 10%. We then look at a 1-sided hypothesis test between a particular treatment group (for example Treatment Group 1) and the Control group. If the sample size is 600, the sample of each group (Control, Treatment 1 and Treatment 2) is approximately 200 people. So, the sample size of a test between the Control Group and just 1 treatment group would be approximately 400 people.

Given the paucity of evidence of our topic, our best guess at the impact is that the treatment has a 'small effect' (Cohen's d = 0.2).

The table below shows the power calculation of our exploratory trials, the power for our primary test will be even greater given the 2 treatment groups are combined. For the exploratory analysis, we calculate that if the survey response rate is at its expected rate of 10%, we would have 76% power to detect a 'small effect'. We view this as sufficient.

Table 2: Power calculations

Sample size	Power assuming 'small effect' (Cohen's d = 0.2)
10% of people respond (n=600)	76%
5% of people respond (n=300)	55%
1% of people respond (n=60)	25%

Power calculations for confirmatory analysis (primary analysis)

For our primary analysis we will be pooling the 2 treatment groups together into a single treatment group, so there will be approximately 200 people in the Control group, and approximately 400 in the treatment group. We have not conducted separate power calculations for these hypothesis tests but note that given the larger sample size, our power would be greater than the 76% expected for the exploratory analysis.

Hypotheses

Main hypothesis: Training Effectiveness will be higher among participants who receive training via micro-learning videos or podcasts compared to those who receive training by the traditional online click-through training module.

Randomisation

Randomisation will occur within the Qualtrics survey software. Survey participants will click through a link to Qualtrics, once on the website all participants will answer a consent question and give their email before randomisation occurs. The randomisation will then occur in the second question block and will be automated by the randomizer feature offered by Qualtrics. This feature will show one of the 3 training options to participants, as a link to those allocated to the traditional e-learning module and as embedded content for the micro-learning video and podcast. The randomizer feature from Qualtrics ensures each training format is randomised as evenly as possible. This means each participant has an equal probability of assignment to each category and across the population there is an approximately equal number of participants in each group.

Method of analysis

The table below outlines how each of the survey response options will be coded for analysis.

Table 3: Dependent variables, survey response options and variable coding

Dependent Variable	Survey response options	Variable coding
Training Effectiveness	Combination of dependent variables to form an index (engagement rating, net promoter score, training completion, and learning comprehension tests)	Continuous variable from 0 to 5. Where each of the secondary outcome variables (engagement rating, net promoter score, training completion, learning outcome 1 and learning outcome 2) are converted to a binary measure [0,1]. Then the binary measures are summed together to create a score out of 5.
Training Completion*	a. Little or none (0%–24%) b. some (25%–74%) c. most (75%–99%) d. all of it (100%)	Binary variable, where {a, b, c} = 0 and d = 1
Net Promotor Score	1–10 scale	Binary variable, where {1–6} = 0 and {7–10} = 1
Engagement Rating	a. Strongly Disagree b. Disagree c. Neither Agree nor Disagree d. Agree e. Strongly Agree	Binary variable, where {a, b, c} = 0 and {d, e} = 1
Learning comprehension	Learning Outcome 1: a. NAPLAN b. Flash cards c. Group A d. Group B Learning Outcome 2: a. Participants in the trial are representative of the overall population of interest. b. People in the treatment group are similar to people in the control group even before they receive the treatment. c. The sample size of the study is large enough.	Continuous variable from 0 to 2. Where learning outcome 1 is converted to a binary variable, where {a, b, c} = 0 and {d} = 1 and learning outcome 2 is converted to a binary variable, where {a, c} = 0 and {b} = 1. Then the binary variables of learning outcome 1 and learning outcome 2 are summed together to create a score out of 2.
Preferred format	a. Short explainer video b. Online click-through module c. Podcast format	Results will be presented descriptively as this question is exploratory

* Note: Trial participants that are randomised but have left at the first page of the survey have their training completion recorded as ‘little or none’.

Dealing with missingness

Some participants may start the survey but not complete it, or not complete all questions. In such cases, we will use the mean value for each incomplete response.

Primary outcome analysis

Model Specification

Our target estimand is the Intent-to-Treat effect. This will be estimated using an ordinary least squares linear regression.

The analysis will use the primary outcome dependent variable of: *Training Effectiveness Index.* The model estimates the effect of being assigned to the treatment groups (micro-learning video or podcast) on *Training Effectiveness*. This estimate, confidence intervals and p-values will be derived from an ordinary least squares regression model using robust (HC2) standard errors and with the following specification:

$$Y_i = \beta_0 + \beta_1 Treatment_i + \in_i$$

Where:

- i is an index for each individual in the trial
- Y is the Training Effectiveness
- β_0 is the intercept
- $Treatment$ is a binary assignment indicator (where {0} = online training module; {1} = podcast format or micro-learning video)
- β_1 is a coefficient representing the average treatment effect
- $\in$ is the heteroskedasticity-consistent type 2 (HC2) standard error term.

Exploratory analysis

Our first exploratory analysis will follow the same model specification as the primary outcome analysis. However, the outcome variable Y will instead be (in separate regressions):

- Training Completion
- Net Promoter Score
- Engagement Rating
- Learning comprehension

Our second exploratory analysis will use the same dependent variable as the primary outcome analysis but follows a different model specification. Instead of pooling the 2 treatment arms we keep them separate and analyse them separately (for example dropping treatment 2 from the sample and comparing treatment 1 to the control group).

Lastly, we will present the results of the ‘Preferred format’ question as exploratory analysis in the form of a chart. This question is not an outcome of the RCT but a question about individuals’ preferences.

Trial threats

Non-compliance

Some participants may not complete the survey after completion of the training materials. This may be due to participant oversight, time constraints or technical difficulty.

Mitigation strategies:

To address this, survey participants will be asked to provide their contact email address, and 1-2 follow-up emails will be sent to remind participants to complete the survey if they haven’t already.

Missing data

Missing data can occur due to:

- Attrition: some participants may not complete the survey at all.
- Technical issues: there could be system errors in the online system or email reminders.

Mitigation strategies:

Cases with missing outcome data will be removed from the analysis. If the case has data for other outcomes, they will be retained for these analyses. If the data appears to be missing at random due to unrelated events like a technical issue, we will proceed with our analysis as pre-specified. Where we suspect missingness is related to treatment assignment we will explore the possible impact of that missingness on our estimates using bounds.

Spillovers

Spillovers are unlikely in this trial, given the outcomes largely relate to the participants experience completing the survey.

Treatment contamination

Treatment contamination may occur if participants are able to access multiple e-learning formats in the trial. This could occur, if, for example, the survey page could be refreshed to get to the e-learning format that the participant wanted to use. Alternatively, it could also occur if participants liked the video, found it online, and shared the link with other participants.

Mitigation strategies:

The Qualtrics survey questionnaire platform will be designed such that participants are only shown one e-learning format for each unique user and cannot regenerate the survey to get different access links.

We do not anticipate sample contamination through participants sharing different training content amongst themselves to be a major issue because the survey questions at the end are clear that they are asking about the training format that the participant has just used (rather than a different one). So even if some participants have already seen the training video, this should not have a major impact on the results.

Interpretation of results

The goal is to assess whether online trainings in the format of micro-learning videos or podcasts can improve training effectiveness compared to regular online click-through training modules.

If no significant effect is observed, this means that we will not have any evidence to suggest alternative e-learning formats to traditional click-through training modules improve training effectiveness. However, this does not necessarily mean micro-learning videos or podcasts do not work to improve training effectiveness–it may reflect other factors, such as that the content itself needs to be reworked, or that changing the format alone is insufficient to change training effectiveness.

Appendices

Appendix A – Qualtrics survey questionnaire

Page	Content/Question	Answer	Mechanics/ Format
1	Participant Information Statement and Privacy Notice (see Appendix B for full statement)	NA	Privacy Statement
2	Please provide your email:		Free text
3 (option 1)	Please complete the learning activity on APS Learn by clicking on **this link,** which will open in a new tab. You will need approximately 25 minutes to complete the course. **Do not close this survey tab.** Please follow these steps: 1. Enter your APS Learn login email (typically your work email ending in ".gov.au") and password in the fields provided and click on "APS Employees Login". 2. If you are not registered, please click on the registration button (the process should take less than 1 minute) 3. Once logged in, click the "Enrol" button on the Randomisation course page and follow the instructions provided to complete the click through learning activity. 4. Return to this tab, click the "next" button and answer the survey questions asked.	NA	Random assignment page
3 (option 2)	Please watch this short video (developed by J-PAL) by clicking on **this link**, which will open in a new tab. **Do not close this survey tab.** The video is less than **3 minutes** in duration. Once you have finished watching the video, return to this tab, click the "next" button and answer the survey questions asked.	NA	Random assignment page
3 (option 3)	Please listen to this podcast by clicking on **this link,** which will open in a new tab. **Do not close this survey tab.** The podcast is less than 15 minutes in duration. Once you have finished listening to the podcast, return to this tab, click the "next" button and answer the survey questions asked.	NA	Random assignment page
4	How much of the training module/video/podcast did you complete?	1. Little or none (0%–24%) 2. some (25%–74%) 3. most (75%–99%) 4. all of it (100%)	Multiple choice
5	How likely would you be to recommend this training to a friend/colleague?	1–10	Scale

Page	Content/Question	Answer	Mechanics/ Format
6	To what extent do you agree with the following statement: “The training content was engaging”	a. Strongly Disagree b. Disagree c. Neither Agree nor Disagree d. Agree e. Strongly Agree	Multiple choice
7	**Consider this scenario:** The Australian Government is studying a new approach to language learning through flash cards for primary school children. The government wants evidence to decide whether this is a good approach. Children from the same suburb, with similar NAPLAN scores, and demographics are randomly assigned to 2 groups. Group A receives the flash cards, while group B doesn’t. In this scenario, what is the control or comparison group?	a. NAPLAN b. Flash cards c. Group A d. Group B	Multiple choice
8	Which **ONE** of the following correctly completes the sentence below? “Randomising between the treatment and control groups helps find the true causal effect of a program because it ensures.	a. Participants in the trial are representative of the overall population of interest. b. People in the treatment group are similar to people in the control group even before they receive the treatment. c. The sample size of the study is large enough.	Multiple choice
9	What format do you think is best to deliver this training?	a. Short explainer video b. Online click-through module c. Podcast format	Multiple choice