Metric Choice

Number of cookies: number of unique cookies to view the course overview page.

Chosen as invariant metric: the number of cookies should remain the same.

Not chosen as evaluation metric: the number of cookies will not change, and thus will not show the effects of the experiment.

Number of user ids: number of users who enroll in the free trial.

Not chosen as invariant metric: in the experiment, after clicking the button, students who say they have fewer than 5 hours to study every week will have be shown two options instead of the checkout process, which means the number of students enrolling this course is likely to reduce. So the number of users will differ.

Not chosen as evaluation metric: though the numbers of cookies of experiment group and control group are roughly the same, the number of user for each cookie may differ a lot, which might result a significant difference between the numbers of users in two groups. So even if there is a change in the number of users, we cannot say for sure that it is due to the change we are investigating.

Number of clicks: number of unique cookies to click the "Start free trial" button (which happens before the free trial screener is trigger).

Chosen as invariant metric: the number of cookies should remain the same.

Not chosen as evaluation metric: the number of cookies will not change, and thus will not show the effects of the experiment.

Click-through-probability: number of unique cookies to click the "Start free trial" button divided by number of unique cookies to view the course overview page.

Chosen as invariant metric: the number of cookies to click the button (number of clicks) and the number of cookies to view the course over (number of cookies) are the same

Not chosen as evaluation metric: click-through-probability will not change, and thus will not show the effects of the experiment.

Gross conversion: number of user ids to complete checkout and enroll in the free trial divided by number of unique cookies to click the "Start free trial" button.

Not chosen as invariant metric: users who claim they have fewer than 5 hours are shown two options. Thus the number of user ids to enroll in the free trial is very likely to change.

Chosen as evaluation metric: to see if the change actually will reduce the number of user ids to enroll in the free trial.

If the number of gross conversion reduces, then I will recommend launching the experiment.

Retention: number of user ids to remain enrolled past the 14day boundary (and thus make at least one payment) divided by number of user ids to complete checkout.

Not chosen as invariant metric: the number of user ids to remain enrolled past 14day boundary and

the number of user ids to complete checkout will differ.
Not chosen as evaluation metric: this metric needs a large number of page views, and thus requires relatively long time.
Net conversion: That is, number of user ids to remain enrolled past the 14day boundary (and thus make at least one payment) divided by the number of unique cookies to click the "Start free trial" button.
Not chosen as invariant metric: the number of user ids to remain enrolled past the 14day boundary is very likely to change.
Chosen as evaluation metric: to see if the change will influence the number of students who finish the free trial and complete the course.
If the number of net conversion does not vary a lot, then I will recommend launching the experiment.

Measuring Standard Deviation

Std of gross conversion: 0.0202
Std of net conversion: 0.0156

For both of them:
Unit of analysis is cookies
Unit of diversion is also cookies
So analytical variability will agree with empirical variability.

Sizing

I won't use Bonferroni correction because Bonferroni correction is mainly used in experiments where multiple metrics are independent, but the gross conversion and net conversion are clearly related.

I'll need 685325 page views.

I'll divert all of the traffic to the experiment and Udacity will need 18 days to run this experiment.
Reasons: the experiment needs 685325 page views, while there are only 40000 page views each day. If I do not divert all of the traffic to the experiment, it will take too much time to run this experiment.
Plus, this experiment is not a risky one for the following reasons:

1. Students who claim they have fewer than 5 hours a week to study can still join the free trial. Thus, the change will not alter the user experience a lot.
2. There is not a significant change on the website. So users do not need much time to adapt to the change.
3. There is no change regarding the database of the website. So no data will be lost.
4. Though users need to enter the information of their credit card, the website clearly has taken some measure to protect the information.

5. It does not have any ethical problem.

Experiment Analysis

Sanity Checks

	Confidence interval	Observed difference	Passes
Number of cookies	[0.4989,0.5012]	0.5006	√
Number of clicks on "Start free trial"	[0.4959,0.5041]	0.5004	√
Click-through-probability on "Start free trial"	[0.0812,0.0830]	0.0821	√

Result Analysis

Effect Size Tests

	Confidence interval	Statistical significance	Practical significance
Gross conversion	[-0.0291,-0.0119]	√	√
Net conversion	[-0.0116,0.0019]		

Sign Tests

	P-value	Statistical significance
Gross conversion	0.0026	√
Net conversion	0.6776	

Summary

The sign tests agree with the effect size hypothesis tests.

Recommendation

Though the net conversion has no statistical or practical significance, the confidence interval indicates it is very likely that the net conversion will reduce (a large area of the confidence interval falls below zero). Also, the decrease might have practical significance (an area of the confidence interval falls behind –dmin). Based on this, I won't recommend launching the experiment.

Follow-up experiment

Change: bold the words indicating people will need more than 5 hours a week to finish the course in the course description.

Hypothesis: this may emphasize the time needed to finish the course, and thus reduce the number of frustrated students who left the free trial because they didn't have enough time.

Metrics I’d like to measure:
Number of cookies: serve as invariant metric
Net conversion: serve as evaluation metric

Unit of diversion: cookies