

# Reducing consumer harm through regulatory experimentation

Protecting vulnerable consumers from risky investing – Financial Conduct Authority



#### Overview

Ineligible consumers were increasingly investing in high-risk products designed for experienced investors. The Financial Conduct Authority (FCA) trialled a range of positive frictions to reduce the number of ineligible consumers accessing these risky products.

## Key finding

Self-certification of ineligible consumers reduced by 36%.

## Outcome

New consumer protection measures proposed, including more robust screening questions and risk warnings on advertisements.

## **Evaluation method**

Randomised control trial.



## The Financial Conduct Authority (FCA) wanted to understand how to protect ineligible consumers from investing in high-risk products.

FCA is responsible for ensuring UK financial markets work well: protecting consumers, promoting competition and enhancing the integrity of the UK's financial system.

During the COVID-19 pandemic, an increasing number of consumers invested in high-risk products without necessarily understanding the associated risks. To invest in these products, consumers must self-certify as either 'high net worth' or 'sophisticated' (experienced with investing).

FCA sought to determine whether positive frictions (intentional obstacles in a process that encourage more thoughtful decision-making) would reduce the number of ineligible consumers self-certifying and consequently accessing high-risk products.

## Intervention and outcome



#### The introduction of positive frictions reduced ineligible self-certification by 36%.

FCA conducted an online randomised control trial to assess consumer self-certification practices. The trial tested a range of positive frictions including commitment checkboxes to confirm that criteria had been met, a requirement for a written evidence declaration, and the imposition of a time delay before consumers could submit their certification statement.

All approaches were successful in significantly reducing self-certification. The checkboxes and evidence declaration in combination were the most effective, cumulatively reducing selfcertification by 36%. These results informed the proposal of a range of consumer protection measures.

These included asking consumers more robust questions about their investment knowledge and experience, and improving risk warnings on advertisements.

## Key steps for successful experiments

<b>Ø-•</b>

#### Communicate early and transparently to secure additional resources.

By communicating the limitations of the online experiment to their policy colleagues early in the research process, the FCA research team was able to bolster the study with additional measures. For example, they added a survey to establish a reliable baseline measure of how many people in the experiment sample would genuinely self-certify.

Consider identifying and communicating potential risks and limitations of your research early on to build support for introducing effective mitigation measures.

#### Consider using an online experiment.

Online experiments are often cheaper and faster to run compared to field experiments. However they typically have less external validity compared to experiments run in the real world.

Consider running on online experiment in situations which require a rapid response to emerging policy challenges, or as a pilot to inform more robust follow-up experimentation.

#### Overcoming lack of reliable baseline

FCA didn't have verified information on whether participants met the certification criteria. This would have reduced the reliability of the findings.

To address this, FCA randomly assigned some participants to a survey to establish a baseline measure of genuine certification. This approach helped them understand the true eligibility of the participants and ensured that the findings were more accurate and relevant for policy insights.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy



## Improving public spaces through regulatory experimentation

#### Reducing cigarette butt littering – NSW Environmental Protection Authority









Location:

Australia



## Overview

Cigarette butts pose a health and safety hazard if not properly disposed of. The NSW Environmental Protection Authority (EPA) partnered with local councils to trial different policy interventions to decrease cigarette butt littering rates.

## Key finding

Butt littering rate decreased from 62% to 42%.

## Outcome 1

Informed state-wide anti-littering guidelines.

## Outcome 2

Led to the development of the Butt Litter Check assessment tool.

## **Evaluation method**

Quasi-experiment.



## The NSW Environment Protection Authority (EPA) wanted to collect evidence on what works to reduce cigarette butt littering.

The EPA is the primary environmental regulator within NSW and is responsible for protecting and enhancing the state's environment. Cigarette butts pose a significant environmental hazard as they leach chemicals and increase fire risks. Available evidence suggested that litter enforcement had been ineffective in deterring cigarette butt littering, with cigarette butts the most commonly littered item in the state (43% of littered items).

The EPA was interested in collecting evidence on alternate litter prevention activities, after conducting a study which found that people who littered were highly influenced by their social and environmental context. To build on this research, the EPA conducted an experiment to investigate the effectiveness of a range of interventions at reducing cigarette butt littering.

## Intervention and outcome



## Changing smoking site infrastructure decreased cigarette butt littering rates from 62% to 42% across all interventions.

In 2018, EPA partnered with 16 local councils to develop a quasi-experiment that assessed the effectiveness of four interventions: creating pathways to bins, instilling a sense of ownership by forming clean demarcated smokers' areas, communicating social norms and emphasising enforcement. These intervention sites were compared to control sites which received no intervention. The project measured littering behaviour and prevalence at baseline, during and three months post-intervention via direct observation and interviews with smokers.

At the end of the experiment, all interventions had reduced butt littering compared to the control sites. The ownership intervention was the most effective. All interventions remained effective after three months, except for the enforcement intervention. This experiment led to the development of NSW anti-littering guidelines, as well as development of an assessment tool called the Butt Litter Check. It also informed the creation of the Cigarette Butt Litter Prevention Program, which aims to reduce the butt littering rate by 50% by 2030.

## Key steps for successful experiments

#### Collaborate with local delivery partners.

EPA collaborated with 16 local councils. This collaboration ensured that the interventions were relevant, addressed local challenges and managed risks.

Consider co-designing a program logic with on-the-ground delivery partners to foster buy-in and overcome challenges to implementation.

#### Build on previous knowledge.

EPA developed evidence-based litter prevention interventions that were based on existing research on the drivers of littering behaviour and the effectiveness of litter prevention activities.

Test and build on existing research to design interventions that effectively address the root causes of specific issues.

#### Addressing partner resource constraints

Councils faced resource constraints. Factors that helped them overcome these included: officer support, assistance with approvals and processes, provision of EPA collateral, and demonstrated effectiveness and cost savings through the EPA methodology.

#### Overcoming risk aversion

Stakeholders expressed concern about creating smoking areas as part of the trial. The EPA engaged extensively with councils and with NSW Health, ensuring that such concerns were heard and addressed.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



# Developing an evidence base through regulatory experimentation

## Understanding consumer passivity in subscription markets – Danish Competition and Consumer Authority







Competition regulation



Denmark



## Overview

As the popularity of subscription-based payment models increased, the Danish Competition and Consumer Authority (DCCA) grew concerned that consumers would become passive and subscribe to services longer than they desired. This could then have negative implications for market competition and innovation.

There was limited evidence available, so the DCCA designed an experiment using subscription payment data and rejected card payments to assess the extent of consumer passivity.

## Key finding

Consumers were 70% more likely to cancel subscriptions when manual renewal was required.

Industry:

## Outcome 1

Informed EU-wide regulatory reform proposal.

## **Evaluation method**

Quasi-experiment.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

#### Outcome 2

Set a new precedent for enforcement cases in other jurisdictions.



## The Danish Competition and Consumer Authority (DCCA) wanted to investigate whether consumers were paying for subscription-based products longer than they desired.

DCCA is responsible for promoting fair and efficient markets in Denmark. Amidst the rapid growth of subscriptionbased payment models, DCCA grew concerned that these payment methods reduced consumer engagement with subscription decisions, leading to passivity. Such passivity could result in consumers retaining subscriptions longer than desired, ultimately reducing market competition, increasing prices and stifling innovation. There was limited available evidence on the harms of subscription-based products.

DCCA wanted to quantify the extent of consumer passivity to help inform new regulation that would better protect consumers and promote market competition.

#### Intervention and outcome



## DCCA found that consumers are 70% more likely to cancel subscriptions requiring manual renewal.

DCCA conducted a quasi-experiment using payment data from 117,000 unique subscriptions. They wanted to quantify the extent of consumer passivity by analysing the impact of rejected card payments on subscription termination. The experiment compared cancellation rates between a control group (who had no payment rejections) and an intervention group (who had at least one rejected payment). The experiment found evidence of consumer passivity.

Consumers were 70% more likely to cancel their subscriptions following a card rejection compared to participants with no card rejection. Over 1,000 days, only 14% of subscriptions in the intervention group remained active, compared to 32% in the control group. DCCA have used the results of this trial to inform the development of an EU-wide regulatory reform proposal. This research has also supported enforcement cases by regulatory agencies in other jurisdictions.

## Key steps for successful experiments



#### Build a convincing rationale to secure internal support.

DCCA highlighted the significance of the problem by drawing lessons from behaviour change in similar situations (by linking the current subscription issues to gym memberships). This helped build internal support from senior leadership and secure resources for the experiment.

## Consider linking your initiatives to existing evidence on well-documented challenges or trends to build a compelling case for experimentation.

#### Foster strategic partnerships.

DCCA collaborated with a third-party subscription payments administrator which enabled access to existing data on subscription behaviours and the identification of a suitable natural experiment.

Consider partnering with external experts or organisations to leverage their specialised knowledge and resources.

#### Overcoming status quo bias

DCCA researchers drew from prior research and behavioural science to develop a compelling business case for investigating subscription markets.

#### Overcoming data and measurement challenges

Subscription behaviours are continuous and therefore difficult to directly measure. DCCA overcame this by identifying and conducting a natural experiment on existing real world data sets.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



# Reducing errors through regulatory experimentation

## Boosting the accuracy of tax returns – Inland Revenue Authority of Singapore







Tax regulation



Singapore



## Overview

The Inland Revenue Authority of Singapore (IRAS) offers a No-Filing Service (NFS) for eligible taxpayers, which pre-fills their tax returns based on available information. Taxpayers are then supposed to check the accuracy of the information before filing. The IRAS was concerned, however, that some NFS taxpayers were not thoroughly reviewing their pre-filled tax returns, leading to downstream issues and revisions.

IRAS tested a two-step solution to reduce inaccuracies by encouraging NFS taxpayers to engage with and carefully review their tax returns.

## Key finding 1

Increased likelihood of taxpayers filing their tax returns by up to 5.9 percentage points (p.p.).

## Key finding 2

Reduced the likelihood of tax returns being reviewed by 0.2 p.p.

## Outcome

IRAS implemented the redesigned SMS notification and self-help navigation prompts for taxpayers.

## **Evaluation method**

Before and after study.



## IRAS wanted to improve tax reporting accuracy by encouraging taxpayers to verify their pre-filled tax returns before filing.

IRAS administers taxes and enterprise disbursements in Singapore. For straightforward tax matters, IRAS offers the No-Filing Service (NFS) to make filing easier for taxpayers. Under this service, tax returns for eligible taxpayers are pre-filled based on existing data sources. Taxpayers are to review the accuracy of these returns and correct any errors by filing a return.

IRAS was concerned that some NFS taxpayers were not thoroughly checking their prefilled returns even when their tax situation had changed, which could lead to potential downstream objections or revisions to their tax assessments. IRAS wanted to promote greater ownership over personal tax matters and reduce incidence of tax assessments finalised on incorrect or outdated claims.

## Intervention and outcome



## IRAS increased the likelihood of NFS taxpayers filing their tax returns by up to 5.9 p.p.

IRAS tested a two-step solution for all NFS taxpayers:

- 1. In step one, IRAS redesigned their SMS notification by making the call-to-action more prominent, drawing on the behavioural science concept of salience. IRAS also sent a new reminder SMS to some taxpayers closer to the filing due date, which emphasised how IRAS had helped to pre-fill parts of the tax return and the consequences of not checking the pre-filled information.
- 2. In step two, IRAS introduced self-help navigation prompts to the tax portal, which encouraged NFS taxpayers to consider the consequences of different options and review their tax returns carefully.

IRAS found that each intervention in the two-step solution helped to encourage NFS taxpayers to verify their tax returns before the filing due date. The redesigned SMS notification increased filing, a proxy for verification, by 2.3 p.p. (compared to the previous year), the reminder SMS by 3.0 p.p. (for those who received it), and the self-help navigation prompts by 5.9 p.p. (for those who accepted it). The solution also reduced the chance of NFS taxpayers being selected for backend review by 0.2 p.p.

IRAS has implemented the redesigned SMS and self-help navigation prompts for all NFS taxpayers, and reminder SMS for selected NFS taxpayers.

## Key steps for successful experiments

#### Establish institutional support to unlock collaboration.

Both IRAS leadership and system teams were supportive of the experiment. This institutional support facilitated effective internal collaboration across the project. For example, the system teams helped to implement the SMS in step one and self-help navigation prompts in step two.

Consider fostering strong institutional support and collaboration within the organisation.

#### Øring together diverse skills.

IRAS formed a multi-disciplinary project team which composed of the team overseeing NFS taxpayers, the central Behavioural Insights (BI) and Design team, and BI and Design community members.

## onsider assembling a diverse team with varied expertise to enhance the effectiveness of experiments and research projects.

#### Overcoming implementation challenges

Running experiments with real-world services can be challenging. IRAS's processes are highly automated with complex system rules. Initially, IRAS wanted to test multiple versions of the SMS notification in step one, but due to the system's complexity, IRAS simplified the experiment and tested only one SMS version. The simpler design allowed IRAS to more clearly identify what worked and what didn't, demonstrating that sometimes simplicity is better than complexity.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

#### Regulatory experimentation case study NSW Productivity and Equality Commission



## Protecting investors through regulatory experimentation

Examining the impact of gamification on online trading -**Ontario Securities Commission** 







Finance



Location:

Canada

## **Overview**

The Ontario Securities Commission (OSC) grew concerned about the rise of mobile-friendly investing platforms and the impact of their engagement tactics on investors' behaviour. The OSC devised an experiment to simulate the real-world so they could better understand how investors respond to common engagement techniques, such as gamification.

## Key finding 1

A gamification technique, consisting of points with negligible value, led to an increase of 39% in trading behaviour.

## Key finding 2

Using stock leader boards increased the likelihood that investors would hold these stocks by 14%.

## Outcome

Provided evidence of potential risks of gamification practices to retail investors.

## Evaluation method

Randomised control trial.

## Background



## The Ontario Securities Commission (OSC) wanted to understand the influence of gamification techniques on investing behaviours.

OSC is an independent Crown agency that regulates Ontario's capital markets by making rules that have the force of law and by adopting policies that influence the behaviour of capital markets participants. A wave of digital, mobile-friendly investing platforms has created new options for retail investors in Canada and around the world. While these platforms have expanded market participation, there is growing interest in some of the digital engagement practices that these platforms use. These tactics, sometimes referred to broadly as gamification, use game-related elements (such as badges, points or leaderboards) to influence investor behaviour. OSC was interested in understanding what, if any, implications gamification techniques had on investor protection.

## Intervention and outcome



## OSC found that the use of gamification techniques significantly influenced trading behaviour, increasing trading by 39%.

OSC conducted a randomised control trial using a simulated online trading platform. They examined the impact of two gamification techniques on trading behaviours. The first technique was a system where points were awarded for trades. These points had negligible real world economic value, but nevertheless led to a 39% increase in trading behaviour. The other technique was the use of leader boards that featured the top traded stocks. OSC found that the use of these leader boards increased the likelihood that investors would hold featured stocks by 14%.

Taken together, these findings demonstrate that gamification techniques can increase trading frequency, which is often associated with poorer investor returns. This experiment therefore provides compelling evidence that gamification techniques may impact investor financial wellbeing. OSC continues to explore this area to understand the impacts of gamification and related techniques on investors.

## Key steps for successful experiments



#### Partner with a research organisation.

OSC partnered with the Behavioural Insights Team (BIT), an organisation specialising in behavioural insights and rigorous online experiments. OSC's internal behavioural insights team collaborated with BIT to develop a carefully designed experiment, execute it swiftly and generate findings with strong generalisability to real-world online trading.

Consider collaborating with specialised external organisations to benefit from their expertise in designing and executing experiments.

#### Cultivate a culture of experimentation.

OSC has a strong culture of experimentation and recognises the value of experimental evidence. This helped secure executive support for the trial early on.

Consider seeking out allies within the organisation who value experimentation, to cultivate an enabling environment for running experiments.

#### Overcoming external validity issues

Online experiments offer significant advantages, including lower costs and faster execution. However, a trade-off is often external validity (how generalisable the results are to the real-world).

By targeting online behaviour (online trading), OSC aimed to reduce the gap between the simulated experimental setting and the real world, thereby improving validity.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



## Better consumer communication through regulatory experimentation

#### Improving energy debt communications – Office of Gas and Electricity Markets

	Your latest energy UM bill is ready Bill date: Bill period: 11 July Date: Bill period: 11 July Solar Sola	How do I pay? Your payment is due by 25th July Pay by this date and you't benefit from our Prompt Payment Discount. EDUID I pay less? Remember - it can be worth thinking about swithing your plan or tariff: about swithing your plan or tariff. Base tariff: Electricity artification of the set of the	
Level:	Industry:	Location:	
Easy	<b>Energy</b>	UK	

#### Overview

Times of financial stress pose significant challenges for households. For households struggling to pay their energy bills, the tone of communications from energy suppliers could impact their willingness to engage and manage their debt effectively.

The Office of Gas and Electricity Markets (Ofgem) conducted an experiment to assess how the tone of debt communications impacted consumers' ability to understand, as well as their willingness to engage with energy suppliers.

## Key finding

Harsh debt communications resulted in poorer comprehension and reduced desire to contact an energy supplier compared to friendlier communications.

## Outcome 1

Informed cross-sector regulatory statements on debt communications.

## **Evaluation method**

Randomised control trial.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

## Outcome 2

Supported work and inspired further research by debt advice charities.

#### Background



#### The Office of Gas and Electricity Markets (Ofgem) wanted to understand how the content and tone of energy debt communications were perceived by energy consumers.

Ofgem is responsible for safeguarding the interests of energy consumers in Great Britain. The COVID-19 pandemic and rising energy prices had left many families with increasing debt and unable to pay their energy bills, some for the first time. Ofgem was concerned that the communications energy companies were sending to families about their debt were overly harsh and ineffective. Ofgem wanted to investigate how the content and tone of initial debt communications from energy suppliers were affecting consumers' willingness to engage with them and effectively manage their debt.

## Intervention and outcome



#### Debt communications that were harsh in tone resulted in worse comprehension and poorer attitudes compared to friendly-toned communications.

Ofgem conducted a two-phased research project. In the first phase, one-on-one interviews with 30 domestic energy consumers (including vulnerable individuals) were conducted over three waves to refine debt communications. The second phase was designed to quantify customers' intended behaviours and attitudes in response to different versions of debt communications, simulating a real-life scenario.

To achieve this, an online behavioural experiment was designed for a demographically representative sample of over 1,500 Great Britain domestic energy customers. Participants were randomly assigned to read one of four debt communications before answering questions about their intentions to act, emotional reactions and comprehension of key information.

Ofgem found that communications with a harsh tone (as described by participants) reduced the likelihood of a participant wanting to contact a hypothetical energy company compared to those with a friendly tone. Harsh communications also resulted in worse comprehension of key information and were less likely to be read in full. This project inspired debt research charities to conduct further research. Moreover, it provided a strong evidence base for regulators to make shared statements on debt communications in the UK Regulators Network Cost of Living Working Group letter.

## Key steps for successful experiments

<b>Ø-•</b>

#### Build internal research capability.

Ofgem have an internal research unit which enabled them to design, program and conduct the research project entirely in-house. This approach was not only more cost effective but also significantly expedited the project and ensured high-quality, tailored research.

Consider building internal research capability to streamline and enhance experimentation.

#### Combine different research methods.

Combining different research methods involves using both qualitative (exploratory, non-numerical data) and quantitative (structured, numeric data) approaches. Ofgem leveraged insights from their qualitative research (the interviews) to inform and optimise the design of their quantitative research (the online experiment).

Consider using both a qualitative and quantitative research approach to enrich your project.

#### Managing risks

Ofgem was initially concerned about the risks associated with handling personal data directly (in-house) for data analysis and storage. The project team addressed these concerns by conducting a Data Protection Assessment and committing to deleting data within three months.

A Data Protection Impact Assessment is a process that evaluates how personal data is collected, stored and used to ensure compliance with data protection laws and regulations.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



# Supporting online safety through regulatory experimentation

Protecting users from harmful online content – Ofcom



## Overview

Content controls are an important tool offered by online platforms that allow users to limit the types of content they are shown. Ofcom partnered with the Behavioural Insights Team (BIT) to conduct two trials to better understand how information and options presented to users influences their use of content controls.

## Key finding 1

Showing information with examples of sensitive content on the decision page encouraged more users to choose stricter content controls.

## Key finding 2

Prompts encouraged users to check their content settings.

## Outcome

Findings contribute to evidence on online safety and media literacy.

## **Evaluation method**

Randomised control trial.



## Of com wanted to build evidence on how platform choice architecture affects engagement with social media content controls.

Ofcom is the regulator for online safety in the UK, under the *Online Safety Act 2023*. Also, Ofcom has a duty to promote media literacy and regulates UK established video-sharing platforms, with respect to rules protecting users from harmful content. Enabling people to take control of their online experiences is important for media literacy and online safety. Content controls, which allow users to reduce the amount of sensitive content they see, are one important tool offered by social media platforms. Behavioural research demonstrates that the way platforms design and present their services and tools ('the choice architecture') shapes how users respond. Ofcom wanted to explore how <u>choice architecture affects engagement</u> with content controls among adults.

## Intervention and outcome



## User decisions on content controls were heavily susceptible to the way that choice was presented.

Of com partnered with the Behavioural Insights Team (BIT) to conduct two randomised control trials using a mock social media platform. The two trials targeted different stages in the user journey.

The first trial explored the choice of content controls ('all content' or 'reduced sensitive content') when setting up a new social media account. Users were randomly assigned to one of five groups to test the impact of a) preselecting a default option, and b) varying how information and choice options are presented. Presenting information with examples of sensitive content on the decision page increased the selection of 'reduced sensitive content' to 29% from 24% in the control condition. When 'all content' was pre-selected, only 15% opted for 'reduced sensitive content'.

The second trial tested whether prompts encourage users to check content settings while browsing, and the impact of prompt timing and message. Prompts increased the proportion of users checking their settings to 17%–23% (4% checked their settings without a prompt).

The prompt which led to the greatest increase in users checking their settings emphasised the ease of changing settings and was sent after a user disliked a sensitive post.

## Key steps for successful experiments

<b>Ö-I</b>

#### Establish a process to identify barriers and develop interventions.

Ofcom followed a structured process to identify priority areas and develop interventions. This included conducting initial desk research, running internal workshops and prioritisation exercises as well as leveraging existing behavioural frameworks and taxonomies.

Consider developing a structured process grounded in established frameworks and practices.

#### lmprove internal research capability.

Ofcom's dedicated Behavioural Insight Hub played a crucial role in developing this experiment. Consider establishing a similar research unit to enhance internal research quality and experimental capabilities.

#### Creating evidence-based guidance

The results of these trials gave Ofcom a set of insights into how users make decisions about their content settings and the way those choices are shaped by platform design. This work builds on the evidence base for Ofcom's online safety work related to content controls. It also contributes to Ofcom's research on on-platform interventions to promote media literacy and informed the <u>Best Practice Design Principles for Media Literacy</u>.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



# Exploring new technology through regulatory experimentation

## Testing the applications of generative artificial intelligence (AI) – Australian Securities and Investments Commission



#### Overview

The Australian Securities and Investments Commission (ASIC) sought to explore the capabilities and limitations of generative AI and its usefulness for their internal processes.

ASIC trained a generative AI large language model (LLM) and conducted an experiment to assess the quality of summaries produced by the LLM against those prepared by humans.

## Key finding

Al-generated summaries did not capture the complexity and nuance as well as human-generated summaries.

## Outcome

Provided valuable insights on the current limitations of a generative AI large language model (LLM).

## **Evaluation method**

Quantitative assessment and qualitative debrief.

## Background



#### The Australian Securities and Investments Commission (ASIC) wanted to explore the effectiveness of AI in summarising public submissions made to a parliamentary inquiry.

ASIC is Australia's independent integrated corporate, markets, financial services and consumer credit regulator. ASIC wanted to conduct an experiment on generative AI (a type of artificial intelligence that creates new content based on prompts provided by the user and informed by large datasets) with a focus on measuring the quality of the generated output rather than the performance of the models. The motivation stemmed from the need to understand the current capabilities and limitations of generative AI in an environment where the technology is rapidly evolving. ASIC also wanted to explore how it could potentially support or enhance internal processes. The trial was exploratory and not for regulatory use.

## Intervention and outcome



## AI-generated summaries were less effective than human-generated summaries, particularly in capturing nuance and context.

ASIC conducted a five week proof-of-concept trial using a specific generative AI large language model (LLM) to summarise a sample of public submissions made to a parliamentary inquiry. The trial aimed to optimise the prompts used to instruct the LLM and then compare the AI-generated summaries to human-generated ones. The final phase of testing involved a blind assessment of AI and human summaries based on a standardised criteria, focusing on the quality and accuracy of the generated outputs, followed by a qualitative debrief with the assessors.

ASIC found that the AI summaries performed lower on all criteria compared to human summaries, particularly in capturing nuance and context. These point-in-time results related to the use of certain prompts, using a specific LLM, for a specific use case and should not be extrapolated more widely. The trial provided valuable insights which can be applied to future AI experimentation to ensure ASIC has a continued understanding of the opportunities and uses of AI as the technology evolves, including its shortcomings.

## Key steps for successful experiments

<b>Ö-I</b>

#### Experiment with novel and emerging technologies.

As a novel and emerging technology, generative AI offers fertile ground for impactful research. By emphasising the benefits of building evidence for AI's usefulness, the ASIC research team successfully gained senior buy-in and support to develop a safe and secure environment for rapid experimentation.

Consider how you might trial new technology to improve your regulatory processes and practices.

#### Using internal and external subject matter experts.

Partnering with a third-party AI expert organisation provided technical AI expertise. This enabled ASIC to blend external AI expertise with internal research and regulatory subject matter expertise, resulting in a more robust exploratory trial.

Consider seeking partnerships with external experts to bolster internal capabilities and gain access to specialised knowledge.

#### Using internal staff as participants

ASIC was interested in assessing the quality of the AI and human submission summaries, a task requiring specific internal expertise. ASIC was conscious that knowledge of the source of the summaries and the order summaries were assessed in could have introduced bias. To address this, ASIC randomised the order ASIC assessors viewed the summaries in and employed a blinded experiment which meant the assessors did not know where the summaries came from.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



## Improving consumer information through regulatory experimentation

## Supporting consumer decisions around add-on insurance – Australian Securities and Investments Commission

Level:





Finance

Industry:



Location:

Australia

#### Overview

The Australian Securities and Investments Commission (ASIC) grew concerned about add-on insurance (insurance products sold alongside other products, such as extended warranties on electronics) due to questionable sales tactics and poor value for money.

ASIC tested the effectiveness of six different information statements to determine how best to support consumers when making purchasing decisions for add-on insurance.

## Key finding

Providing an information statement led to a 24% reduction in hypothetical purchases of add-on insurance.

## Outcome

Supported the effectiveness of using an information statement as part of a deferred sales model and helped inform ASIC's regulatory guidance.

## **Evaluation method**

Randomised control trial.



## The Australian Securities and Investments Commission (ASIC) was concerned about poor value add-on insurance products.

ASIC is Australia's independent integrated corporate, markets, financial services and consumer credit regulator. ASIC was concerned about add-on insurance. Examples include insurance policies sold with car loans or extended warranties on electronics. ASIC surveillances and consumer research had found several issues including poor value for money, high-pressure sales techniques and selling to consumers who could not actually make claims.

The 2017 Financial Services Royal Commission underscored these concerns, with the recommendation for the government to implement a deferred sales model (DSM). The DSM requires a waiting period between the purchase of a primary product and the offer of add-on insurance. ASIC wanted to investigate the effectiveness of information statements provided as part of the DSM.

## Intervention and outcome



## A 24% reduction in add-on insurance purchases when given an information statement compared to not receiving one.

ASIC partnered with the Behavioural Economics Team of the Australian Government (BETA) to conduct an online experiment testing the effectiveness of six different information statement designs. ASIC and BETA tested a) whether providing an information sheet influenced consumers decisions, including opting out of future solicitations for add-on insurance, and b) how different design and information features (for example, claims ratio) affected decisions.

The experiment involved three hypothetical shopping scenarios: buying a phone, buying flights or getting a loan. Participants first decided which product they would like to buy and were then randomly assigned to see one of six information statements or proceeded with no information (control condition).

A 24% reduction in hypothetical add-on insurance purchases was found for those given an information statement compared to those who were not. However, there was no evidence to suggest some of the specific features such as a claims ratio affected participants decisions. The findings contributed to ASIC's regulatory guidance, and demonstrated how experimentation can support consumer protection interventions.

## Key steps for successful experiments

<b>Ö-</b>

#### Capitalise on high-level directives.

Because their research was based on a recommendation from the Financial Services Royal Commission and subsequent law reform, ASIC were able to prioritise the project and more easily secure resources and support.

When a project is backed by a high-level directive, consider using it to gain momentum and secure necessary resources early on.

#### Collaborate with a research organisation.

By collaborating with BETA, ASIC gained valuable support and expertise, enhancing the study's design and implementation.

Consider collaborating with research organisations to bolster research capabilities and leverage external expertise for more effective outcomes.

#### Limits of disclosure

Even though this experiment focused on testing information provided to consumers, when it comes to the complex landscape of financial services there can be limitations to the effectiveness of disclosure (see <u>ASIC's previous work</u>).

In this example, the information statement tested was part of a broader intervention designed to break up the sale of add-on insurance, which also included a mandated time lag between product purchase and add-on insurance sale. This meant the information statement was not relied on in isolation.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy Want help designing your own experiment? Go to productivity.nsw.gov.au/contact-us

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



## Understanding market preferences through regulatory experimentation

Understanding how consumers would value a central bank digital currency– The Reserve Bank of Australia





## Overview

The Reserve Bank of Australia wanted to know how consumers would value having a digital currency issued by the central bank. Since this does not exist yet, it is difficult for people to accurately place a value on this option. The RBA overcame this issue by conducting a discrete choice experiment, using survey questions to simulate how people would respond in real-world situations.

## Key finding

Consumers on average would not value the added safety of a central bank digital currency, but do value the privacy characteristics it could potentially offer.

## Outcome

Informed wider discussions on the costs and benefits of introducing a central bank digital currency.

## **Evaluation method**

Discrete choice experiment.

#### Background



## The Reserve Bank of Australia (RBA) wanted to assess how much Australians would value a central bank digital currency compared to existing and emerging forms of money.

RBA is Australia's central bank and is responsible for the stability of the financial system. As digital transactions are becoming more commonplace, RBA wanted to explore the costs and benefits of introducing a new form of money, known as a central bank digital currency. This new money would be issued by RBA, instead of commercial banks, and as such, it would offer more safety and potentially more privacy for consumers when making digital transactions. RBA was uncertain to what extent consumers value the added privacy and safety, and therefore whether they would likely adopt such a digital currency.

#### Intervention and outcome



## RBA found that consumers on average do not value the additional safety of a central bank digital currency, but do value some of the privacy characteristics it could have.

RBA conducted a discrete choice experiment. They chose this design because they couldn't directly observe the behaviour of interest (since a central bank digital currency doesn't yet exist). Additionally, standard survey questions often fail to reflect real-world behaviours.

The experiment asked respondents to choose between two hypothetical bank accounts. Each account had several attributes that varied randomly, including the fees involved, the level and nature of privacy offered, and whether the money was safeguarded by RBA or a commercial bank. RBA wanted to find out the extent to which respondents were willing to choose the higher fee option when a) the money is safeguarded by RBA and therefore safer, and b) when the privacy levels were higher.

Results showed that consumers on average were unwilling to pay extra for the added safety of using a central bank digital currency, but would pay \$5 more per person per year for the privacy characteristics that the central bank digital currency could have. These results helped inform ongoing discussions in RBA on the costs and benefits of introducing this new form of money in Australia.

## Key steps for successful experiments

<b>Ö-I</b>

#### Partner with external experts.

Discrete choice experiments are challenging to design well. As such, RBA engaged an external research expert that specialises in this type of experiment.

## Consider similar partnerships to enhance the rigour and reliability of your experimental designs and statistical analyses.

#### Use existing infrastructure and resources.

RBA was able to minimise the cost of running a large, nationwide survey by adding the discrete-choice question onto a pre-existing RBA survey.

Consider using existing infrastructure and resources to reduce costs and streamline data collection.

#### Overcoming resource constraints

The project involved a collaboration with an academic expert, which helped RBA to overcome resource constraints.

#### Testing in a hypothetical market

A central bank digital currency doesn't exist in Australia yet. The RBA was concerned that self-reported attitudes towards an unfamiliar financial product would have limited behavioural validity. Running a discrete choice experiment enabled RBA to simulate real world behavioural responses to a hypothetical product offering.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



## Continuous improvement through regulatory experimentation

## Increasing survey response rates to better assess education providers – NSW Department of Education







Education



Location:

Australia



## Overview

The NSW Department of Education (DoE) wanted to improve the volume of student survey responses so that it could better assess education providers. Through an iterative series of A/B tests, DoE experimented with the timing, frequency and medium of their survey to optimise response rates.

## Key finding

Increased survey response rates by approximately 100%.

## Outcome 1

Increased the number of providers with sufficient survey sample from approximately 100 to 250.

## **Evaluation method**

A/B testing.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

## Outcome 2

Online format significantly reduced yearly costs for survey administration.



## The NSW Department of Education (DoE) wanted to improve the assessment of education providers by increasing student survey response rates.

DoE oversees and manages the state's educational policies and programs, including the allocation of funding to various training providers. These funding decisions are based on provider performance, which is assessed through the annual nationwide Student Outcomes Survey. Recently, DoE transitioned from traditional phone and paperbased surveys to an online format. This shift has enabled the department to access real-time data on student outcomes, improve response rates and achieve significant cost savings. As part of this transition, DoE's research team aimed to optimise and test multiple methods of administering survey reminders to further enhance response rates. While not a direct test of a regulatory approach, increased survey engagement is a universal benefit that can help DoE to more accurately evaluate provider performance, particularly for smaller providers.

## Intervention and outcome



## DoE doubled survey response rates by changing the timing, frequency and medium of reminders.

DoE conducted a series of iterative A/B tests to determine the optimal timing, frequency and medium (email and text) for survey reminders to maximise response rates. They randomly allocated students to either a) the control group, receiving the standard reminder, or b) the treatment group, which varied across the different tests. For example, one experiment tested different lengths of time between the first and second reminders sent to the students.

Each experiment led to incremental improvements in response rates. By the end, the team had doubled the survey response rates through minor changes in the way in which the reminders were sent. As a result, DoE now has more accurate and reliable performance data of smaller providers. Moreover, this work inspired other teams within the department to conduct similar A/B tests of their own.

## Key steps for successful experiments

<b>\$-•</b>

#### Conduct experiments at opportune moments.

During the transition to the online survey format and before full automation was implemented, DoE's research team had to manually administer the survey. This provided a window of opportunity to conduct the A/B tests, where the research team had access to and could easily manipulate survey delivery variables.

## Consider timing experiments during moments that present unique opportunities for intervention and experimentation.

#### Iterate and improve the experiment design.

The DoE research team was able to achieve a significant increase in survey response rates by adopting an iterative approach to their experimental design, where they tested and implemented a variety of minor changes to the survey reminders.

Consider adopting a similar iterative strategy to optimise experiment results.

#### Overcoming resistance to change

DoE had been using phone and paper surveys for decades. Transitioning to and experimenting with a new process can be challenging in a large organisation. DoE's culture of innovation and strong support from senior managers enabled the research team to build support for the experiments and access the required resources.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.



## Process improvement through regulatory experimentation

#### Improving early childhood education and care quality assessments – The NSW Department of Education, Early Childhood Education and Care Regulatory Authority



Level: Moderate



Education



Location:

Australia



#### Overview

The NSW Early Childhood Education Care Regulatory Authority (NSW ECEC Regulatory Authority) oversees the provision and regulation of early childhood education care services in NSW. The NSW ECEC Regulatory Authority was shifting from a process of comprehensive quality assessment towards more narrow and frequent partial reassessments.

The NSW ECEC Regulatory Authority used an early adopter model to test the impact of the partial reassessment process on quality ratings and to refine their risk-based regulation.

## Key finding

Partial reassessments saved an average 12 hours per assessment.

## Outcome 1

Optimised the implementation of partial reassessments before wider rollout.

## **Evaluation method**

Quasi-experiment.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

## Outcome 2

Collaboration with other regulators supported implementation.



## The NSW Early Childhood Education and Care Regulatory Authority (NSW ECEC Regulatory Authority) wanted to evaluate the impact of changes to quality assessments of early childhood education and care services.

The regulatory authority oversees the provision and regulation of early childhood education care services across the state, rating services through a national quality rating system. Services are given an overall rating based on their performance against seven quality areas. In late 2022, the Australian Children's Education and Care Quality Authority announced changes to the assessment process at a national level, with each state responsible for implementation in their own jurisdiction.

One of the changes was the introduction of a greater use of partial reassessments, where a small subset of quality areas are assessed, rather than all seven. The rationale behind this change is that partial reassessments would be less labour-intensive and allow regulators to conduct assessments more frequently. This would enable the provision of more current information to families. NSW ECEC Regulatory Authority wanted to test whether partial reassessments would impact the accuracy of quality ratings compared to full assessments and better support risk-based regulation.

## Intervention and outcome



## NSW ECEC Regulatory Authority found that, on average, partial reassessments save 12 hours per assessment.

The regulatory authority first designed the implementation of partial reassessments with input from the sector. This informed a decision to allow services to choose one quality area in which they would be reassessed. The remaining quality areas were chosen by authorised officers based on risk and sector trends. An early adopter model was implemented, in which services could opt-in to experience partial reassessments before statewide roll-out. This enabled the regulatory authority to collect a wide range of qualitative and quantitative data (such as surveys, timestamps, internal reports) to evaluate the initiative. They found partial reassessments saved an average of 12 hours per assessment. Preliminary findings suggest partial reassessments may be as accurate in picking up changes to quality, when compared to full assessments for lower risk services.

## Key steps for successful experiments

<b>Ö-I</b>

#### Sandbox and prototype small processes within a larger initiative.

NSW ECEC Regulatory Authority developed a tool to assist officers in selecting which quality areas to reassess for a specific service. This tool supported a risk-based approach, ensuring partial reassessments accurately capture quality. To minimise the risk of introducing bias into decision-making, the regulatory authority created a sandbox and prototyped various versions of the tool.

Consider adopting a similar iterative approach to optimise processes before full-scale implementation.

#### Collaborate with other jurisdictions and regulators.

The changes to quality assessments were mandated across Australia. In the early stages, NSW ECEC Regulatory Authority engaged with other states to share ideas about how the changes could be implemented. Once implementation was underway in NSW, they were able to share their insights with other states to assist their transition.

Consider looking to other jurisdictions or agencies for similar guidance and collaboration.

#### Overcoming resistance to change

NSW ECEC Regulatory Authority anticipated resistance from the sector to changes in the assessment process. To address this, they used an early adopter model, gathering evidence and insights before implementing the new processes sector-wide. The regulatory authority also proactively communicated the changes to stakeholders through various events, such as roadshows and webinars, well in advance of wider implementation.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission. Image credit: Adobe Stock.

## Glossary of experimental designs

## Level: Easy

#### A/B testing

A simplified form of a randomised control trial (RCT) where participants are randomly shown one of two versions (A or B) of a webpage, app or other product.

These tests often focus on optimising a single outcome. They can also be iterative: once one version wins, it may be compared with a new version C. This type of study is typically easy to implement, especially with digital tools and platforms that automate the process.

#### Before-and-after study

A study that measures and compares outcomes before and after a treatment or intervention.

These studies are relatively simple to design and analyse. However, a key limitation is that results can be influenced by external factors (for example, changes in the weather across the study period might affect outcomes) and this can't be distinguished from the effect of an intervention.

#### Level: Moderate

#### Discrete choice experiment (DCE)

A research method used to elicit preferences by asking participants to choose between sets of alternative options that have varying features.

This approach helps in understanding decision-making and the value placed on different features. While DCEs are not as resource-intensive as RCTs, they require careful design and sophisticated analysis to accurately interpret the trade-offs participants are willing to make.

#### Online experiment

A study conducted over the internet where participants complete tasks or respond to questions remotely. This can allow researchers to reach a larger and more diverse group of participants very rapidly, compared to traditional lab settings.

A common limitation of online experiments is reduced external validity – the extent to which experimental findings can be generalised to real-world settings. This is because online experiments often don't resemble the real world context of the behaviour being studied. However, when the behaviour of interest is performed online, this gap is reduced. External validity is less of an issue in these cases, especially when using a simulated online environment that closely mimics real-world settings.

#### Quasi-experiment

A research design that resembles an experiment but does not involve randomly assigning participants to different groups.

Due to the lack of randomisation, quasi-experiments are susceptible to various biases. One common bias is self-selection bias, where participants who choose to take part might be more motivated and therefore perform differently, either better or worse, based on their reasons for participating. These biases can complicate the establishment of cause-and-effect relationships.

#### Level: Advanced

#### Randomised control trial (RCT)

A study in which participants are randomly assigned to either a treatment group or a control group.

This randomisation allows researchers to measure the causal effect of the treatment by comparing outcomes between the groups. RCTs are considered the gold standard for determining causal relationships.

Want more regulatory experimentation resources? Go to productivity.nsw.gov.au/regulatory-policy

Want help designing your own experiment? Go to productivity.nsw.gov.au/contact-us

This material was produced by the Behavioural Insights Team on behalf of the NSW Productivity and Equality Commission.









