Road safety messaging trial with food delivery workers in the gig economy

Centre for WHS

THE BEHAVIOURAL INSIGHTS TEAM

MACQUARIE University SYDNEY · AUSTRALIA

This report and the work it describes were funded through the Workers Compensation Operational Fund. Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and does not necessarily reflect SafeWork NSW policy.

Prepared by:
Dr Elizabeth Convery[1]
Dr Bowen Fung[1]
Erin Howard[1]
Saul Wodak[1]
Zoe Powell[1]
Dr Ben Searle[2]
Angelica Vårhammar[3]

June 2021

[1]The Behavioural Insights Team, Sydney 2000

[2]Macquarie University, Macquarie Park 2109

[3]Centre for Work Health and Safety, New South Wales Government, Sydney 2000

Executive summary

Background and method

Food delivery workers (FDWs) in the gig economy are at risk of illness or injury while working. While the number of incidents involving FDWs that have been reported to the health and safety regulator in NSW is low, there has been an exponential increase over the past three years, from one incident reported in 2017 to 19 in the first half of 2020 alone. Tragically, a further five FDWs died in traffic accidents on Australian roads between September and November 2020, underscoring the pressing need to address this problem.

FDWs are commonly defined as independent contractors and are thus primarily responsible for their own work health and safety (WHS). However, food delivery platforms (FDPs) have significant influence over the work of their FDWs. FDPs and FDWs thus share a duty of care toward each other and anyone involved in or impacted by their work. However, lack of direct oversight of FDWs, in combination with workforce demographic factors and limited experience, means that FDWs may lack the capacity to fully and effectively manage WHS risks on their own.

The Behavioural Insights Team, in collaboration with Macquarie University and the Centre for WHS, undertook a four-phase project that aimed to improve the WHS of FDWs in the gig economy. Phase 1 aimed to describe the characteristics of FDWs and their WHS behaviours, knowledge, and concerns, while Phase 2 focused on exploring the WHS behaviours, knowledge and concerns of FDPs. The primary WHS concern nominated by FDWs and FDPs was traffic accidents. FDWs also identified verbal abuse and robbery or vandalism of delivery equipment as key concerns. Risk exacerbation behaviours among FDWs included using mobile phones while riding or driving, working while fatigued, wearing dark clothing at night, cycling on footpaths and in other pedestrian-only areas, failing to follow COVID-safe guidelines on social distancing and mask-wearing, and speeding or rushing. These behaviours are extremely widespread - only 5% of the FDWs we surveyed reported "never" rushing to deliver orders - and are driven largely by the desire to achieve and maintain high in-app ratings, satisfy customers, and maximise income. Detailed results of Phases 1 and 2 can be found on the Centre for WHS' Knowledge Hub.

This report details the results of Phases 3 and 4. In Phase 3, we worked closely and extensively with FDWs, FDPs, and an e-bike rental company to co-design a proactive risk reduction intervention to improve FDWs' safety on the road. In Phase 4, we evaluated the intervention in a randomised controlled trial (RCT) with a major FDP, who rolled out the intervention to their fleet in Greater Sydney and Greater Melbourne for a one-month trial period.

Results and discussion

An iterative co-design process with five FDWs, representatives of four major FDPs, and the owner of an e-bike rental shop was conducted to create the intervention. The co-design process involved 1:1 interviews, a workshop, and co-creation and iteration of the intervention and trial design. A key issue that emerged during co-design was that many FDPs intend delivery times to be an *estimation*, while FDWs perceive them as an *expectation*, beyond which they could incur penalties for failing to consistently meet the delivery times shown in the app. As a result, we co-designed an intervention with our stakeholders that aimed to reframe FDWs' perception of delivery times as estimations rather than expectations. The intervention was a suite of four messages, each underpinned by a different behavioural science principle: beliefs about consequences, anchoring, descriptive social norms, and implementation intentions. The messages were then evaluated in a randomised controlled trial (RCT) that was conducted in partnership with a major FDP in Greater Melbourne and Greater Sydney. In the trial, the FDP sent one message per week via the FDP's in-app messaging function to 7,066 of their active FDWs, with the remaining 7,215 FDWs acting as the control group.

FDWs who opened at least one message showed 3% slower average travel speeds relative to the control group, which was a statistically significant difference. This finding suggests that the intervention was modestly effective. For the fleet of 14,281 FDWs as a whole, however, the messages had no significant effect on average travel speed, the subjective belief that the FDP penalises FDWs for late deliveries, or FDWs' perception of safety on the road while delivering for the FDP. This can be at least partially explained by low engagement with the messages and the post-trial survey, which is consistent with past challenges that FDPs and researchers have faced when attempting to engage FDWs as a singular cohort. The findings of this trial suggests that the greater challenge for implementation and scaling of messaging interventions is first determining how best to ensure these interventions are received, read, and acted upon by FDWs.

Table of Contents

List of Figures

List of Tables

Introduction

Background and objectives

Technological advancement and the emergence of app-based business structures have created flexible working opportunities for workers across a range of expanding sectors (Geissinger et al., 2020). These changes have also provided better access, flexibility, and prices for customers and consumers. However, despite such advancements, the gig economy has challenged and generated confusion around work rights, legal roles, and responsibilities (e.g., De Stefano, 2015; Johnstone, 2019).

To date, research into work health and safety (WHS) concerns in the gig economy has been limited (Christie & Ward, 2019) and often related to industrial relations and associated working condition challenges (e.g., Apouey et al., 2020). Similarly, the bulk of research into the WHS of those engaged in other types of insecure work has, to date, focused on people who are self-employed, work casually, or are on temporary contracts (e.g. Koranyi et al., 2018). While the trends identified among these populations may be broadly applicable to workers in the gig economy, it is unknown to what extent this is true and whether the same risks and controls apply and can be implemented in its digital platform-based business structures.

In order to support the WHS of gig economy workers, this project focuses on the workers and digital platforms of one of the most visible sectors of the gig economy: food delivery. Food delivery workers (FDWs) in the gig economy are at risk of illness or injury while working. While the number of incidents involving FDWs that have been reported to the health and safety regulator in NSW is low, there has been an exponential increase over the past three years, from one incident reported in 2017 to 19 in the first half of 2020 alone (SafeWork NSW, 2020). The majority of these incidents are road- and traffic-related, but previous research has highlighted additional harms that FDWs risk on the job, including physical assault, intimidation, and verbal abuse (e.g. Bright & Fitzgerald, 2019; Keoghan, 2020). The global COVID-19 pandemic has posed an additional and complex challenge to WHS. FDWs cannot work from home and, in some cases, are putting their own health at risk by delivering to those in self-isolation (Amin, 2020; Ortiz-Prado et al., 2021). As more inexperienced workers are becoming FDWs in response to a widespread rise in unemployment, the risk of WHS harms may increase.

FDWs are commonly defined as independent contractors and are thus primarily responsible for their own WHS. However, food delivery platforms (FDPs) have significant influence over the work of their FDWs. FDPs and FDWs thus share a duty of care toward each other and anyone involved in or impacted by their work. While sole traders in other industries, such as construction, operate under similar contracting arrangements and must also manage their own WHS (e.g. procuring their own safety equipment and training), construction contractors' work environment exerts a greater degree of control (e.g. rules about personal protective equipment (PPE) upon entering the site) and direct supervision (e.g. from other on-site staff, contractors and managers). These environmental characteristics, in combination with workforce demographic factors and limited experience, means that FDWs may lack the capacity to fully and effectively manage WHS risks on their own (Convery et al., 2020a).

The Behavioural Insights Team, Macquarie University, and the Centre for Work Health and Safety have conducted a four-phase research project that seeks to improve the WHS of FDWs and those impacted by their work. Phase 1 and 2 were the first to describe the WHS understanding, behaviours, and concerns of Australian FDWs (Convery et al., 2020a) and FDPs (Convery et al., 2020b). The main WHS concerns identified among FDWs were traffic accidents, verbal abuse by customers and members of the public, and robbery or vandalism of delivery equipment. Risk exacerbation behaviours included using mobile phones while riding or driving, working while fatigued, wearing dark clothing at night, cycling on footpaths and in other pedestrian-only areas, failing to follow COVID-safe guidelines on social distancing and mask-wearing, and speeding or rushing. Speeding and rushing behaviours were particularly widespread: only 5% of the FDWs surveyed in Phase 1 reported "never" rushing to deliver orders. This appears to be driven largely by the desire to achieve and maintain high in-app ratings, satisfy customers, and maximise income (Convery et al., 2020a). Of particular note is the observation that many FDWs viewed failure to meet "prescribed" delivery times as negatively affecting their ability to access future orders and preferred shifts or even their tenure with the FDP. In Phases 3 and 4, the results of which are detailed in this report, we engaged with end-users, industry partners, and other stakeholders to co-design a behavioural intervention that aims to reframe FDWs' perception of delivery times as estimates rather than expectations and evaluate its effectiveness in reducing risk exacerbation and/or increasing risk mitigation behaviours on the road.

Literature review

The key findings of our literature review suggest that FDPs' apps and web pages may provide an efficient and cost-effective medium for delivering digital messages and prompts that leverage psychological principles to influence safer worker behaviour. The sections below describe the digital aspects of the food delivery industry and how FDPs are already applying the principles within their systems to drive productivity and performance.

Algorithmic management

Delivery of goods is a longstanding profession, and short-distance commercial delivery drivers range from truck and car couriers to those using smaller vehicles, such as mopeds and bicycles. Technological advancement and innovation has revolutionised the profession and changed how work is organised. To solve the issues of time sensitivity relating to food delivery, online FDPs have evolved to use sophisticated algorithms to manage demand (customers), multi-restaurant offerings and just-in-time delivery workers at a moment's notice. As the majority of the interactions between workers and platforms occur through the use of an app or website, almost no human contact is needed to sign up, assess the level of experience, and start completing deliveries (De Stefano, 2015).

While online platforms have no or little control over the safety of the physical environment and the quality of gear used by its workers, platforms do have control over the digital infrastructure. Algorithms allow platforms to set the pace and flow of the delivery process and to monitor and evaluate worker performance and customer satisfaction. The design and tracking capabilities of the technical infrastructure, the information symmetry, the wording of messages and prompts, and the design of the user interface therefore play important roles in how work performance expectations are communicated, perceived and acted upon by workers (Veen et al., 2020).

Tracking and monitoring

The tracking capabilities commonly include the gathering of geospatial information feeds received from the GPS systems in the workers' smartphones as the workers are logged onto the app or website during their shift. The data capture informs platforms of the worker's every movement and includes information such as selected travel routes, travel speeds and job completion rates. This data is then used to inform decisions such as gig allocation, remuneration, and sometimes even punishment or termination of a worker (Newlands, 2020). In addition to geographical tracking, platforms also monitor customer satisfaction through customer-reported feedback and ratings of worker performance, e.g. a thumbs-up or thumbs-down, or a rating on a 1-to-5 scale. Both types of monitoring can create an increased sense of vigilance and fear of negative consequences, which can have significant impacts on workers' risk-taking behaviours and psychological health (Bajwa et al., 2018, Convery et al. 2020a).

Using information to influence behaviour

The data collected through tracking and monitoring systems are also generally not shared outside the FDP unless for specific purposes. Strategic uses of information can include managing customer expectations by, for example, allowing them to track the worker on a map as they travel from the restaurant to the customer. Another example is providing the customer rating to the worker after the delivery to encourage improved performance (Sun, 2018). As such, data can be used to inform or influence the beliefs and behaviours of workers and customers.

As the data is not shared equally, it can create information asymmetries between the FDP and the worker, resulting in a lack of understanding of how the algorithms work, how gigs are allocated and remunerated, how work performance is measured, what the expectations are and what the consequences are for not meeting those expectations. While some information asymmetries may simply result from the complexity of explaining the inner workings of algorithms, others are deliberately leveraged in order to nudge or otherwise influence worker decision making and performance (Veen et al., 2020). For example, by not providing the delivery address until after the worker has accepted the gig helps avoid gigs being rejected due to undesirable delivery locations or distances (Goods, Veen & Barratt, 2019). Making it easy for workers to do the desired behaviour or harder to do undesired behaviours are other ways to nudge improved performance. Examples of this include changing from an “opt-in” to an “opt-out” automatic gig acceptance feature in the app to increase the number of gigs that are accepted (Convery et al., 2020a). Similarly, adding a sense of urgency through limiting the time the job is available for acceptance or rejection forces workers to make decisions on the fly (Veen et al., 2020).

Another approach to drive productivity includes gamification. For example, Sun (2018) reported a platform in China whose workers could gather points based on their customer satisfaction ratings. At certain levels of accumulated points, high-performing workers would be promoted through increasing levels of “knighthood”. Advancement was rewarded by higher levels of pay, while lower performance would see workers demoted to lower levels of knighthood. Gamification and workers’ desire to maintain their current level of knighthood thus adds another dimension to more commonly used rewards for good performance. In Australia, it has been reported that pay rates often are adjusted through “boosts” or “bonuses” to encourage workers to take on more gigs at specific times or locations, or during periods of inclement weather (Goods, Veen & Barratt, 2019). Similarly, in Germany, Plöger & Keuneke (2021) reported the use of push notifications to alert workers to the opportunity for bonuses and nudge increased participation during periods of high demand. Push notifications were also used to give non-financial rewards such as praise for a job well done.

Influencing safer behaviours

As the examples in the previous section suggest, messages, prompts, and user interface designs have already been widely employed by FDPs to leverage psychological principles and influence worker behaviour in an efficient and low-cost way. Similar principles and interventions can be used to drive safety behaviours. For example, past research has shown that trip hazards in a factory could be decreased by using gold-coin stickers to change workers’ perceptions on housekeeping (Wu & Paluck, 2018). Research has also shown how behaviourally-based safety email messages had a larger and longer-lasting impact if sent to employees rather than managers; how stickers encouraging speaking-up improved safe driving; and helmet-wearing while riding a bike increases when others are also wearing a helmet (Behavioural Insights Team, 2019).

The study described in this report aims to leverage the FDP’s messaging infrastructure to use behavioural principles to reframe workers’ perception of delivery times as estimates rather than expectations and evaluate its effectiveness in reducing FDW risk exacerbation and/or increasing risk mitigation behaviours on the road.

Method

In the sections below, we describe the process of distilling the insights gleaned from Phases 1 and 2 into a co-designed behavioural intervention that addresses a priority safety issue (Phase 3), which was then trialled in the field with a major FDP (Phase 4).

Phase 3a: Selection of priority safety issue

Phase 1 identified road accidents, verbal abuse from customers and restaurant staff, and robbery or vandalism of delivery equipment as the primary WHS concerns of FDWs, while Phase 2 identified road accidents and COVID-19 as the primary WHS concerns of FDPs. Since road safety was the top common concern among stakeholders, it was chosen as the focus area for a behavioural intervention.

Internal workshop

The recommendations to improve road safety that were proposed in Phases 1 and 2 were discussed in a workshop attended by members of the project team, with representation from BIT, Macquarie University, and the Centre for WHS. Each idea was discussed in terms of its potential impact, feasibility, and "testability," i.e. its ability to be turned into a behavioural intervention that could be evaluated in a trial. The ideas were then ranked by the workshop attendees using the nominal group technique, an approach to the generation, discussion, and ranking of ideas that aims to reduce the effects of interpersonal and power dynamics within the group (Gallagher et al., 1993). This technique was applied in the current context by asking workshop attendees to anonymously assign asterisks representing votes to their preferred ideas in a shared online document.

The top six ideas were subsequently mapped to the food delivery workflow. As shown in Figure 1, two ideas were aimed at changing the platform itself (*change order acceptance parameters* and *make travel times more realistic and flexible*), two at improving the onboarding process (*increase the depth and breadth of safety knowledge* and *increase the uptake and retention of safety training*), and two at creating a culture of safety post-onboarding (*keep safety front-of-mind before a shift* and *keep safety front-of-mind during a shift*).

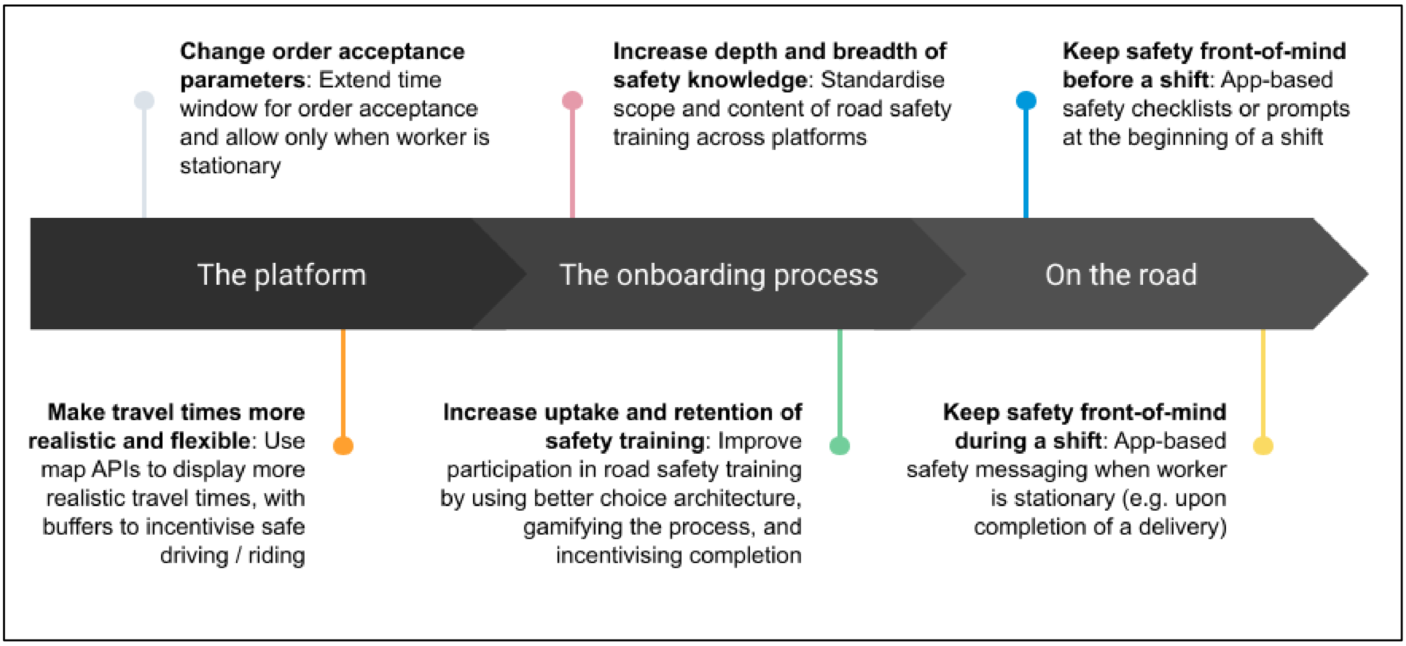


Figure 1: The six ideas derived from the Phase 1 and 2 findings.

1:1 interviews with FDWs and FDPs

The six candidate ideas formed the basis for 1:1 interviews with five FDWs and representatives of four major FDPs, recruited from among those who had participated in Phases 1 and 2 of the project. The interviewees were invited to review the ideas and to provide input during a 30-minute telephone conversation on what they liked and did not like about each idea and what they perceived as the broad appetite for the idea across their stakeholder group. The likelihood that each idea could be implemented in their platform was additionally discussed with the FDP representatives. A summary of their feedback is provided below. Since many of the interviewees considered the two ideas relating to improving the onboarding process (*increase the depth and breadth of safety knowledge* and *increase the uptake and retention of safety training*) as essentially inseparable, they are discussed below under a single subheading.

Change order acceptance parameters

Most FDWs and FDPs were receptive to the idea of extending the order acceptance window and could see the potential safety benefits of doing so, particularly the FDWs who deliver for platforms with narrow windows. Some FDWs proposed complementary strategies, such as presenting the order information in a more visually salient way or allowing orders to be accepted via voice interface. Support for only being able to accept orders when stationary was more polarised. Some FDWs believed that this would be very helpful, while others expressed concerns that it would be impractical, inefficient, or difficult to implement. Some FDPs further highlighted the potential unintended consequences, such as promoting unsafe stopping behaviour. Implementation barriers raised by the FDPs included the challenge of implementing locally relevant changes to the app when product changes are largely controlled by the global team.

Make travel times more realistic and flexible

Most FDWs were in support of an intervention to address this issue, expressing a belief that more realistic travel times would reduce stress and rushing. They suggested building in time for parking and locking up bikes. In contrast, FDPs perceived they already provided accurate travel times, which are derived from a combination of app data and map application programming interfaces (APIs). Further probing of these opposing perspectives revealed that while most FDPs intend delivery travel times to be an estimation, FDWs perceive them as an expectation. The FDPs described the purpose of the time window as information to enable FDWs to make an informed choice about whether or not to accept an order. However, the FDWs believed that if they did not consistently meet the time window shown in the app, they could be penalised by the platform in a range of ways, including losing access to prime shifts, being deprioritised for delivery offers, or being released from the platform. In actuality, FPDs employ a range of performance metrics, some of which are linked to delivery times, but the majority do not directly penalise FDWs for deliveries made within a reasonable time of the estimate displayed in the app. This misalignment in understanding highlighted an area for further discussion at the subsequent co-design workshop, which is described in detail in a subsequent section.

Increase the depth, breadth, uptake, and retention of safety knowledge

Almost all the FDWs we spoke to strongly believed that a well-designed and comprehensive safety training during onboarding, particularly focusing on Australian road rules and defensive riding/driving strategies (i.e. driving or riding under the assumption that others on the road are not thinking about your safety), would make the job safer for new recruits. Many FDWs viewed the current training as being piecemeal and wanted to ensure that future training programs were practically focused and had concrete learning outcomes. The FDPs also acknowledged the benefits of standardised training across the sector, noting that there would still need to be platform-specific training that each FDP would need to be responsible for. Many FDWs believed that training should be mandatory across all platforms. However, the FDPs noted constraints around mandating training for workers classified as "contractors" (like FDWs) as opposed to "employees." There were mixed opinions among the FDWs about how effective incentives would be at encouraging training. Some expressed that incentives may be attractive only to certain FDW subgroups, such as those only interested in the incentive, not in genuinely improving their safety knowledge, and FDWs who complete safety training regardless.

Keep safety front-of-mind before a shift

The FDPs were largely supportive of this idea, with most of them reporting that they already have a pop-up functionality within their apps that would enable the implementation of pre-shift safety information and checklists. The FDWs believed that pre-shift checklists would help them remember important safety information and form better habits, particularly if they focused on a diverse range of safety issues. However, both FDWs and FDPs noted that keeping content fresh was critical to the success of the initiative; otherwise, engagement with the information and checklists was highly likely to diminish rapidly.

Keep safety front-of-mind during a shift

The timing of any content sent out during a shift was highlighted by both the FDWs and the FDPs as an important consideration. Both groups agreed that content should only be sent when the FDW is stationary to avoid creating additional safety risks and at points during the order acceptance, pickup, and delivery process when the FDW would have the most time and cognitive resources to absorb and apply the information. FDWs were particularly interested in within-shift messages that were directly relevant to their current circumstances, including reminders to take breaks after working for a certain number of consecutive hours, wet weather warnings, turning on headlights at dusk, and reminders about variable speed limits and school zones.

Stakeholder co-design workshop

An online co-design workshop was convened with the aim of selecting the top two preferred ideas from the six candidate ideas and co-designing practical ways in which the ideas could be implemented and tested. The workshop was attended by the five FDWs who had provided initial input during the 1:1 interviews, seven representatives of four major FDPs (the four interviewees and three additional colleagues), and one representative of an e-bike rental shop. The six candidate ideas were reviewed with the workshop attendees, and a summary of the feedback received during the interviews was provided. All workshop attendees were given the opportunity to request clarifications or ask questions. Attendees were then asked to vote for a maximum of three preferred ideas on Slido, an anonymous online voting system. The results of the vote are shown in Figure 2.

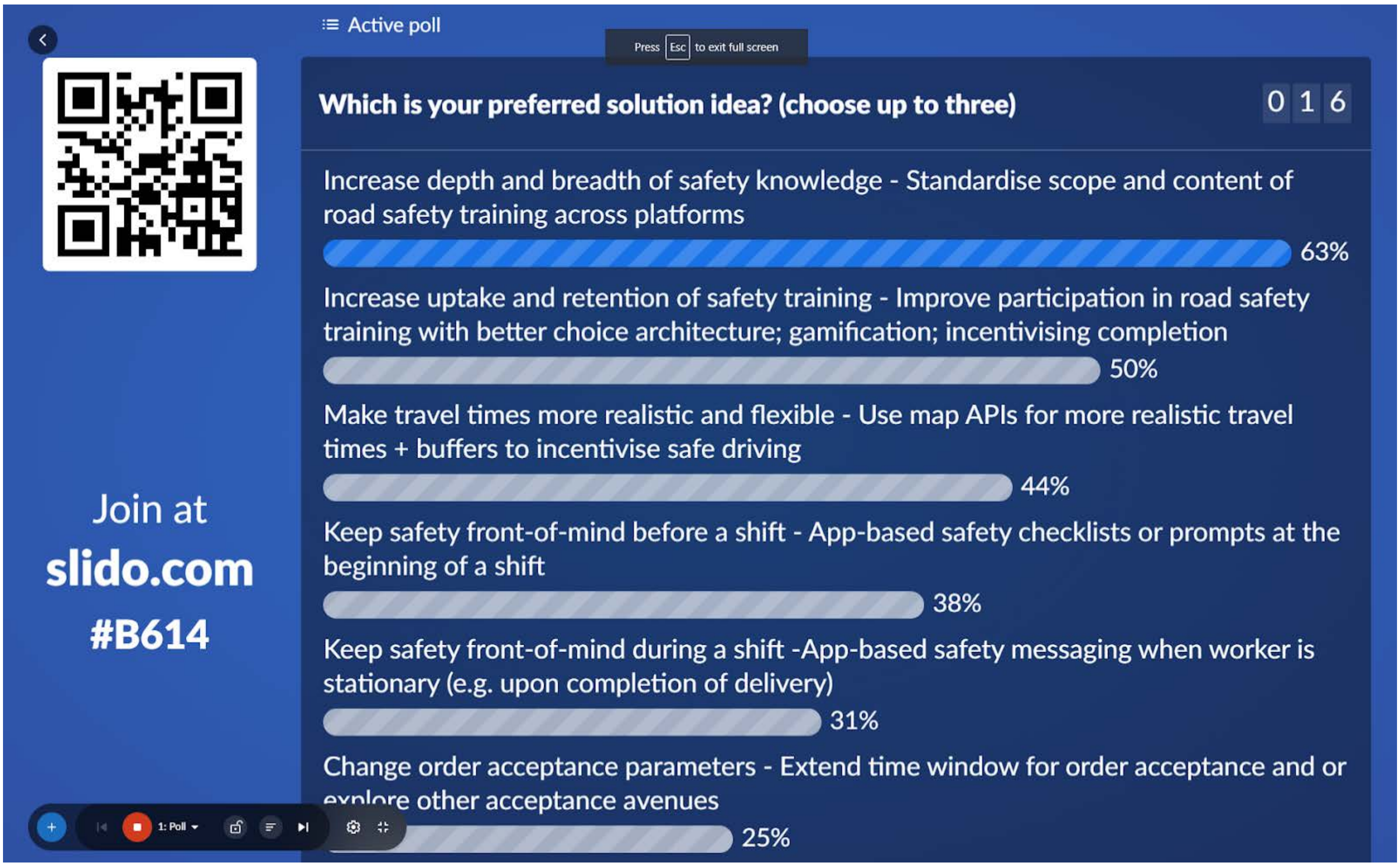


Figure 2: Results of the vote by FDWs and FDPs to identify the two highest priority areas for subsequent intervention co-design.

As a result of a discussion following the vote, it was decided that the top two ideas were very similar and should therefore be workshopped together. This meant that the focus of the subsequent co-design activities was on *increasing the quality and participation of road safety training* and *making travel times more realistic and flexible*.

Workshop attendees were assigned to three- or four-person breakout groups to begin the co-design process. Each breakout group was facilitated by a BIT member of the project team and included at least one FDW and one FDP representative. No group contained more than one representative from the same FDP. Two breakout groups were tasked with co-designing the first idea; the other two co-designed the second idea. The guiding questions for the breakout groups were:

1. What would the idea look like in practice? Think about different perspectives: what would it mean for food delivery workers? For e-bike suppliers? For platforms? For restaurants and customers?
2. What kind of impact do you think the idea would have on worker safety? Are there other impacts it could have?
3. What kind of obstacles might we face putting the idea into practice? How could we address them?

Increase the quality and participation of road safety training

The consensus of the breakout groups tasked with discussing this topic was that road safety training should be tailored to the individual FDW's vehicle type, with separate modules for bicycles, cars, and motorcycles/scooters. Defensive driving and riding tactics were viewed as an important component of any training. Comprehensiveness was seen as less important than focusing on issues that are specifically relevant to workers (e.g. riding safely with cargo). Crowdsourcing topics and approaches from FDWs could increase the relevance of, and engagement with, training materials, which could, in turn, be populated with information from vetted sources, such as SafeWork NSW and Transport for NSW.

At least some face-to-face training was favoured to increase engagement, trust, and knowledge retention. In-app messages were preferred over email messages since they are more likely to be read and absorbed, with the caveat that the timing of these in-app messages would be critical. A key theme of the training discussion was that road safety training should be ongoing, not just confined to onboarding.

Incentives for participating in training that were proposed by the group included "shock" messaging, similar to safe driving campaigns put out by the Transport Accident Commission in Victoria; safety-related incentives such as protective gear, hand sanitiser, and vouchers for vehicle maintenance; and in-app badges and flairs, although there was some question about how effective or appealing this would be.

In terms of feasibility for intervention development and a subsequent trial, a national rollout would require substantial resources, particularly from the government. In the shorter term, however, piloting new content and approaches in one state (i.e. as a proof of concept test) was considered feasible, with the caveat that time and resources would be required to build the technical infrastructure, prepare the content, and customise its "look" to align with individual FDP branding guidelines. Trials to evaluate optimal incentives for engaging in training were deemed highly feasible. However, there were concerns raised about the appropriateness of incentivising FDWs to participate in current training initiatives if those initiatives were already considered in need of improvement.

Make travel times more realistic and flexible

A key implementation challenge that emerged early on during the breakout session was the fact that each FDP uses a different combination of app data and map APIs to calculate delivery times, meaning that each FDP would have a different baseline for a trial. The need to involve the FDPs' global product teams to make a locally-focused change to the app compounded the challenge. Ways of accurately communicating the purpose and meaning of existing delivery window times thus became the focus for subsequent discussion within the two breakout groups assigned to this topic.

As noted above, FDPs consider delivery windows to be an estimation, while FDWs perceive them as an expectation, beyond which penalties could be incurred for failing to consistently meet the delivery times shown in the app. To address this problem, potential ideas discussed in the breakout group included communicating delivery times in terms of averages (e.g. "Ordinarily, workers make this type of delivery in around 15 minutes") or drawing on behavioural science principles such as anchoring[1] to reframe the concept of delivery windows as estimations.

Another option discussed in the breakout session focused on changing the messaging received by the customer, since managing customer expectations contributes to the pressures that FDWs feel on the job. Communications to the customer could let them know that the delivery window is only an estimation and that factors outside the FDW's control can result in food being delivered late. Customer communications could be framed to emphasise the importance platforms place on worker safety (e.g. "[Worker name] has encountered some unexpected traffic delays. Your food will now be arriving at 7:10 so he can navigate the heavy traffic safely.").

In terms of feasibility for intervention development and a subsequent trial, FDPs already have multiple methods of communicating with their FDWs and customers (e.g. in-app messages, email) that are under local, rather than global, control, meaning the infrastructure for a messaging trial would already be in place. Since the primary outcome of any trial would relate to FDW safety, the breakout groups agreed that a message intervention targeted at FDWs, rather than customers, would probably stand a greater chance of having a positive impact on FDW safety.

Phase 3b: Co-design of behavioural intervention

A messaging trial targeting FDWs' belief that delivery windows represent expectations, rather than estimates, was ultimately chosen as the focus for intervention development. This was because the project team judged it to offer the best balance of impact, feasibility, and testability relative to the other ideas. The sections below describe the theory of change underlying the intervention design, the behavioural principles that underpin each intervention message, and the process of co-designing and refining the messages with FDWs.

Theory of change

In order to develop evidence-based strategies to target the safety behaviours of FDWs, we first created a theory of behaviour change. Theories of behaviour change map out how and why a change in behaviour is expected to happen in a particular context. They summarise causal pathways, describing what is known about various psychological and behavioural constructs to explain and predict how change may occur. The behaviour change literature points to the importance and value of theory-based interventions, both to develop evidence-based behavioural

[1] *Anchoring* is a cognitive bias in which people overweight the first piece of information they receive when making subsequent estimates or decisions (Tversky & Kahneman, 1974).

interventions and to communicate these to other researchers in the field in a way that facilitates interpretation, reproduction, and scalability (Michie et al., 2018).

Our theory of change was informed by Phase 1 and 2 qualitative and quantitative research into the road safety behaviours of FDWs (Convery et al., 2020a) and the wider behavioural science literature. As illustrated in *Appendix A*, the theory of change specifies several key opportunities for in-app messaging interventions (the red boxes labelled 1 and 2), linking behavioural mechanisms of action (grey diamonds) to FDW attitudes, beliefs, and behaviours targeted for change (red and blue diamonds). Our theory of change also specifies where on the causal pathway we planned to measure behaviour change (see *Phase 4: Randomised controlled trial of behavioural intervention* below for further detail on the outcome measures that were ultimately selected for our trial of the messaging intervention).

Influence through behavioural messaging

We designed four messages for our subsequent trial. The first message targeted FDW beliefs about the consequences of making late deliveries. The remaining three messages aimed to reinforce this updated belief and provided strategies to help FDWs reduce risk-taking behaviours in the moment.

Four of the FDWs who participated in Phase 3 were asked to review the draft messages and to participate in a short telephone interview with a member of the project team. During the interview, the FDWs were asked what they thought each message was trying to convey, whether the messages made sense to them, and whether they thought the wording was appropriate and understandable. The feedback participants provided on the messaging was primarily positive. Messages were referred to as "a really important message," "great," and that they "do [their] job well". One FDW said:

> *"[The messages are] mostly talking about your health and wellbeing as a driver... things happen along the way while you're delivering... it'd give you a morale boost, it'd mean you don't need to worry."*

In contrast, one FDW noted that he rarely reads the messages he receives through the FDP app at the beginning of the shift. He liked the central idea of each message, but believed more intensive changes would be needed for them to be optimally effective in existing FDWs. Changes suggested by the FDW primarily focused on aligning the timing of the message to an actual delay, rather than sending a more general message at the beginning of a shift:

> *"If it appears at the beginning of my shift, I would be calm momentarily. I think this message is important more if it's used at the correct time. When a problem is happening. More than just a reminder at the beginning of the shift."*

Overall, the four FDWs believed that the messages would reduce uncertainty because they gave concrete strategies and reassurances that unexpected delays were normal, which in turn would reduce stress and the likelihood of rushing and speeding behaviours.

The sections below describe each message in detail, including the theoretical mechanisms of action and supporting evidence for each of the messaging strategies.

Message 1: Beliefs about consequences

Our first message (Figure 4) focused on altering FDWs' belief that not meeting the delivery windows provided by FDPs led to negative consequences or penalties, such as negative ratings. Updating individuals' beliefs that a behaviour will lead to a specific outcome is a behaviour change strategy included in the Behaviour Change Technique (BCT) taxonomy, a resource of 93 behaviour change techniques developed by expert behaviour change practitioners (Michie et al. 2013). Updating individuals' beliefs have been shown to change dietary behaviours (Geraerts et al., 2008) and promote smoking cessation (Brown et al., 2019).

Our previous research found that many FDWs perceived that they would be penalised for not meeting the delivery windows provided by FDPs, thus providing an incentive to rush (Convery et al., 2020a). However, in our co-design workshops, we found that most major FDPs did not penalise late deliveries, except in cases of significant and repeated late behaviour.

FDWs were also concerned about customer satisfaction if the delivery was running late (Convery et al., 2020a), which potentially incentivises risk-taking. While FDPs employ various strategies to mitigate this risk, such as providing regular updates to the customer, our co-design workshops revealed that FDWs were unaware of this messaging. Consequently, our first message (Figure 3) focused on correcting the misperception that FDPs penalised FDWs for not meeting delivery times by telling FDWs that the FDP would advise the customer if the food was running late.

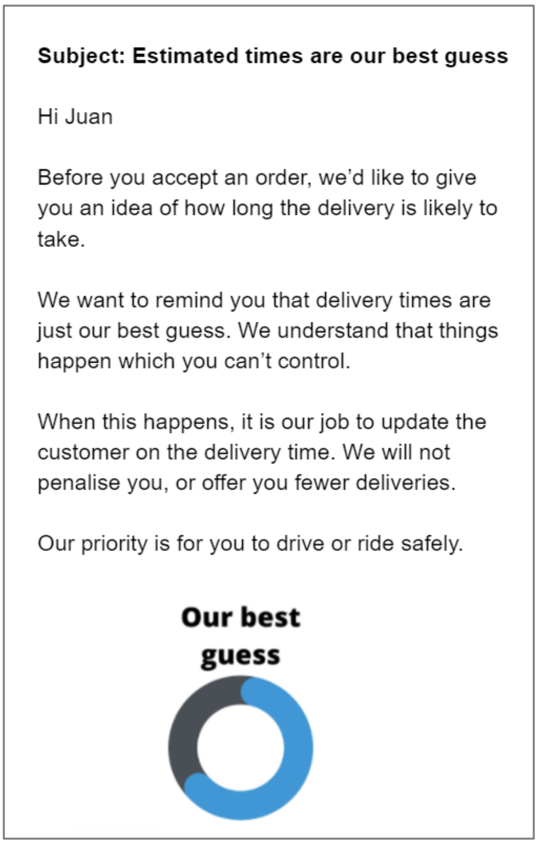


Figure 3: The first message sent to FDWs targeted the belief that not meeting delivery times would result in negative consequences or penalties.

Reinforcing beliefs about consequences and addressing the intention-action gap

In the remaining three messages, we sought to achieve two primary goals: (1) to reinforce the beliefs about consequences message; and (2) to address other known barriers to FDWs performing their job safely. However, we know from the behavioural science literature that changing people's beliefs is often not sufficient to change behaviour (Glasman & Albarracín, 2006) and that people do not always do what they intend to do, a phenomenon known as the *intention-action gap* (e.g. Sheeran & Webb, 2016). Since our previous research showed that FDWs largely *intend* to engage in safety-promoting behaviours (Convery et al., 2020a), we sought to provide them with pragmatic strategies for *acting* on those intentions in Messages 2-4, thus closing the intention-action gap. The adopted behavioural strategies are described below, and included anchoring, communicating descriptive social norms, and implementation intentions.

Message 2: Anchoring

In our previous research, we discovered that the delivery windows provided by FDPs did not always account for unanticipated delays, such as finding a parking space or locking up a bike (Convery et al., 2020a). In combination with the belief that delivery windows were firm expectations, delays can become a significant source of stress for FDWs and can incentivise unsafe road behaviour. To address this, we introduced the concept of *safety time* in Message 2 (Figure 4), giving FDWs explicit "permission" to factor in the added time necessary to ride or drive safely during the delivery. In framing the delivery window as a range, rather than a fixed value, we drew on the behavioural principle of anchoring. As noted above, anchoring refers to a cognitive bias that describes the tendency to rely heavily on reference points when making decisions (Tversky & Kahneman, 1974).

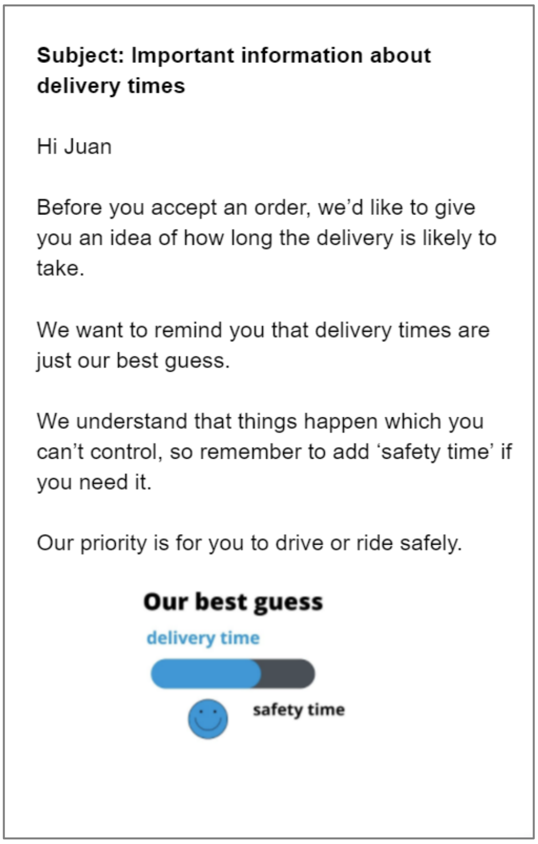


Figure 4: Applying the behavioural principle of anchoring, the second message sent to FDWs introduced the concept of 'safety time' to encourage FDWs to allow extra time when unexpected delays occur.

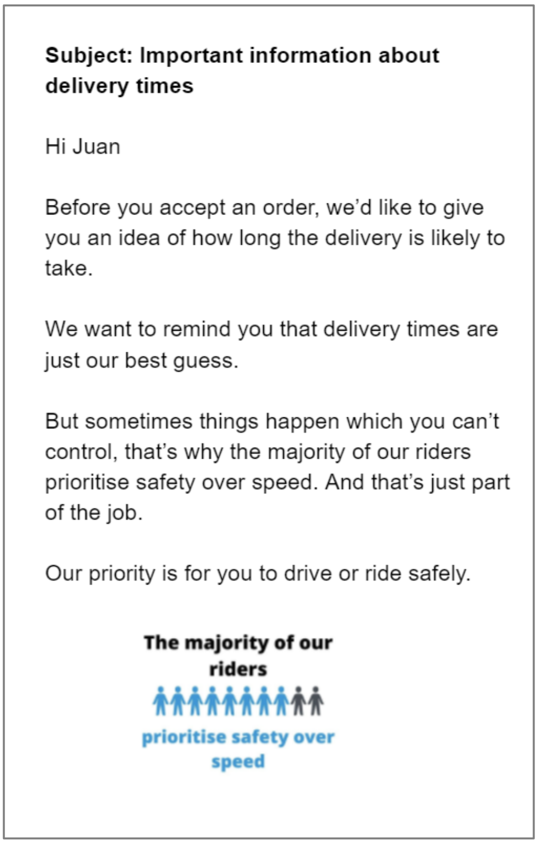
Message 3: Descriptive social norms

People are heavily influenced by what others think and do, particularly when they perceive a shared identity with others they encounter (e.g. Cialdini, 2007; Deutsch & Gerard, 1955). This is a reflection of our innate tendency for cooperation and conformity, which can be harnessed by leveraging the influence of peer groups and social norms.

Our previous research identified various mechanisms through which FDWs safety behaviour may be influenced by other FDWs. These included directly taking up advice from other FDWs (e.g. provided on social media channels), as well as more indirect opportunities for influence, such as witnessing multiple FDWs working in the same area (Convery et al., 2020a).

Describing what most people do in a particular situation encourages others to do the same, a behaviour change strategy known as communicating *descriptive norms*. Communicating descriptive norms has been applied to positively change various behaviours, including the use of child safety seats (Jeffrey et al., 2016), energy use (Schultz et al., 2007) and tax compliance (The Behavioural Insights Team, 2014). One study found that communicating descriptive norms about the recycling rates of hotel towels by other guests led to increased recycling rates (Goldstein et al., 2008).

When we communicate descriptive social norms about a target behaviour, it is important to advertise the behaviour we want to see (e.g. Most FDWs perform their job safely) rather than the behaviour we want to avoid (e.g. Many FDWs take risks). This encourages people not currently performing the target behaviour to shift toward the behaviour most other people are doing. To apply this strategy, we communicated the positive descriptive social norms uncovered in the first phase of our research: that the majority of FDWs prioritised safety over speed (Figure 5).



Subject: Important information about delivery times

Hi Juan

Before you accept an order, we'd like to give you an idea of how long the delivery is likely to take.

We want to remind you that delivery times are just our best guess.

But sometimes things happen which you can't control, that's why the majority of our riders prioritise safety over speed. And that's just part of the job.

Our priority is for you to drive or ride safely.

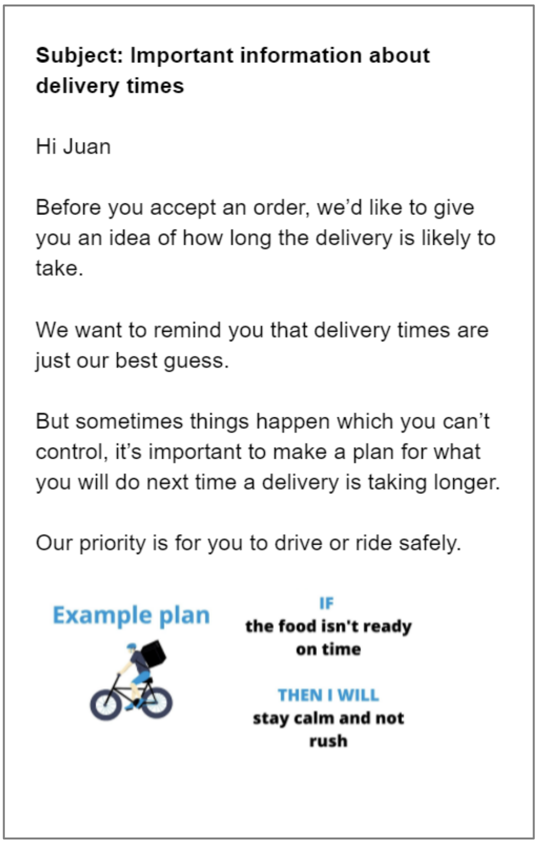
Figure 5: The third message sent to FDWs encouraged social comparison of their behaviour with their peer group to reinforce our observation that the majority of FDWs perform their job safely.

Message 4: Implementation intentions

We often have good intentions but behave differently in the moment. This is partly because people underestimate the influences of visceral drives on their own attitudes and behaviours when they are in aroused or heightened emotional states, a cognitive bias known as the hot-cold empathy gap (Van Boven et al., 2013). While there are various potential motives and incentives for FDWs to rush, such as financial incentives due to piece-rate pay (Convery et al., 2020a), one trigger for speeding was delays outside the FDWs control (e.g. the restaurant running late). These delays can cause stress for FDWs, who in turn take risks to make faster deliveries.

Advance planning helps people respond in the moment in a way that moves them closer to their true goals rather than away from them (Gollwitzer, 1999). Implementation intentions are exercises that specify when, where and how a person intends to complete a goal. They have been used to change behaviour across various domains (e.g. Gollwitzer & Sheeran, 2006; Belanger-Gravel et al., 2011) and are useful for overcoming both expected and unexpected hurdles or obstacles to goal attainment (such as, in this instance, restaurant delays). A meta-analysis of findings from 94 studies found that implementation intentions had a medium to large (d = .65) positive effect on goal attainment (Gollwitzer & Sheeran, 2006). For instance, one study found that women who were prompted to make a plan of where and when they would conduct a breast self-examination were 4.5 times more likely to have done one a month later (Orbell et al., 1997).

To apply this principle, we prompted FDWs to create a plan for how to ensure their own safety when they experience in-the-moment time pressures. We did this through a combination of direct instructions and the provision of an example plan (Figure 6).


Subject: Important information about delivery times

Hi Juan

Before you accept an order, we'd like to give you an idea of how long the delivery is likely to take.

We want to remind you that delivery times are just our best guess.

But sometimes things happen which you can't control, it's important to make a plan for what you will do next time a delivery is taking longer.

Our priority is for you to drive or ride safely.

Figure 6: Applying the behaviour change technique of implementation intentions, the fourth message sent to FDWs encouraged them to make a plan for how to perform their job safely should unexpected delays occur.

Phase 4: Randomised controlled trial of behavioural intervention

A two-armed, four-week RCT of the co-designed messages was conducted in April and May 2021 in collaboration with one of the FDPs that participated in Phase 3 of the project (referred to hereafter as the *trial partner FDP*). The overall aim of the trial was to determine whether messages that communicate pickup and delivery time windows as estimates rather than expectations reduce risk exacerbation behaviours and/or increase risk mitigation behaviours on the road. We focused on travel speed as a proxy for unsafe riding or driving behaviour. Specifically, we sought to determine the effect of the messages on:

1. Average travel speeds on a per-delivery basis
2. FDW beliefs that the trial partner FDP penalises FDWs for late deliveries
3. FDW perceptions of how safe they feel on the road while delivering for the trial partner FDP
4. Performance metrics (delivery times and customer satisfaction)

The messages were sent directly to the FDWs by the trial partner FDP via the messaging function of their delivery app.

The process through which the messages were co-designed, and the behavioural science principles that underpin the messages, was described above in *Phase 3: Co-design of behavioural intervention.* The details of the trial design and implementation are explained below.

Power calculations

We expected a meaningful change in average travel speed would be, at a minimum, 2 percentage points. Concretely, for a vehicle travelling at a maximum average speed of 25 km/h, we aimed for our intervention to decrease the travel speed by 0.5 km/h at a minimum. This is our target for a minimum detectable effect size (MDES), marked on the chart in Appendix B. For these calculations, we assumed a standard deviation in our outcome measure at 25% of the total travel speed, an alpha of .05, and 80% power. The calculations indicated that to meet these requirements, a sample size of 5,000 FDWs would be necessary. Given that our main analyses focused on delivery-level measures (and a higher sample size), this represents a conservative estimate (i.e. accounting for the unlikely event that each FDW made only a single delivery).

Study population

In order to achieve sufficient statistical power as per our *a priori* calculations, the study participants consisted of FDWs across all vehicle types who deliver for the trial partner FDP in Greater Sydney and Greater Melbourne and who had been active on the platform between 1 January and 21 April 2021 (the day before the first day of the trial). Participants were allocated to either the treatment or control group, stratified proportionately by vehicle type to ensure both groups contained the same proportion of FDWs who use motorcycles (including scooters), bicycles (including both pushbikes and e-bikes), and cars.

Ethical approval to conduct the trial (Application number: 2021-03-252) was granted by Bellberry, Inc., a private, not-for-profit organisation that is accredited by the National Health and Medical Research Council (NHMRC) to provide scientific and ethical review of Australian human research projects. The treatment of participants in the trial conformed in all respects to the NHMRC's National Statement on Ethical Conduct in Human Research (2007; updated 2018).

The trial was pre-registered on the American Economic Association (AEA) RCT Registry on 15 March 2021 (RCT ID AEARCTR-0007360).

Study procedure

The trial partner FDP sent the messages to the FDWs in the treatment group via their in-app messaging function on four consecutive Thursdays in the following order: (1) beliefs about consequences; (2) anchoring; (3) descriptive social norms; and (4) implementation intentions. Messages were timed to ensure they appeared when the FDW first logged on to the app to begin a shift. This was to ensure the messages did not exacerbate safety risks by appearing while FDWs were actively riding or driving. The control group did not receive any messages.

At the end of the trial, all participants (both treatment and control) were sent a two-question survey. The wording of both questions was reviewed with two of the FDWs who participated in Phase 3 to ensure they accurately captured the constructs we aimed to measure and used accessible, understandable vocabulary. The questions are specified in the next section.

Outcome measures

Table 1 describes the pre-specified outcome measures for the trial, including details about how the data for each measure was collected.

Table 1: Description of the trial's outcome measures and the source of these data.

Type of outcome	Outcome measure	Data source
Primary	*Average travel speed on a per-delivery basis*	*Derived by dividing time points by distance travelled from the FDP's telemetry recordings*
Secondary	*Subjective belief that the FDP penalises FDWs for late deliveries*	*In-app survey question 1: Do you think that [name of FDP] will penalise you if you're a few minutes late for a delivery? Yes / No / Prefer not to answer*
	Perception of safety on the road while delivering for the FDP	*In-app survey question 2: How safe did you feel on the road when delivering for [name of FDP] in the past month? Very safe / Safe / Neither safe nor unsafe / Unsafe / Very unsafe / Prefer not to answer*
Exploratory	*Average delivery duration*	*FDP data*[2]
	Customer ratings	

[2] *FDP data* refers to administrative data that is routinely collected by the FDP in the course of normal business operations. These data were provided to two of the co-authors (E.C. and B.F.) in de-identified form.

Covariates

Table 2 describes the covariates that were used in the analysis. As with the exploratory outcomes listed in Table 1, all covariate data was obtained as de-identified administrative data from the FDP.

Table 2: The covariates used in the analysis, including a description of each variable and its purpose.

Variable name	Variable description	Purpose
Vehicle type	*Which vehicle the FDW uses to deliver (bicycle or e-bike, motorcycle or scooter, car)*	*To account for differences in the average travel speed of each vehicle type*
Experience	*How long the FDW has delivered for the FDP*	*To account for differences in delivery behaviour as a result of experience*
Customer rating	*Current customer rating of the FDW*	*To account for differences in behaviour as a result of performance*
Deliveries	*Number of deliveries made by the FDW*	*To act as a proxy for fatigue*

Data analysis

The primary analysis consisted of a simple mixed-effects regression model and a full mixed-effects model that additionally included the covariates listed in Table 2 above. These approaches modelled the individual identifier for each FDW as a random intercept in order to account for individual differences in baseline delivery speeds. The secondary and exploratory analyses consisted of a simple ordinary least squares (OLS) regression model including only the dependent variable and treatment allocation and a full regression model that included the covariates listed in Table 2 above. A binary OLS regression model was used for the binary secondary outcome measure (the subjective belief that the trial partner FDP penalises FDWs for late deliveries). Two approaches were used:

1. Intention-to-treat analysis. An intention-to-treat (ITT) analysis includes all participants in a prospective RCT in the analysis, regardless of which treatment, if any, they actually received (McCoy, 2017). For the current analysis, this meant including the full treatment group (and the full control group) in the regression model.
2. Effect of treatment on the treated analysis. Effect of treatment on the treated (ETT) is an analytical method in which only those trial participants who actually receive the treatment are included in the analysis (Geneletti & Dawid, 2011). For the current analysis, this meant including only the FDWs who opened at least one of the four messages (and the full control group) in the regression model.

Results

In this section, we first report descriptive statistics about the FDWs who participated in the trial. We then report and discuss the effect of the messages on the primary, secondary, and exploratory outcomes.

Descriptive statistics

The dataset represented 1,616,016 deliveries completed by 14,281 individual FDWs during the four-week trial period. Of these, 7,066 FDWs had been allocated to the treatment group and 7,215 to the control group. As shown in Figure 7, the majority of FDWs used cars, with smaller proportions of FDWs using bicycles (including pushbikes and e-bikes) or motorcycles (including scooters).

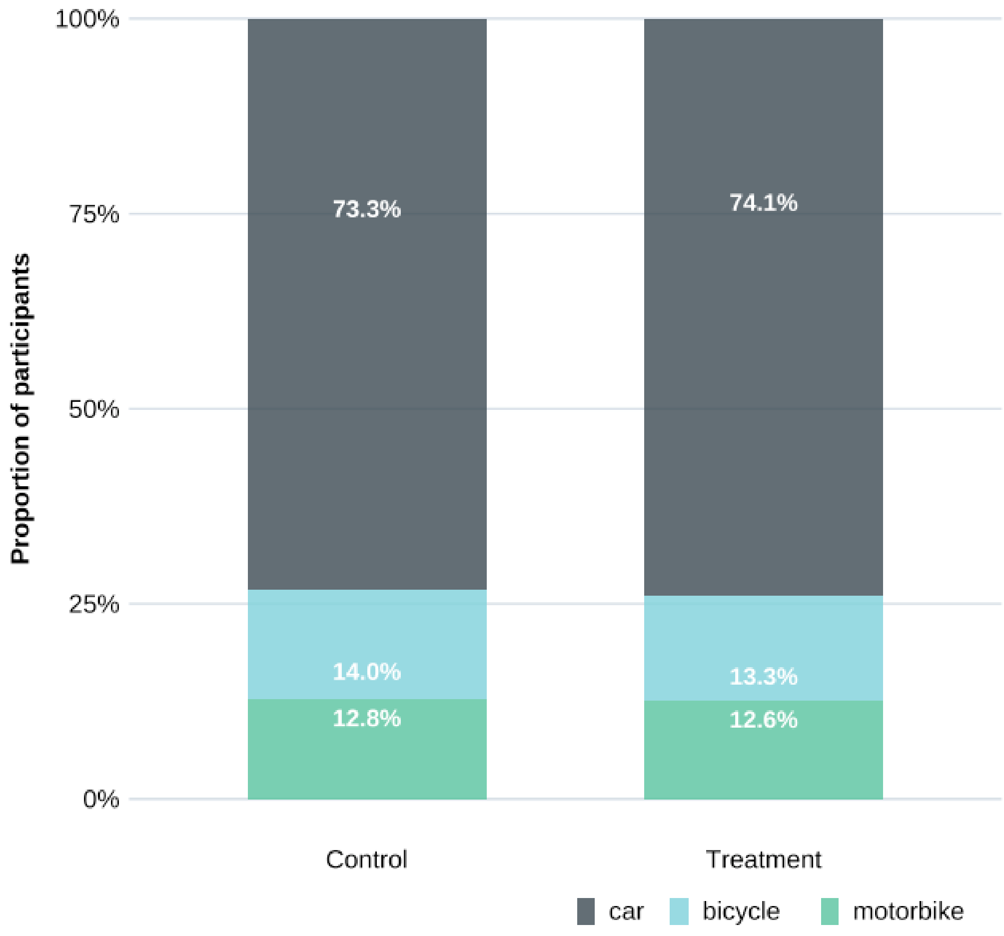


Figure 7: Distribution of vehicle type for the treatment and control groups.

There was no significant difference between the treatment and control groups on the proportion of vehicle types, mean length of experience, or mean customer rating. There was a significant difference in the number of deliveries undertaken during the trial period between the treatment group (average of 110 deliveries per FDW) and the control group (average of 117 deliveries per

FDW). However, the greater average number of deliveries in the control group did not disproportionately affect the outcome since this covariate, as well as the other covariates (vehicle type, experience, and customer rating), was controlled for in the regression analyses.

Regression analyses

In the sections below, we describe the results of the regression analyses that were conducted to determine the effect of the intervention on the primary, secondary, and exploratory outcomes.

Primary outcome

The primary ITT analysis aimed to determine the effect of the messages on average travel speed using OLS regression, estimating heteroscedasticity robust standard errors. The analysis was conducted at the *delivery level*, meaning that average travel speed was calculated for each individual delivery. Vehicle type, experience, customer rating, and deliveries[3] were included as covariates in the regression model.

The messages had no effect on average travel speed at the delivery level

There was no significant difference in average travel speed between FDWs in the treatment and control groups at the delivery level (p = .49). Table 3 shows the regression tables for the models with and without covariates. As expected, there was a greater likelihood of faster travel speeds among FDWs using cars and motorcycles (relative to those on bicycles), FDWs with more experience, and FDWs with higher customer ratings. There was a greater likelihood of slower travel speeds among FDWs who make more daily deliveries, which could be a consequence of greater fatigue in this group.

[3] For each delivery-level observation, the covariate *deliveries* was calculated as the number of deliveries made by the FDW on the day of that delivery.

Table 3: Regression table showing coefficients for the models with and without covariates.

Variable name	Univariate model (no covariates)	Multivariate model (with covariates)
(Intercept)	*14.041 ****	*8.654 ****
	(0.050)	*(0.117)*
Treatment	*0.102*	*0.037*
	(0.071)	*(0.058)*
Cars		*6.148 ****
		(0.084)
Motorcycles		*3.780 ****
		(0.106)
Experience (quartile)		*0.146 ****
		(0.027)
Customer rating (quartile)		*0.196 ****
		(0.027)
Deliveries		*-0.030 ****
		(0.001)
N	*1,586,563*	*1,586,563*

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; + $p < 0.1$

Figure 8 shows the results of the multivariate analysis. The control estimate in the graph is taken from the intercept in the simple regression model, and the treatment estimate is added to this.

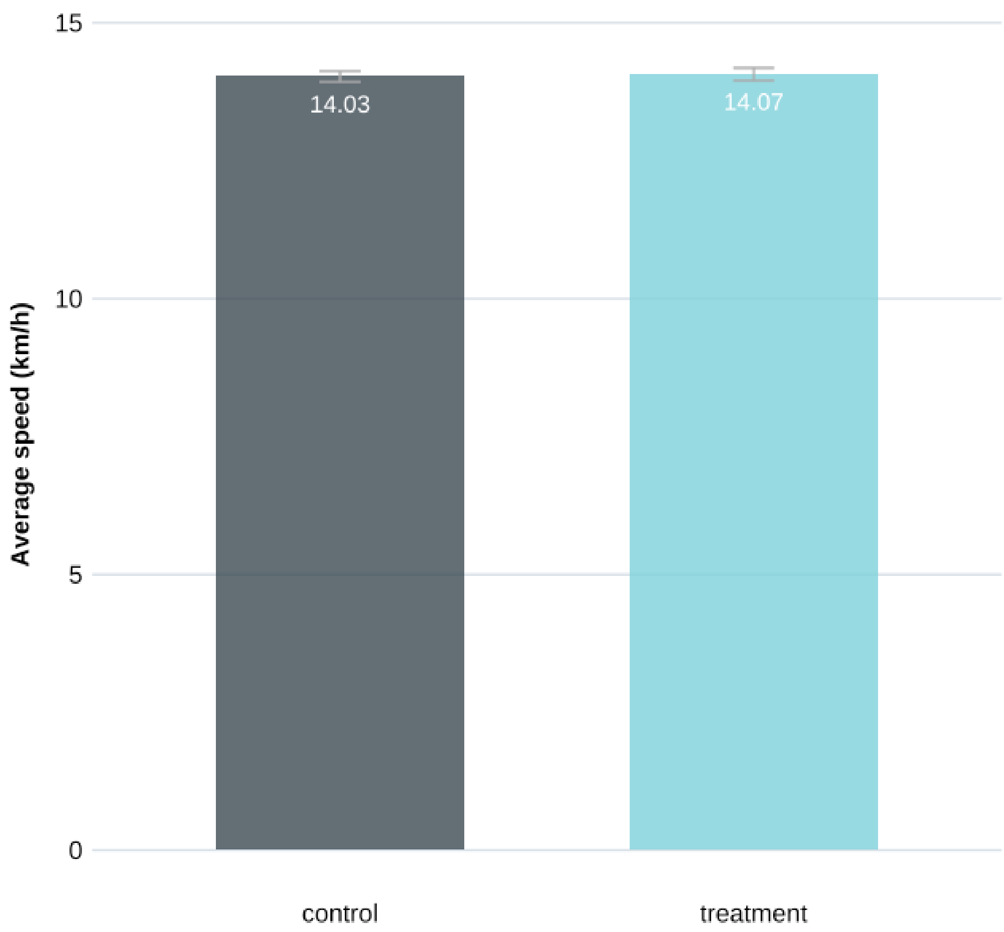


Figure 8: Results of a regression analysis showing no effect of the messages on average travel speed at the delivery level.

Only a minority of FDWs in the treatment group engaged with the messages

Messages can only have an effect on the recipient if they are opened and read. We, therefore, examined the proportion of FDWs in the treatment group who were recorded as having opened the messages each week. As shown in Figure 9, 1,051 FDWs (14.9% of the treatment group) opened the first message. The proportion of individual message opens was very similar for each of the four messages, with 16.5%, 19.2%, and 19.1% opening the second, third, and fourth message, respectively. 67.1% of the FDWs in the treatment group did not open any messages, 12% opened a single message, 8.8% opened two messages, 8% opened three messages, and 4.4% opened all four messages.

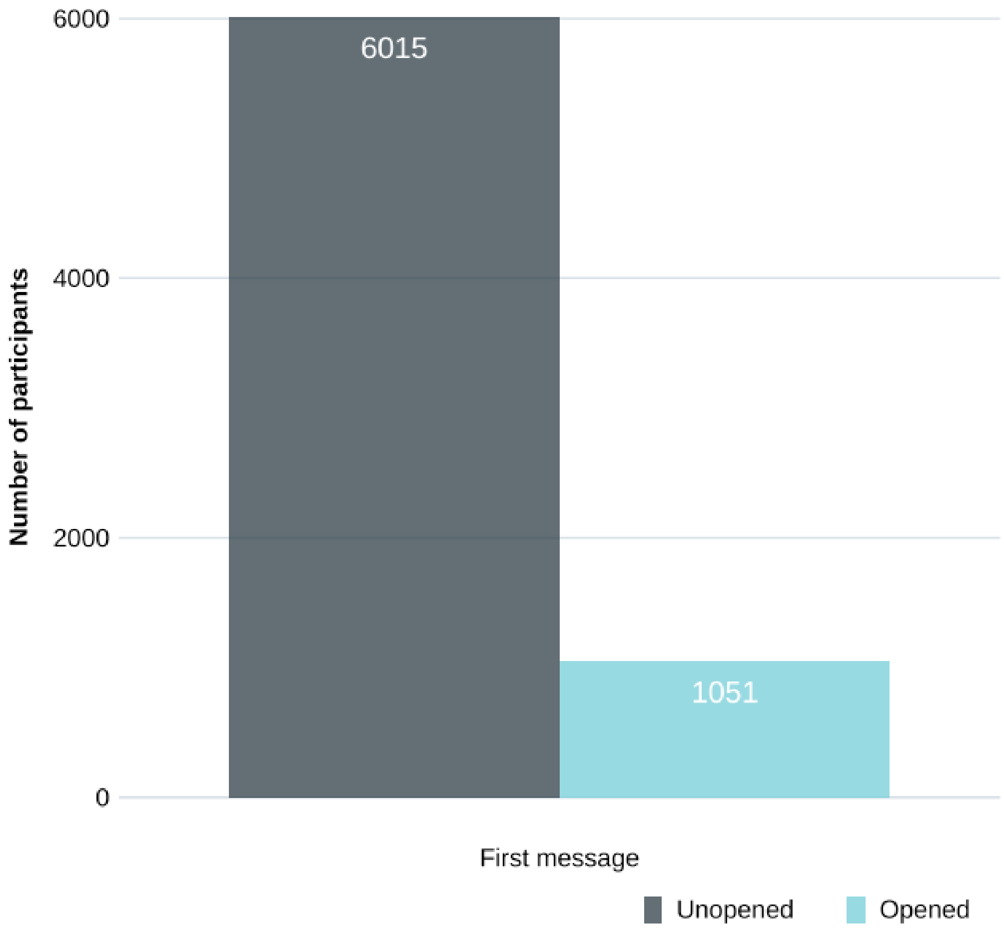


Figure 9: Number of FDWs in the treatment group who opened the first of the four messages relative to the number of FDWs in the treatment group who opened no messages.

FDWs in the treatment group who opened the messages, FDWs in the treatment group who did not open the messages, and FDWs in the control group were compared to determine whether they differed from each other on any of the measured covariates (Table 4). Experience and customer ratings are shown in the table as the proportion of FDWs in each quartile. For the experience and customer rating covariates, Quartile 1 (Q1) represents FDWs with the greatest experience or highest customer ratings, while Quartile 4 (Q4) represents FDWs with the least experience or lowest customer ratings.

The comparison shows that the distribution of vehicle types, experience, and customer ratings were similar for each of the three groups, with the exception of greater representation of car drivers in the group of FDWs who did not open the messages. This may be due to differences in how FDWs who use different vehicle types interact with their mobile phones on the job.

Table 4: Distribution of vehicle types, experience, and customer rating for FDWs in the treatment group who opened the messages, FDWs in the treatment group who did not open the messages, and FDWs in the control group.

Covariate	Category	Treatment group (opened messages)	Treatment group (did not open messages)	Control group
Vehicle type	*Bicycle*	*0.172*	*0.114*	*0.140*
	Car	*0.667*	*0.777*	*0.732*
	Motorcycle	*0.162*	*0.109*	*0.128*
Experience	*Q1*	*0.282*	*0.234*	*0.253*
	Q2	*0.256*	*0.239*	*0.255*
	Q3	*0.224*	*0.273*	*0.247*
	Q4	*0.238*	*0.254*	*0.246*
Customer rating	*Q1*	*0.330*	*0.328*	*0.312*
	Q2	*0.227*	*0.230*	*0.239*
	Q3	*0.277*	*0.262*	*0.276*
	Q4	*0.165*	*0.179*	*0.172*

The low rate of message opens highlights the challenges of engaging FDWs as a singular cohort

The low rate of message opens in this trial is consistent with previously reported challenges engaging FDWs as a singular group. In Phase 2 of this project, FDPs reported a similarly low rate of message opens for other in-app communications they send out to their FDWs and low rates of engagement with non-mandatory initiatives such as WHS training (Convery et al., 2020b). Similarly, surveys conducted among FDWs typically have low response rates, regardless of whether they are conducted by independent researchers (Convery et al., 2020a; Young Workers' Centre, 2018) or by FDPs (Convery et al., 2020b).

These past experiences highlight a key barrier to engaging FDWs: lack of a strong group identity. As identified in Phase 2 of this project, the majority of FDWs are temporary visa holders (largely students) who have worked as FDWs for less than one year and who view the money they earn from food delivery work as supplementary, rather than primary, income (Convery et al., 2020a). This suggests that for many FDWs, their identity may be more strongly tied to other aspects of their lives. For example, an FDW may identify primarily as a university student, a future professional in their current field of study, or a member of a particular cultural or linguistic group rather than as an FDW.

Social identity theory states that social identities are most influential when people consider their membership in a particular group to be central to their self-concept (Tajfel, 1978). People who see themselves as belonging to a particular group will behave in accordance with that social identity, conforming to the norms of that group. Engaging people on the basis of their membership in a group with which they only loosely identify thus makes changing the behaviour of that group more challenging. For example, since social identity theory is the basis for using descriptive social norms to change behaviour (as was done in the third message in the trial), this behaviour change strategy may be less effective when deployed within a group of people who do not share a strong group identity.

Average travel speed decreased among FDWs who engaged with the messages

An ETT regression analysis was conducted to determine the effect of the messages on average travel speed among FDWs who had opened at least one message, estimating heteroscedasticity robust standard errors. As with the ITT analysis described above, this analysis was conducted at the delivery level, and the covariates included in the model were vehicle type, experience, customer rating, and deliveries. The results of the ETT analysis showed a modest but significant ($p < .001$) effect: the average travel speed among the FDWs who read any messages travelled an average of 0.41 km/h slower than the FDWs in the control group, or a 3% decrease in speed (Figure 10). Although the results are statistically significant, they should be interpreted with caution, given that this is an ETT analysis. It is plausible that the demonstrated effect is being driven by underlying factors that were not measured as part of this trial. For example, it is possible that personal factors, such as high conscientiousness or high levels of risk aversion, may have driven both a higher rate of message opens as well as slower travel.

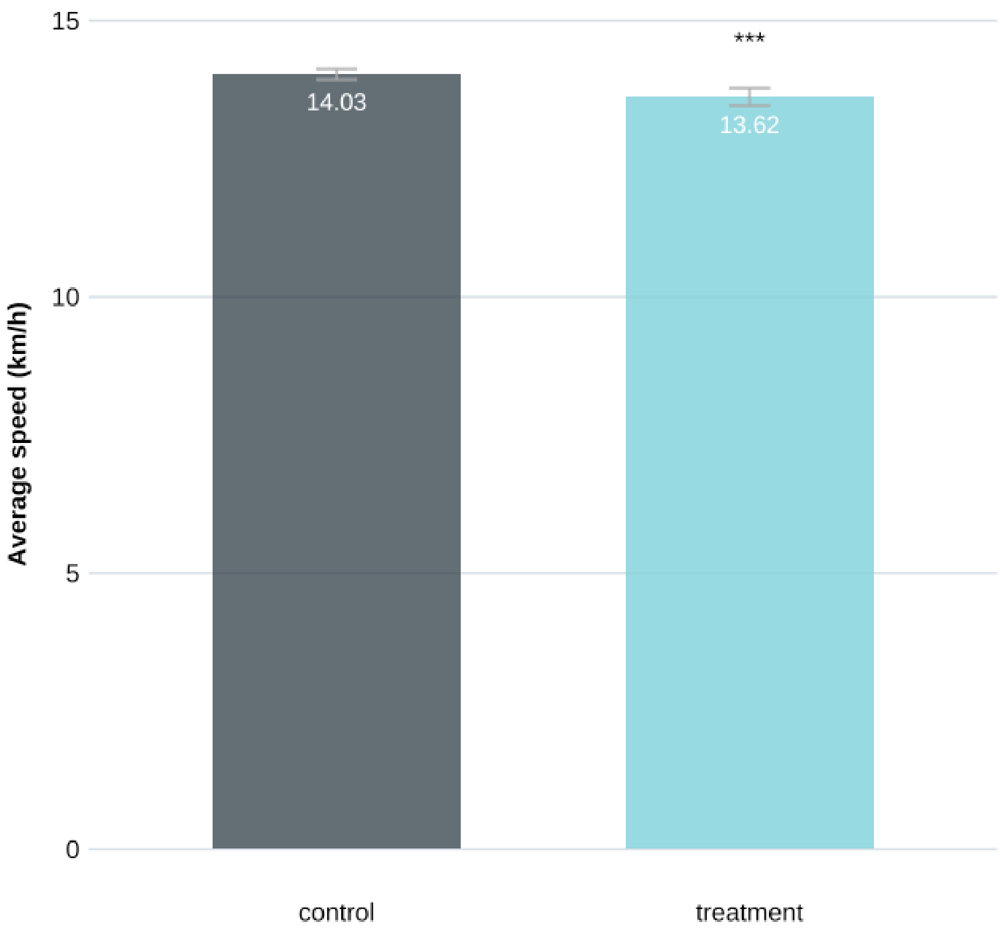


Figure 10: Results of a regression analysis showing a modest effect of the messages on average travel speed among FDWs in the treatment group who opened at least one of the four messages.

Subgroup analyses on primary outcome

ITT and ETT subgroup analyses were conducted on the primary outcome. The aim of the subgroup analyses was to determine whether there were any differential effects of the messaging on FDWs with different vehicle types or levels of experience. For both ITT and ETT analyses, there were no differential effects by vehicle type or level of experience.

Secondary outcomes

The secondary analyses aimed to determine the effect of the messages on (1) the subjective belief that the trial partner FDP penalises FDWs for late deliveries, and (2) FDWs' perception of safety on the road while delivering for the trial partner FDP during the trial period. Since these data were collected via a survey, the analysis was conducted at the *worker level*. The survey response rate was extremely low: the first question (beliefs about consequences) was answered by 559 FDWs (3.9% of the trial participants), and the second question (perception of safety) was answered by 549 FDWs (3.8% of the trial participants).

Figures 11 and 12 show the distribution of responses to each survey question. Responses were compared between three groups: (1) FDWs in the control group, (2) FDWs in the treatment group who opened at least one message, and (3) FDWs in the treatment group who did not open any messages. There were no significant differences in subjective beliefs about being penalised or perceptions of road safety between these three groups.

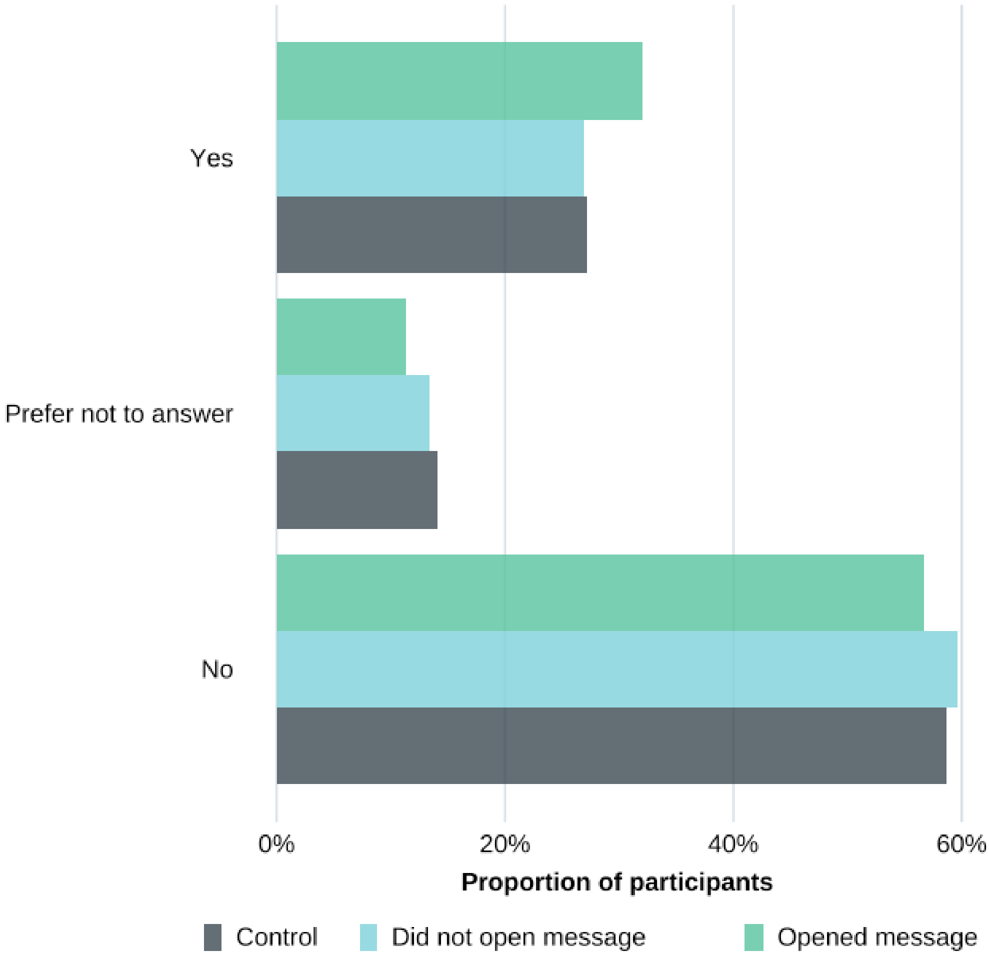


Figure 11: Distribution of responses to the survey question "Do you think that [name of FDP] will penalise you if you're a few minutes late for a delivery?"

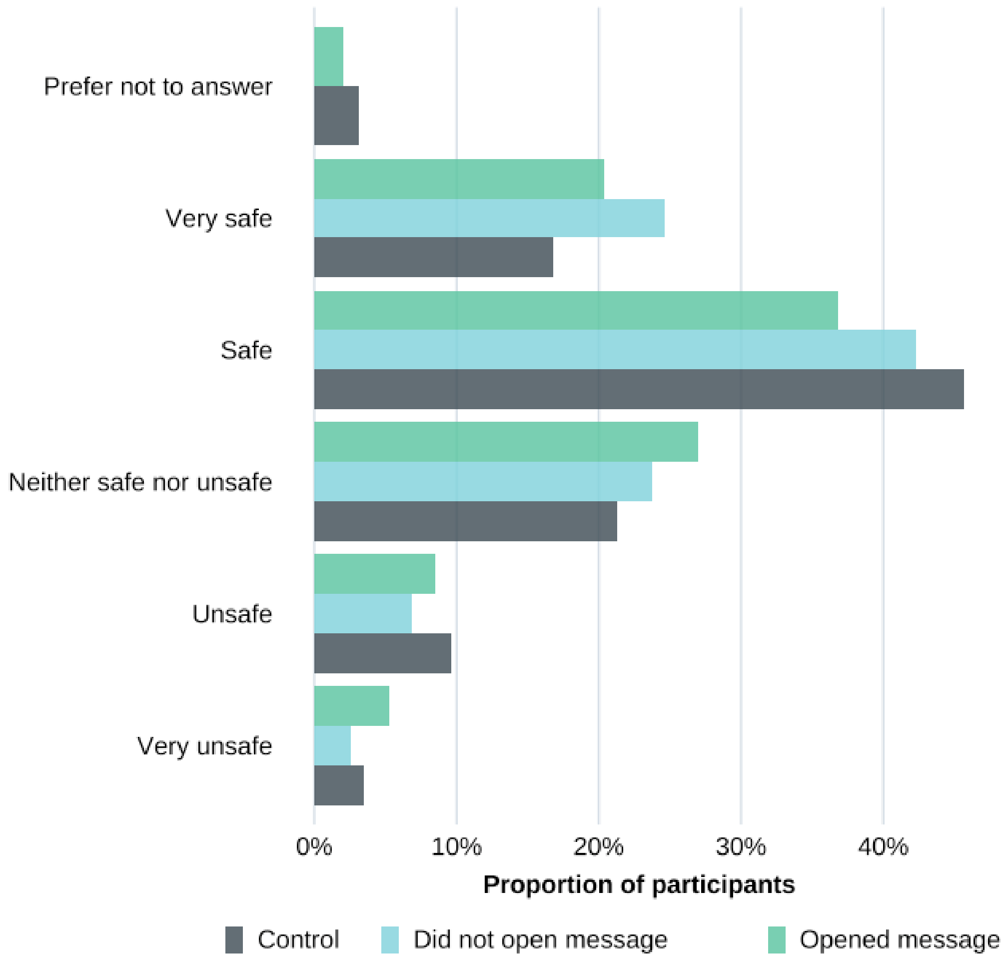


Figure 12: Distribution of responses to the survey question "How safe did you feel on the road when delivering for [name of FDP] in the past month?"

Vehicle type, experience, and customer rating had no effect on beliefs about being penalised. However, FDWs using cars were more likely than FDWs using bicycles or motorcycles to feel safe, and FDWs with more experience were more likely to feel unsafe. Customer rating had no effect on FDWs' perception of safety.

It is worth highlighting that the majority of survey respondents (regardless of whether they were in the treatment or control groups) reported a belief that the trial partner FDP would not penalise them for a late delivery (58% of respondents) and that they felt "safe" or "very safe" while delivering for the trial partner FDP (62% of respondents). However, it is equally important to note that there are still FDWs who believe they would be penalised for a late delivery (28% of respondents) and who feel "unsafe" or "very unsafe" on the road (12%). This held true even for a proportion of FDWs who engaged with the trial messages, which aimed to combat those beliefs. This finding highlights the influence of *confirmation bias*, or the tendency of people to favour information that is congruent with their existing beliefs and to discount information that is incongruent, particularly for beliefs that are deeply held (Plous, 1993). In the context of this trial, it is possible that simply receiving information about penalisation and safety was insufficient to shift beliefs about these topics for a minority of FDWs.

Exploratory outcomes

The exploratory analyses aimed to determine the effect of the messages on (1) the average time taken to make a delivery, controlling for vehicle type, and (2) the customer ratings received by FDWs during the trial period. All analyses were conducted at the worker level.

The messages had no effect on delivery times or customer satisfaction

Both ITT and ETT analyses showed no significant difference between the treatment and control groups on either of the exploratory outcomes. Figure 13 shows the effect of the messages on delivery time, broken down by vehicle type, for both the treatment and control groups.

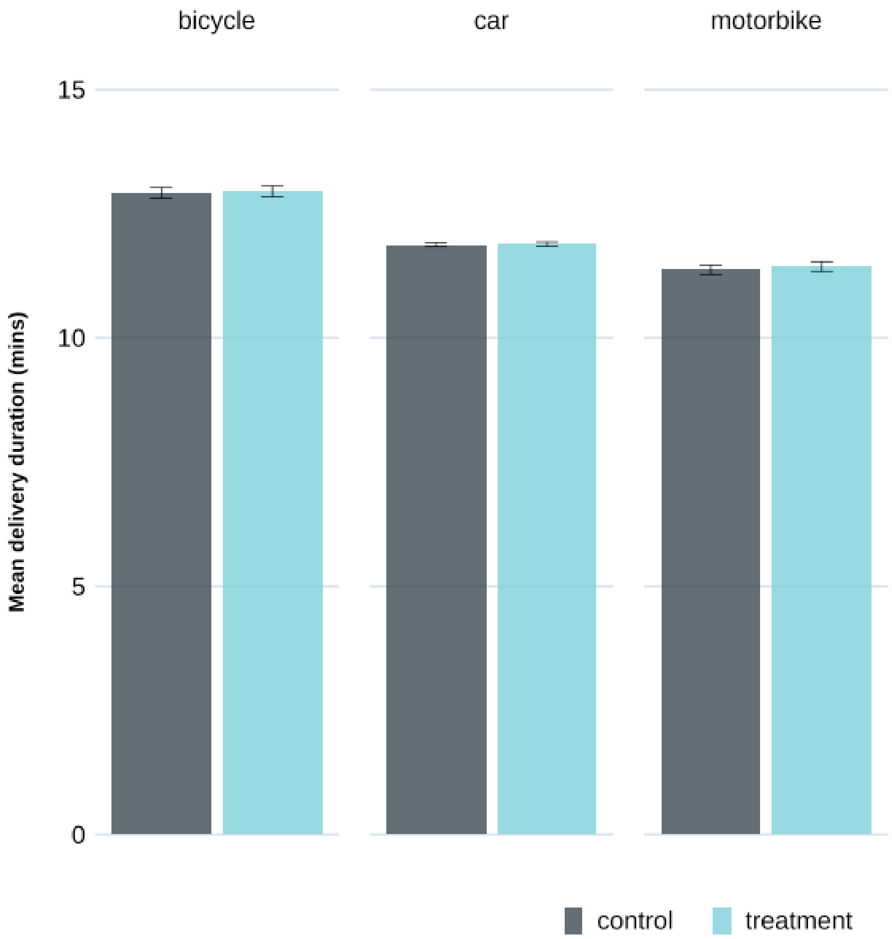


Figure 13: Effect of the messages on mean delivery duration by vehicle type in an ITT analysis. Error bars represent the standard error of the mean.

Figure 14 shows the effect of the messages on mean customer rating (a binary measure with “thumbs up” or “thumbs down” as the two rating options). The figure includes only those FDWs who received at least one rating during the trial period (n = 4,368 for the treatment group and n = 4,512 for the control group).

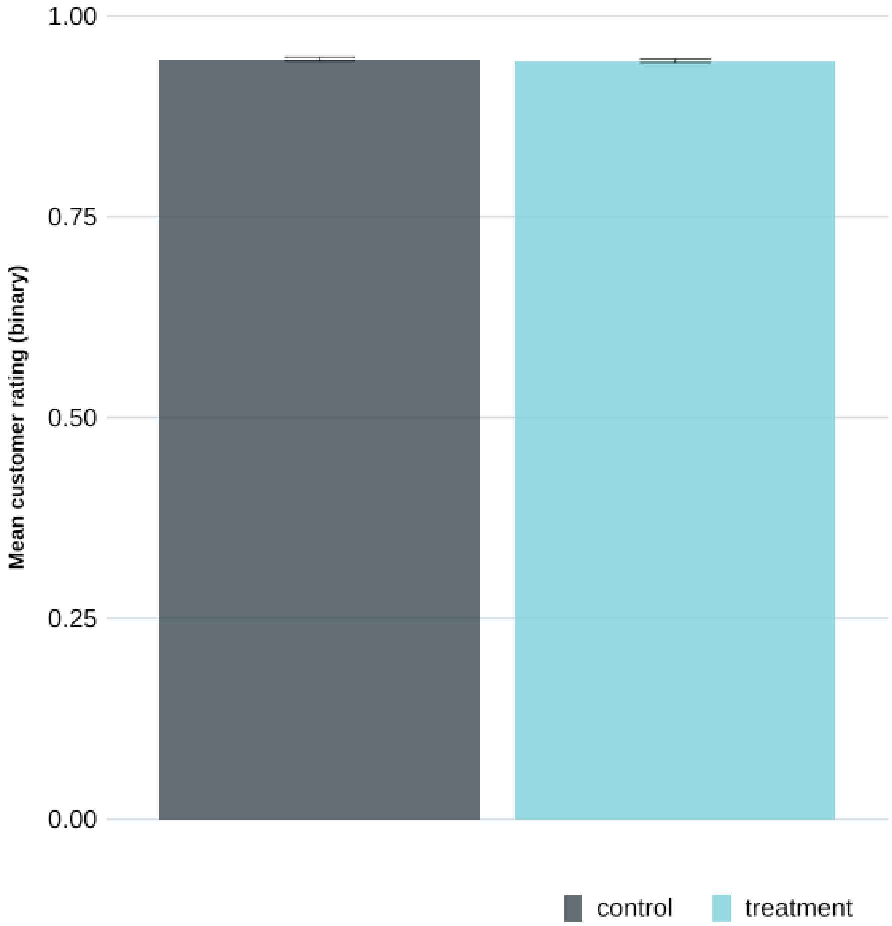


Figure 14: Effect of the messages on mean customer rating in an ITT analysis. Error bars represent the standard error of the mean.

Both exploratory findings highlight the fact that there were no adverse effects of the messages on two of the metrics commonly used by FDPs to track worker performance.

Conclusions

An iterative co-design process was undertaken with FDWs, representatives of four major FDPs, and an e-bike rental company to improve FDW road and traffic safety, the number one WHS priority among this group of stakeholders. To address this challenge, a suite of four behaviourally informed messages was developed that aimed to reframe FDWs' perceptions of delivery times as estimations rather than expectations, thus reducing the need for rushing and speeding to meet delivery deadlines. Rushing and speeding, both identified in Phases 1 and 2 of this project as common risky behaviours, can lead to both psychological (e.g. stress) and physical harms (e.g. traffic accidents).

The suite of four messages was evaluated in an RCT in partnership with one of the FDPs. In the trial, the FDP sent one message weekly to 7,066 of its Sydney- and Melbourne-based FDWs (with another 7,215 of its FDWs acting as the control group) via its in-app messaging function. Average travel speeds during deliveries were analysed and used as a proxy for testing the ability of the intervention to reduce WHS risk among FDWs. As described in this report, the messaging intervention was modestly effective in reducing average travel speeds, an important first step toward creating a positive impact in the WHS space. However, this effect was seen only among the minority of FDWs who opened and read the messages. While there was no significant effect of the messages on the fleet as a whole, there were also no adverse effects on key FDW performance metrics (overall delivery time and customer satisfaction).

The trial findings show promise in reducing WHS risks among FDWs through behaviourally informed messaging, but also highlight the greater, more pressing challenge of increasing FDW engagement. This is a known and persistent problem faced by FDPs when attempting to engage FDWs in voluntary initiatives, such as WHS training and worker surveys. As a result, we designed our trial to maximise FDW engagement by:

- Collaborating with FDWs, as well as FDPs and an e-bike rental company, to select a relevant problem to address with our intervention and trial
- Co-designing and iterating the messaging intervention with FDWs, as well as the other stakeholders
- Partnering with a major FDP as the trial partner to evaluate the messaging intervention with a large FDW population under real-world conditions
- Designing the trial such that the trial partner FDP sent the intervention messages directly to their FDWs via their in-app messaging system

Despite the use of these strategies to build FDW engagement, the rates at which FDWs in the trial opened the messages remained low. This outcome underscores the fact that engagement must be addressed before attempting further trials with this cohort.

Recommendations

In order to tackle the challenge of low FDW engagement, a range of behavioural insights principles could be leveraged to lift FDW engagement and create a path toward ultimately improving their WHS. Each of these principles is described below, accompanied by an example of how they could be implemented in the context of FDW messaging. As a whole, the principles and implementation recommendations share a common approach that is best illustrated with the following quote:

> *"Instead of asking, 'How can I get him or her to do it?', it starts with a question of 'Why isn't he or she doing it already?' Then… you ask, 'What can I do to make it easier for that person to change?'"*[4]

Messenger effects

People tend to give different weight to information depending on who is communicating it to them (Wilson & Sherrell, 1993). A review of almost 200 studies published from 1950-2000 found that highly credible messengers (defined as those who had a high level of relevant expertise and were perceived as trustworthy) were more effective than low-credibility messengers at changing attitudes and behaviours (Pornpitakpan, 2004).

How to implement: Future messaging interventions targeting WHS risk behaviours may be more effective if not addressed from the FDP via the in-app messaging function. Instead, communications to FDWs could be sent from a more credible and trusted source that is relevant to the focus of the messaging initiative. For example, when aiming to influence implementation intention, send the message from a real FDW who has recently avoided a traffic accident or near-miss through advance planning.

Reciprocity

Reciprocity is a strategy in which an unconditional gesture or gift is offered before asking the recipient to do something in return. For example, a handwritten note from the CEO of an investment bank, accompanied by a lolly, increased charitable donations among the bank's staff (Behavioural Insights Team, 2015). The "gift" may also be intangible in the form of information that increases operational transparency. For example, hospital outpatients who were told how much missed appointments cost the healthcare system were significantly more likely to attend their scheduled appointments (NSW Behavioural Insights Unit, 2016).

How to implement: Give FDWs information that helps them understand how the platform works. For example, in addition to anchoring, the FDP could tell FDWs about 'safety time', highlighting

[4] Daniel Kahnemann, *How to Launch a Behavior-Change Revolution.* Freakonomics podcast, Ep. 306, 2017.

its purpose and how much the FDP is investing in supporting additional delivery time for FDWs to ensure their safety.

Incentives

Providing incentives can increase motivation and spur behaviour change. Incentives do not necessarily need to be financial. In fact, there is a robust body of research demonstrating that non-monetary incentives, including those that are tangible, intangible, or even symbolic, are effective in changing behaviour. For example, one study found that a symbolic reward - a congratulatory card highlighting top performance on a data entry task - increased performance by an average of 12% (Kosfeld & Neckermann, 2011).

How to implement: Offer appealing incentives to FDWs to engage with the messages and/or complete the survey. These may be tangible incentives (e.g. vouchers for purchasing safety equipment or for a bicycle service) or intangible incentives (e.g. add game elements to the FDP app such that opening messages and participating in surveys are fun and rewarding).

Dynamic social norms

Apply dynamic instead of static social norms. Dynamic social norms differ from static social norms in that they communicate the extent to which people are adopting a new behaviour over time (Latané, 2000). “More and more people are getting vaccinated against COVID-19” is an example of a dynamic social norm (cf. the static social norm of “XX% of people have been vaccinated against COVID-19”). Since social norm messaging should highlight the behaviour you want people to adopt, rather than the behaviour you want them to discontinue, dynamic social norms are particularly effective when the desired behaviour is currently performed only by a minority of the group.

How to implement: Instead of using statements such as “The majority of our riders prioritise safety over speed”, consider highlighting the increasing number of FDWs who are doing the right thing. For example, “More and more of our riders take time to make sure they ride safely when on the road.”

Acknowledgements

This work was funded by the NSW Government’s Centre for Work Health and Safety, who also oversaw the work, reviewed the manuscript and approved its publication. We would also like to acknowledge the contributions of the following:

The FDW and FDP representatives who participated in Phase 3

The FDP that acted as the trial implementation partner in Phase 4

The FDWs who participated in the trial

Dr Karen Tindall and Dr Alex Gyani (The Behavioural Insights Team) for ongoing policy and research advice

Travis McMullen (Centre for WHS) for ongoing project support

References

Amin, M. (2020, 4 April). *Food delivery riders say employers are not doing enough to protect them from coronavirus*. Australian Broadcasting Corporation. Retrieved from https://www.abc.net.au/news/2020-04-04/food-delivery-riders-not-being-supported-incoronavirus-response/12120480

Apouey, B., Roulet, A., Solal, I., & Stabile, M. (2020). Gig workers during the COVID-19 crisis in France: financial precarity and mental wellbeing. *Journal of Urban Health, 97*(6), 776-795

Bajwa, U., Gastaldo, D., Di Ruggiero, E., & Knorr, L. (2018). The health of workers in the global gig economy. *Globalisation and Health, 14*, 124.

Bélanger-Gravel, A., Godin, G., & Amireault, S. (2013). A meta-analytic review of the effect of implementation intentions on physical activity. *Health Psychology Review, 7*(1), 23-54.

Behavioural Insights Team (2014). *EAST: Four simple ways to apply behavioural insights*.

Behavioural Insights Team (2015). *Update report 2013-2015*.

Behavioural Insights Team (2019). *EAST for Health and Safety*.

Bright, S. & Fitzgerald, A. (2019). *Snapshot: On Demand Food Delivery Drivers*. The Transport Workers Union. Retrieved from https://d3n8a8pro7vhmx.cloudfront.net/victorianunions/pages/1411/attachments/original/1517351126/ODFDR_Snapshot.pdf?1517351126

Brown, T. J., Hardeman, W., Bauld, L., Holland, R., Maskrey, V., Naughton, F., ... & Notley, C. (2019). A systematic review of behaviour change techniques within interventions to prevent return to smoking postpartum. *Addictive Behaviors, 92*, 236-243.

Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science, 12*, 105-109.

Cialdini, R. B. (2007). *Influence: The psychology of persuasion*. New York: HarperBusiness, Revised Edition.

Convery, E., Morse, A., Fung, B., Wodak, S., Powell, Z., Quinn, V., Taylor, M., Searle, B., & Vårhammar, A. (2020a). *Work health and safety of FDWs in the gig economy*. Report prepared for the Centre for Work Health and Safety (NSW). Retrieved from http://www.centreforwhs.nsw.gov.au/knowledge-hub/work-health-and-safety-of-food-delivery-workers-in-the-gig-economy-report

Convery, E., Howard, E., Powell, Z., Wodak, S., Fung, B., Quinn, V., Taylor, M., Searle, B., & Vårhammar, A. (2020b). *Work health and safety perceptions of food delivery platforms in the gig economy*. Report prepared for the Centre for Work Health and Safety (NSW). Retrieved from http://www.centreforwhs.nsw.gov.au/knowledge-hub/work-health-and-safety-perceptions-of-food-delivery-platforms-in-the-gig-economy

Christie, N., & Ward, H. (2019). The health and safety risks for people who drive for work in the gig economy. *Journal of Transport and Health, 13*, 115-127.

De Stefano, V. (2015). The rise of the just-in-time workforce: On-demand work, crowdwork, and labor protection in the gig-economy. *Comparative Labor Law & Policy Journal, 37*, 471.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology, 51*, 629-636.

Geraerts, E., Bernstein, D. M., Merckelbach, H., Linders, C., Raymaekers, L., & Loftus, E. F. (2008). Lasting false beliefs and their behavioral consequences. *Psychological Science, 19*(8), 749-753.

Geissinger, A., Laurell, C., & Sandström, C. (2020). Digital disruption beyond Uber and Airbnb - tracking the long tail of the sharing economy. *Technological Forecasting and Social Change*, 155.

Geneletti, S., & Dawid, A. P. (2011). Defining and identifying the effect of treatment on the treated. *Causality in the Sciences*. 10.1093/acprof:oso/9780199574131.003.0034.

Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: a meta-analysis of the attitude-behavior relation. *Psychological Bulletin, 132*(5), 778.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research, 35*(3), 472-482.

Gollwitzer, P. M. (1999). Implementation intentions: strong effects of simple plans. *American Psychologist, 54*(7), 493.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology, 38*, 69-119.

Goods, C., Veen, A., & Barratt, T. (2019). "Is your gig any good?" Analysing job quality in the Australian platform-based food-delivery sector. *Journal of Industrial Relations, 61*(4), 502-527

Jeffrey, J., Whelan, J., Pirouz, D. M., & Snowdon, A. W. (2016). Boosting safety behaviour: Descriptive norms encourage child booster seat usage amongst low involvement parents. *Accident Analysis & Prevention, 92*, 184-188.

Johnstone, R. (2019). Regulating work health and safety in multilateral business arrangements. *Australian Journal of Labour Law, 32*(1), 41-61.

Keoghan, S. (2020, 15 April). *'They cross in front of the car': Surge in Sydney food couriers raises safety concerns*. Sydney Morning Herald. Retrieved from https://www.smh.com.au/national/nsw/they-cross-in-front-of-the-car-surge-in-sydney-foodcouriers-raises-safety-concerns-20200415-p54jza.html

Koranyi, I., Jonsson, J., Rönnblad, T., Stockfelt, L., & Bodin, T. (2018). Precarious employment and occupational accidents and injuries - a systematic review. *Scandinavian Journal of Work, Environment & Health, 44*(4), 341-350.

Latané, B. (2000). Pressures to uniformity and the evolution of cultural norms: Modeling dynamic social impact. In D. R. Ilgen & C. L. Hulin (Eds.), *Computational modeling of behavior in organisations: The third scientific discipline* (p. 189–220). American Psychological Association.

McCoy C. E. (2017). Understanding the intention-to-treat principle in randomised controlled trials. *The Western Journal of Emergency Medicine, 18*(6), 1075-1078.

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., ... & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine, 46*(1), 81-95.

Michie, S., Carey, R. N., Johnston, M., Rothman, A. J., De Bruin, M., Kelly, M. P., & Connell, L. E. (2018). From theory-inspired to theory-based interventions: a protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action. *Annals of Behavioral Medicine, 52*(6), 501-512.

Newlands, G. (2020). Algorithmic Surveillance in the Gig Economy: The Organization of Work through Lefebvrian Conceived Space. *Organisation Studies*. doi:10.1177/0170840620937900

National Health and Medical Research Council. (2007, updated 2018). *The National Statement on Ethical Conduct in Human Research*. Canberra: NHMRC.

NSW Behavioural Insights Unit. (2016). *Spotlight on health results: behavioural insights short report*.

Orbell, S., Hodgkins, S., & Sheeran, P. (1997). Implementation intentions and the Theory of Planned Behavior. *Personality and Social Psychology Bulletin, 23*, 945-954.

Ortiz-Prado, E., Henriquez-Trujillo, A. R., Rivera-Olivero, I. A., Lozada, T., Garcia-Bereguiain, M. A., & team, U.-C.-. (2021). High prevalence of SARS-CoV-2 infection among food delivery riders. A case study from Quito, Ecuador. *Science of the Total Environment*, 770.

Perkins, H. W., Linkenbach, J. W., Lewis, M. A., & Neighbors, C. (2010). Effectiveness of social norms media marketing in reducing drinking and driving: A statewide campaign. *Addictive Behaviors, 35*(10), 866-874.

Plöger, J., & Keuneke, F. (2021). Arbeit per App – neue Abhängigkeiten in der Gig Economy. Standort. doi:10.1007/s00548-021-00695-w

Plous, S. (1993). *The psychology of judgment and decision making*. New York: McGraw-Hill.

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*(2), 243-281.

SafeWork NSW, 2020, unpublished data.

Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., & Griskevicius, V. (2007). The constructive, deconstructive, and reconstructive power of social norms. *Psychological Science, 18*(5), 429-434.

Sheeran, P., & Webb, T. L. (2016). The Intention–Behavior gap. *Social and Personality Psychology Compass, 10*(9), 503-518. doi:10.1111/spc3.12265

Tajfel, H. (Ed.). (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Academic Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science, 185*(4157), 1124-1131.

Van Boven, L., Loewenstein, G., Dunning, D., & Nordgren, L. F. (2013). Changing places: A dual judgment model of empathy gaps in emotional perspective taking. In *Advances in experimental social psychology* (Vol. 48, pp. 117-171). Academic Press.

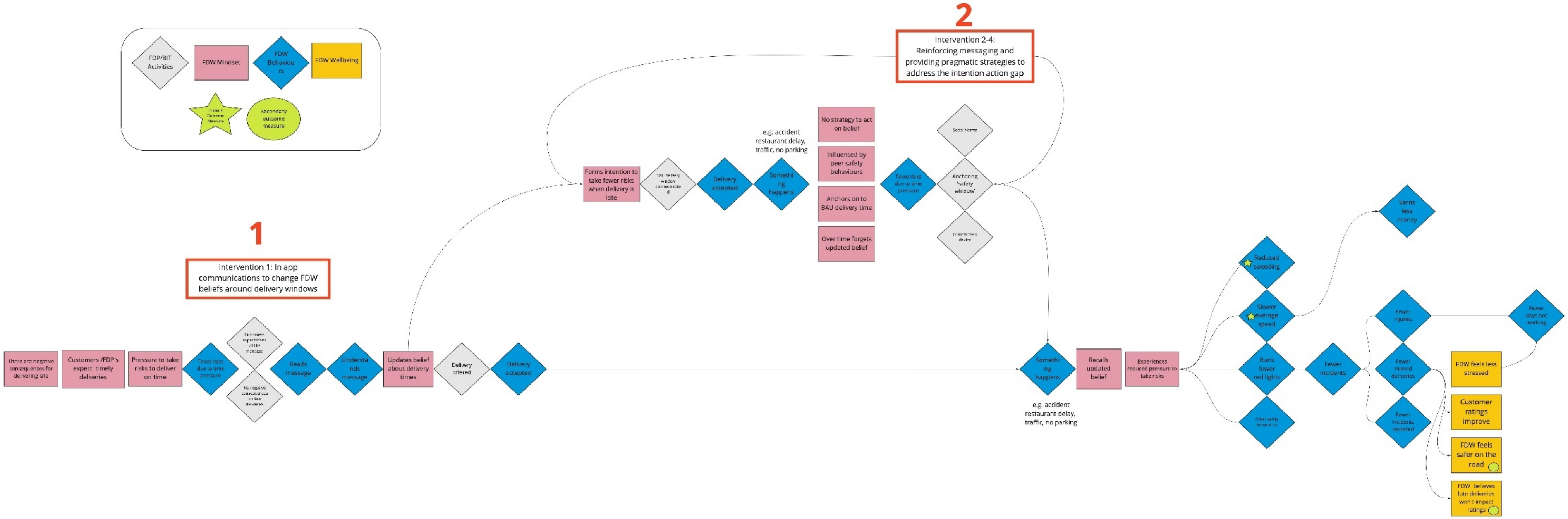
Veen, A., Barratt, T., & Goods, C. (2020). Platform-capital's 'app-etite' for control: a labour process analysis of food-delivery work in Australia. *Work, Employment and Society, 34*(3), 388-406.

Wilson, E.J., Sherrell, D.L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science, 21*, 101.

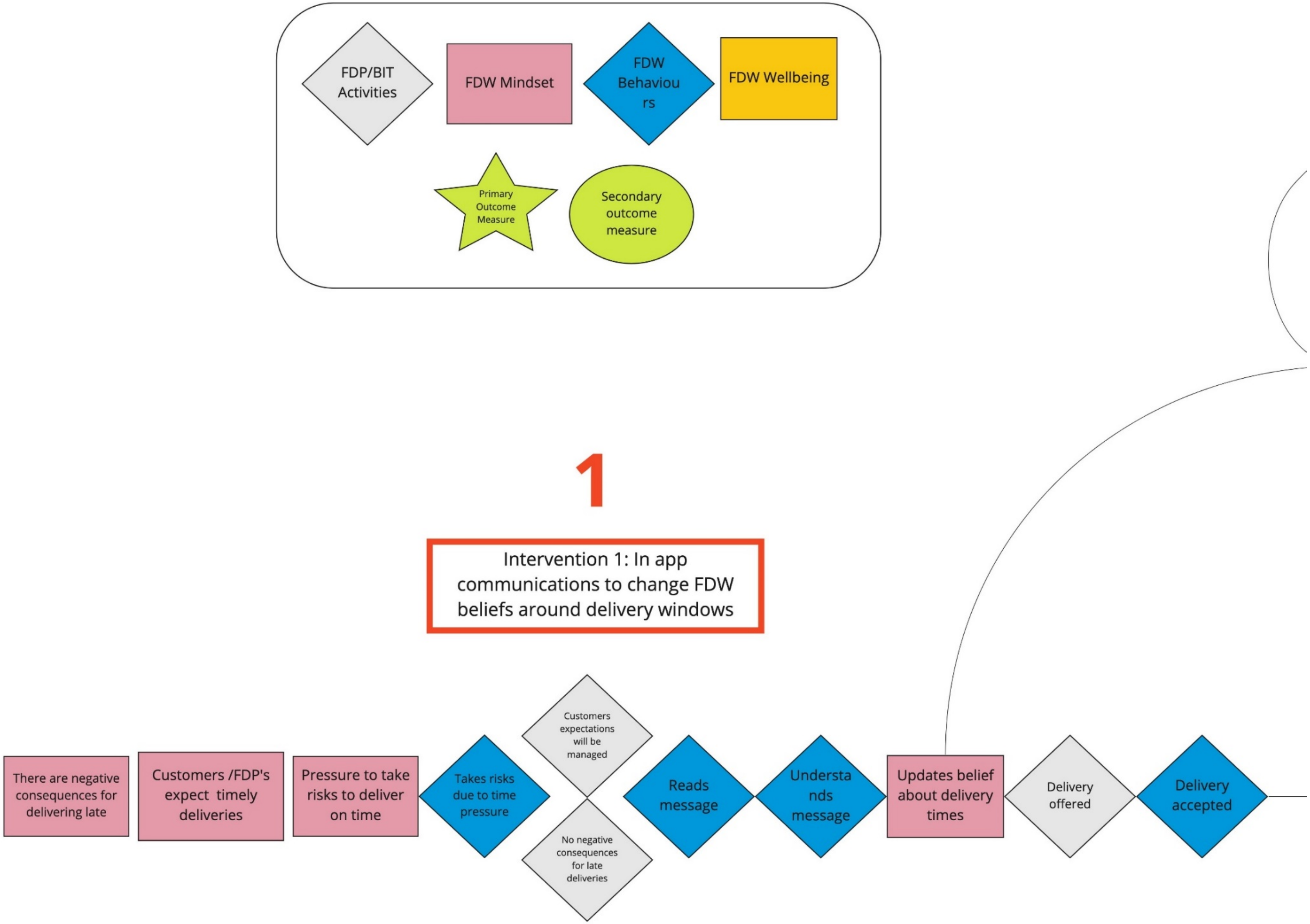
Appendices

A: Theory of Change

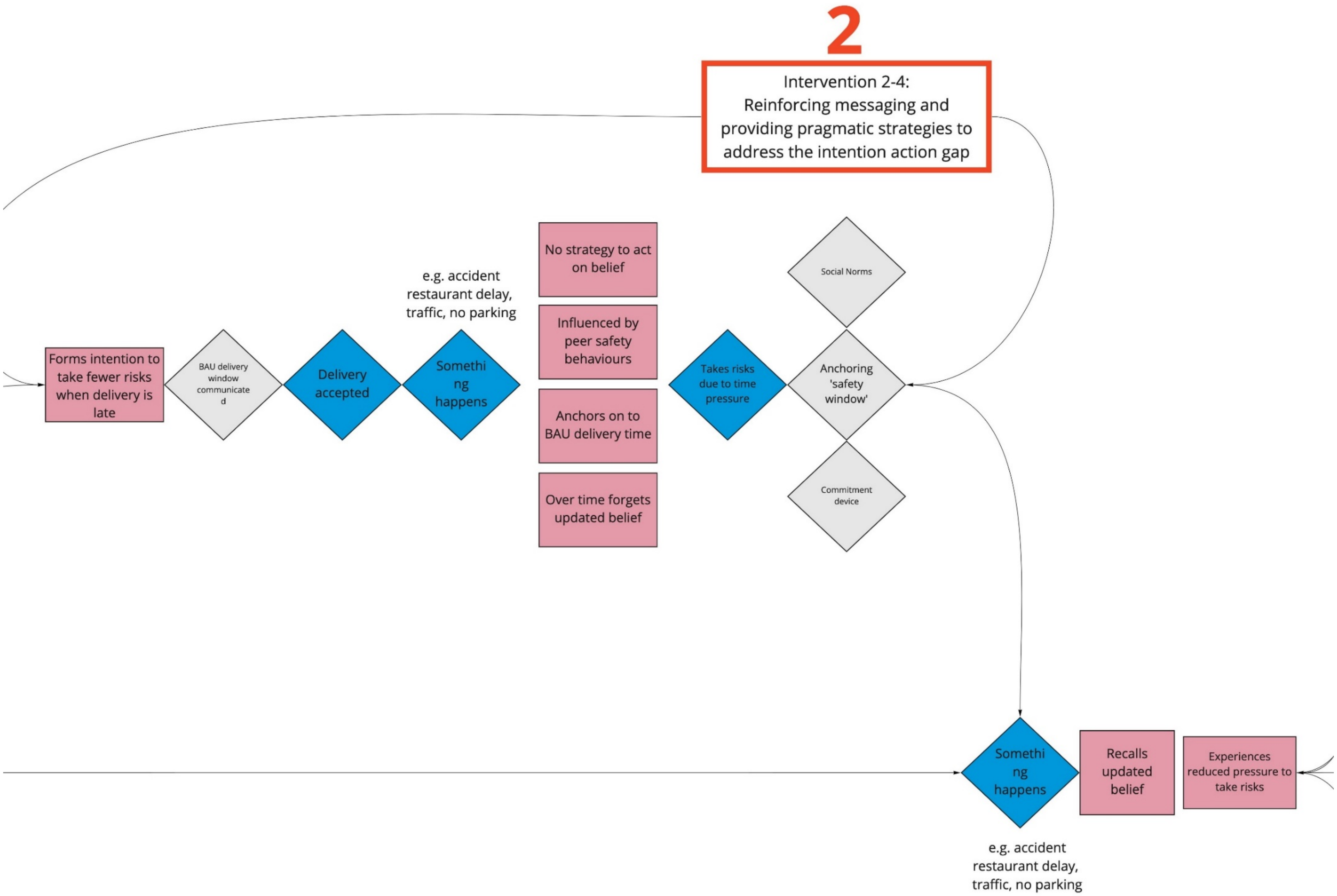
The theory of change for this project specifies key opportunities for in-app messaging interventions (the red boxes labelled 1 and 2), linking behavioural mechanisms of action (grey diamonds) to FDW attitudes, beliefs, and behaviours targeted for change (red and blue diamonds). The yellow boxes specify where along the causal pathway we measured behaviour change.



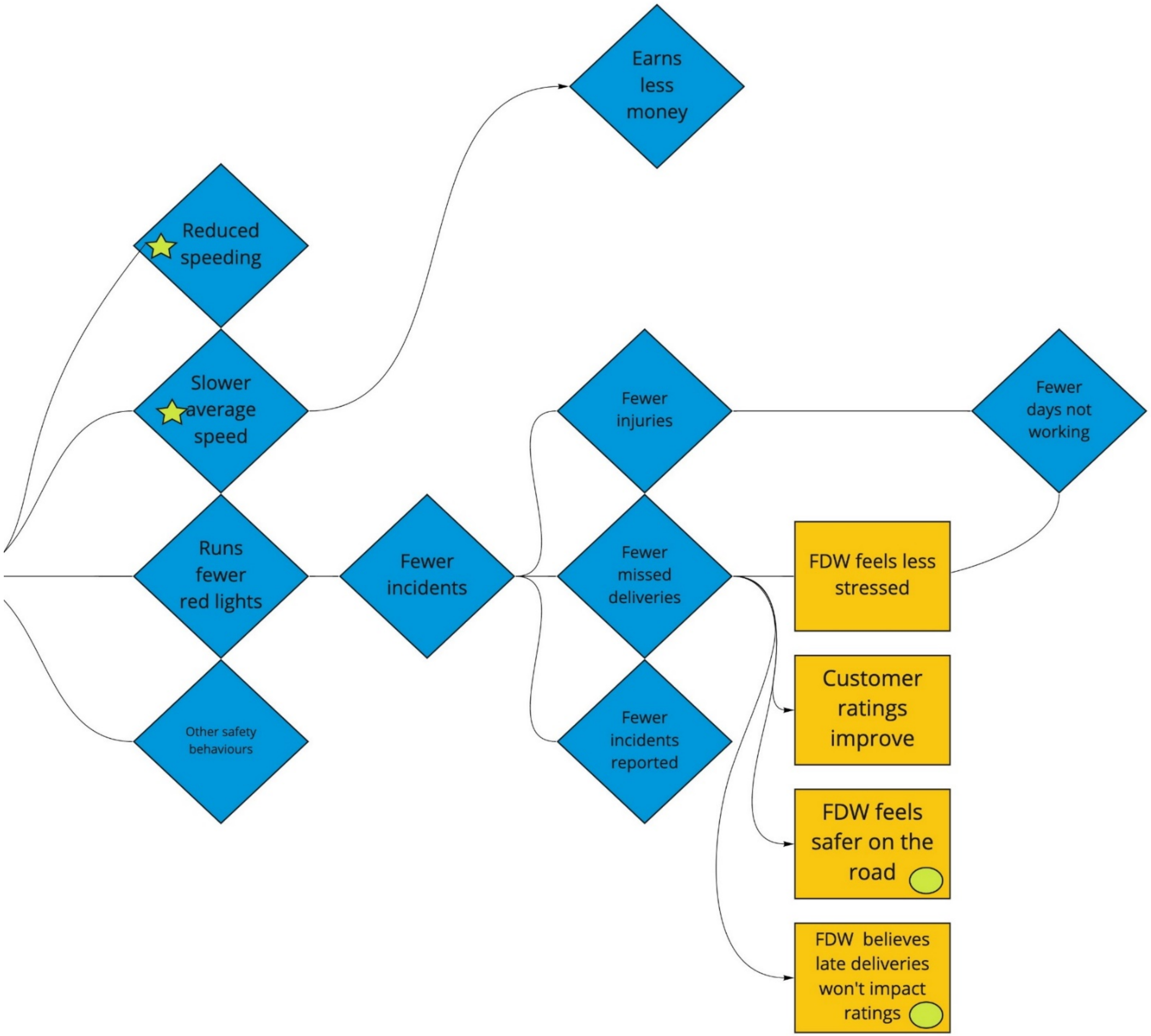
The first part of the theory of change, zoomed in to show detail.



The second part of the theory of change, zoomed in to show detail.



The third part of the theory of change, zoomed in to show detail.



B: Power calculations

The minimum detectable effect size (MDES) for primary outcome measure of the trial is shown below, assuming a standard deviation at 25% of the total travel time/speed, an alpha of .05, and 80% power. The calculations indicated that to meet these requirements, a sample size of 5,000 FDWs would be necessary.