Financial Conduct Authority

Occasional Paper 47

December 2018

Blackbird's alarm call or nightingale's lullaby? The effect of tweet risk warnings on attractiveness, search, and understanding

Timothy L. Mullett, Laura Smart, Neil Stewart

FCA Occasional Papers in financial regulation

The FCA Occasional Papers

The FCA is committed to encouraging debate on all aspects of financial regulation and to creating rigorous evidence to support its decision-making. To facilitate this, we publish a series of Occasional Papers, extending across economics and other disciplines.

The main factor in accepting papers is that they should make substantial contributions to knowledge and understanding of financial regulation. If you want to contribute to this series or comment on these papers, please contact Karen Croxson at karen.croxson@fca.org.uk.

Disclaimer

Occasional Papers contribute to the work of the FCA by providing rigorous research results and stimulating debate. While they may not necessarily represent the position of the FCA, they are one source of evidence that the FCA may use while discharging its functions and to inform its views. The FCA endeavours to ensure that research outputs are correct, through checks including independent referee reports, but the nature of such research and choice of research methods is a matter for the authors using their expert judgement. To the extent that Occasional Papers contain any errors or omissions, they should be attributed to the individual authors, rather than to the FCA.

Authors

Timothy L. Mullett (University of Bath and Warwick Business School)

Laura Smart (Financial Conduct Authority)

Neil Stewart (Warwick Business School)

Acknowledgements

We would like to thank Josie Flight for organising experiments 1A-C on behalf of the FCA, Paul Adams, Dr Pete Lunn, Professor John Gathergood, Professor Colin Camerer, Dr Stefan Hunt and Dr Jeroen Nieboer for expert guidance and review, Tom Robertson for helping to design the stimuli, as well as Kate Studholme, Tom Carter, Claudine Hilton and Sue Newman for providing specialist input on financial promotions.

This work was supported by grant ES/N018192/1 from the UK Economic and Social Research Council. We are grateful for this support.

All our publications are available to download from www.fca.org.uk. If you would like to receive this paper in an alternative format, please call 020 7066 7948 or email publications_graphics@fca.org.uk or write to Editorial and Digital Department, Financial Conduct Authority, 12 Endeavour Square, London E20 1JN.

Contents

Abstract

Modern advertising platforms, such as those on social media, often place restrictions on the size of adverts and the length of messages. These restrictions can pose difficulties for financial products which are required to be advertised in a "standalone compliant" way. Standalone compliance means that the advert must be balanced, for example including risk warnings, or information about unusual features, alongside positive information about the product. We ran a series of behavioural experiments in simulated social media environments which examine the design and timing of risk warnings and other balancing information. The results show that, for character-limited social media, standalone compliance reduces consumers' information search and understanding of risks, ultimately leading to them choosing less suitable products.

Unsurprisingly, products advertised through standalone compliant tweets are less attractive to potential customers. They are then less likely to be clicked on and have their webpages explored relative to non-compliant tweets. Even when viewing a social media feed where all tweets are compliant, participants are less likely to shop around compared with a feed where all tweets are non-compliant.

In the final two experiments, participants were given a scenario in which they needed to meet specific needs and levels of risk, while avoiding certain outcomes. Only one product type from a range of options was suitable for the consumer in each scenario, because of product features which were stated in the risk warning. In these experiments, standalone compliance reduced the likelihood that participants chose the most suitable product by 7-14%. This is best explained by reduced search activity; because standalone compliant products were less likely to be explored, participants had a reduced understanding of the products and were less able to make a good decision.

The results also demonstrate an important role for well-placed risk warnings in educating consumers. More detailed risk warnings on product webpages were found to be most effective at increasing understanding, and they were significantly more effective when designed using insights from behavioural science.

1 Executive summary

Background

In 2017, 39.4 million (58%) of the UK population used social media - and we are increasingly using it to buy products.[1] In Europe, an estimated 35% of consumers used social media for product research.[2] Given its popularity, it is no surprise that financial services firms have enthusiastically taken it up. Social media spending as a percentage of total marketing budgets across all sectors globally has grown between 2009 and 2016 from an average of 3.5% to 11.7%.[3]

Social media can be a valuable way for firms to reach new customers and advertise products which might enrich their lives. It also allows firms to compete with each other to offer better deals. However, along with opportunities, the nature of relatively new channels of advertising presents risks. For example, some social media channels, like Twitter, limit the number of characters per message, making it more difficult to impart crucial product information. There is also emerging evidence that people behave differently online, making faster decisions based on less information (Benartzi & Lehrer, 2015) – and these might not always be good ones.

The growth of social media and online advertising has led many regulators to revisit advertising rules, many of which were written for a world where the majority of adverts appeared in print. For example, the UK Advertising Standards Authority extended the Code of Non-Broadcast Advertising and Direct & Promotional Marketing (CAP) to include content on a company's own social media channel in 2011.[4]

Similarly, the Financial Conduct Authority (FCA) published guidance on social media advertising in 2015, which reminded firms that any form of communication (including character-restricted social media adverts) are capable of being a financial promotion, and are therefore required to contain certain information, including any necessary risk warnings (Financial Conduct Authority, 2015).[5] The FCA considers that such **standalone compliance (SC)** of adverts is important, creating a consistent approach for firms offering different product types and ensuring consumers receive a balanced explanation of a product.

There is evidence from behavioural theory that standalone compliance is necessary to prevent harm. People tend to make judgements based on initial information and then find it difficult to update them when more information becomes available (Kahneman & Tversky, 1982; Gigerenzer & Goldstein, 1996; Reyna & Brainerd, 1995; Edwards & Potter, 1992). In advertising, when presented with more (often negative) information about a product later in the purchase process, consumers may not change their mind.

[1] https://www.statista.com/statistics/278341/number-of-social-network-users-in-selected-countries/

[2] GlobalWebIndex. Europe: Region Report. 2017.

[3] http://deloitte.wsj.com/cmo/2016/11/15/cmo-survey-why-is-social-media-falling-short/

[4] https://www.asa.org.uk/resource/Extending-the-digital-remit-of-the-CAP-Code.html#.VX7KzaROVi4

[5] Link to the paper here. Risk warnings may also be provided through images, where characters are not limited.

This raises the risk that they may buy a product on the basis of incomplete information contained in an advert.

While this theory has been thoroughly tested in the laboratory, little is known about this effect in practice for financial adverts in character-limited social media. Despite this, social media is increasingly prominent in firms' marketing strategies and many argue that it presents opportunities for more innovative ways of interacting with customers online.

Given the increasing importance of social media, it is important that the FCA understands the impact of its rules in this area. We tested several design features of character-limited risk warnings, as well as how standalone compliance affects consumers' judgements across 10 products.

Method

We carried out 10 experiments ('artefactual field experiments' in the taxonomy of Harrison and List (2004)) designed to simulate an increasingly realistic choice environment. We slowly built evidence at each stage, designing each subsequent experiment to answer the questions raised in the last one, gradually adding complexity to simulate real-life decision making.

The first set of experiments (1A-C) investigated the impact of specific features of social media posts and balancing information, including risk warnings, on consumer understanding and product preferences for 10 different product types. Features we tested included the effect of improved readability, pictures, and an FCA label.

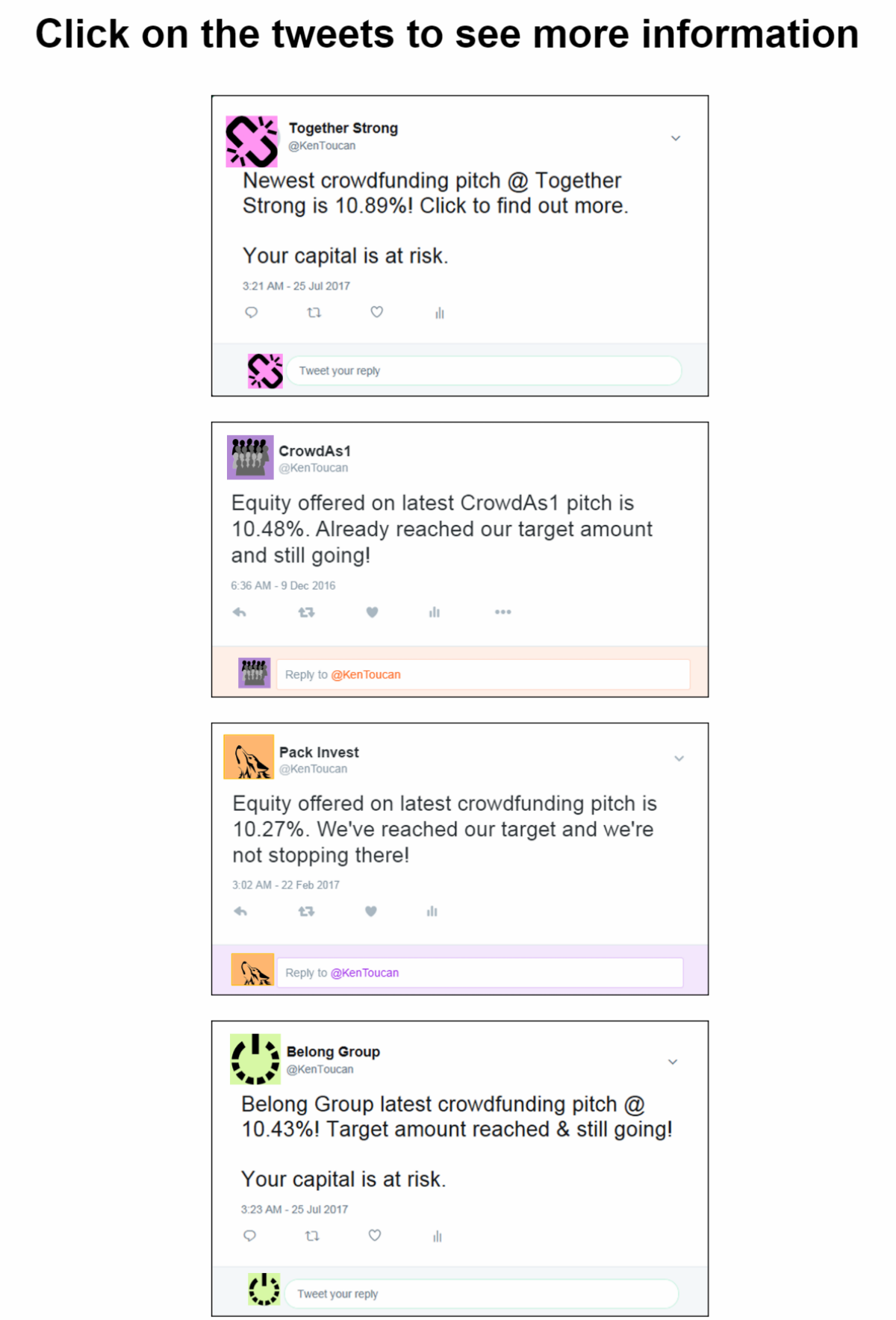
The second set of experiments (2A-F) investigated the impact of standalone compliance on understanding, preferences and search activity. In the final experiments (2E and 2F), we create a hypothetical scenario with only one objectively suitable product choice and test the impact of standalone compliance on choosing the most suitable product, a proxy for understanding.

In Experiments 1A-C and 2A, participants rated products based upon individual tweets followed by their associated webpages (see Figure 1.1).

Figure 1.1: Example of a tweet (including risk warning) and associated webpage

In Experiments 2B-E, participants saw a simulated social media feed showing a number of tweets (see Figure 1.2). Participants could click through to each product's associated webpage and then either choose that product or click back to the social media feed. To investigate standalone compliance, we compared tweets which contained balancing information with the same tweets which had this information removed ("non-compliant tweets"). Balancing information included mandated risk warnings (e.g. "Your capital is at risk") or other risk information, (for example, eligibility for compensation schemes). For shorthand, this information is described in this paper as "risk warnings".

Figure 1.2: Example of a social media feed used in Experiments 2B-E



To test participants' **preferences**, we measure their ratings of different product tweets and webpages, and which product they ultimately chose. To investigate **search activity**, we measure clicks on tweets and the time spent on their associated webpages. For **understanding**, we use a set of comprehension questions and measure participants' improvement in accuracy after exposure to the product tweets and webpages. We also measure whether participants choose the most suitable product from a set of options given a specific scenario in the final two experiments (2E and 2F).

Results

Design of risk warnings: the impact of readability, pictures and labels

Experiments 1A-C show us that modifying risk warnings has a modest effect on individuals' understanding of the associated risks and their preferences for and against products.

- Simplified (more readable) risk warnings improved understanding, but only if they appeared on product webpages (not on social media posts).
- The inclusion of pictures linked to the tweet risk warning increased preference for the product, but did not affect understanding.
- Labelling tweet risk warnings with 'FCA warning' decreased understanding and did not affect preferences.

Timing of risk warnings: the impact of standalone compliance

Experiments 2A-F show that standalone compliance in a character-restricted medium reduces consumers' preferences for the product, and reduces search activity (e.g. clicking on adverts to get more information or investigating alternatives). Standalone compliance also has a negative effect on the understanding of risks, leading consumers to select less suitable products for a given scenario. In particular, in Experiments 2E and 2F, participants were 7-14% less likely to choose the most suitable product for a given scenario when adverts were standalone compliant. This was true when all possible adverts were standalone compliant and when only some were.

Table 1.1 summarises the results of each experiment on each outcome measure used.

Table 1.1: Summary of results

What was tested	#	Preferences			Search		Understanding	
		Tweet ratings	Product ratings	Choice	Clicks	Webpage time	Comprehension questions	Suitable product chosen
Readability	1A:	◯	◯				↑ / ◯	
Pictures	1B:	↑	↑				◯	
FCA label	1C:	◯	◯				↓	
Standalone Compliance	2A:	↓	↓					
	2B:			◯	↓	◯		
	2C:			↓	↓	↓		
	2D:			↓	↓	↓ / ◯	◯	
	2E:				↓	↓ / ◯		↓
	2F:				↓	↓ / ◯		↓

Discussion

Changing the design of risk warnings resulted in changes to preferences and understanding. Risk warnings displayed on character-restricted social media posts reduced preference and information search and led to consumers choosing less suitable products.

Design

The design of risk warnings can affect consumer understanding and preferences. For example, improving the readability of webpages, but not tweets, has a modest effect on comprehension. This means that it may be worth making efforts to simplify risk information, but that careful attention needs to be paid to the location of this information. We found that pictures increase preference for products even if they are displayed within the risk warning. This suggests that pictures may not be the most effective way to encourage consumers to take heed of risks. Our finding that an FCA label on risk warnings actually decreases understanding of the product reinforces the FCA's and other regulators' practices of not requiring this.

Timing

In the experiments, standalone compliance (SC) in social media posts reduced preference and search activity and resulted in lower understanding and less suitable product choices.

Of course, if the advertised product is risky, unsuitable or poor value, reducing preference (and subsequently purchase) would be a good one for consumers. To address this, our final two experiments measured the quality and suitability of consumers' choices, a proxy for outcomes. We found that when consumers were given scenarios for which a particular advertised product in a list was designed to be the only one suitable, standalone compliance meant that they were less likely to choose the most suitable one.

One explanation is that the presence of risk warnings in social media posts makes products less appealing to participants, meaning they are less likely to explore a product and consider fewer alternatives. Because they click on these adverts less, they are less likely to read the webpage text. In turn, this makes them less likely to get the information necessary to understand the product, buy it or choose a suitable alternative.

An alternative explanation for this effect is that risk warnings give the impression that all products are similar. As broadly similar warnings are applied within a category of products, brief early warnings within a social media post may not distinguish well between higher and lower risk alternatives. If these similar risk warnings are interpreted as a signal that all products are similar to one another, they may cause consumers to reduce information search and comparison and make them more likely to leave with nothing.

Since compliant adverts are less attractive than non-compliant ones, requiring that all, rather than just some, adverts are compliant prevents some firms from having an unfair advantage. However, in a compliant social media environment, participants were still less likely to search for more information and ultimately made less suitable product choices, than one in which all tweets were non-compliant. This could have important implications for competition, as well as consumers – less shopping around may decrease pressure on providers to improve offerings.

The experiments indicate that the webpage which is displayed when participants click on the tweet and on which the product can be purchased plays an important part in consumers' understanding of the risks and likelihood of making a good decision. It is therefore important that risks are presented in a way that attracts attention and improves understanding. Our experiments used a simplified webpage (no images or significant colour contrast, short, bullet-pointed information), which may not accurately reflect the variety of webpages in the real world. Monitoring standards and carrying out research, such as eye tracking, in this area could be an important role for the regulator.

2 Background

From pigeon, to pen, to tweet: regulating today

Adverts have nearly always faced restrictions on length, due to the cost of advertising space, the size of billboards and the relative success of short messages in grabbing attention. This has never been more the case with the increasing use of social media. Platforms like Twitter, Facebook, Instagram and LinkedIn commonly limit the number of characters in messages, meaning that marketers need to be ever more sparing with words.[6] Despite this, social media has become increasingly popular.

Some commentators argue that new ways of communicating with customers present valuable educational opportunities, which could help people to receive and remember relevant information (see consultation responses in Financial Conduct Authority, 2016). For example, social media adverts can direct consumers to explanatory videos or webchat services more easily than traditional adverts.

Character constraints mean that marketers have to work harder to get crucial product information across. This includes mandatory risk warnings, which aim to balance headline claims with information about exclusions or potential losses, as well as other balancing information, such as eligibility for compensation or information on the cost of borrowing. In this paper, "risk warning" may be understood to include all types of balancing information, whether this would formally be described as a risk warning or not. Risk warnings form part of the FCA's policy of standalone compliance (SC) by ensuring a balanced product description in every promotion - which is intended to protect consumers from misleading adverts.

On the other side, fair and effective advertising, including through social media, can facilitate competition between firms, thus improving what consumers are offered. In this case, how should regulators balance the consumer protection goal of standalone compliance with that of supporting innovation?

Try to understand: how we make decisions about adverts

Evidence from decades of research in behavioural economics, marketing and psychology demonstrates the huge power of advertising. Good adverts can inform, persuade and inspire people to buy products and services that they want or need. But when adverts confuse or mislead, particularly in relation to high value, long term decisions about financial products, they can cause significant harm.

Adams and Smart (2017) examined the role of advertising in financial consumer decisions, in particular how consumers evaluate adverts and use the information provided. This review identified three major components to describe how consumers react and use adverts: See, Interpret and Act. **See** describes how consumers notice adverts. **Interpret** refers to the next stage of processing – comprehension and preferences about an advert. **Act** describes the actions a consumer might take, for

[6] Note that all these social media providers allow users to add images or videos which are then unlimited by characters, as discussed in the FCA's social media guidance.

example, purchasing the advertised product. We address each of these three components throughout the experiments, whilst examining the effect of risk warning information.

See: Our research investigates what properties make adverts and risk warnings most prominent, and how the timing of risk warning information affects whether it is seen or not.

Interpret: We investigate the effect of different risk warning presentations upon participants' overall impression of the product: does the presence of risk warning information, either early or late in the product selection process, affect participants' preference for a product? Furthermore, do different types of risk warning presentations improve participants' understanding of the risks conveyed by this information, and do participants demonstrate this understanding by selecting appropriate products given different situations and constraints?

Act: The action taken in response to an advert can be multi-faceted. For the purposes of our research and policy questions, we focus upon two: whether an advert encourages participants to seek out additional information about the product, and whether it increases participants' preference for (or likelihood of purchasing) a product. Crucially, we examine the effect of when and how risk warning information is presented.

Risk warnings are used across a range of industries from manufacturing, to aviation or food. In advertising, they can highlight the possible negative consequences of a purchase or (mis)use of a product at the point a decision is made and are intended either to discourage use (e.g. cigarettes) or reckless purchase (e.g. gambling products). Despite their prevalence, relatively little is known about their effectiveness in different contexts. The most well-documented evidence is for smoking labels, and shows that graphic, larger and more comprehensive warnings on packets increase knowledge of health risks and encourage smokers to quit (Hammond, Fong, McNeill, Borland, & Cummings, 2006; Fong, Hammond, & Hitchman, 2009), though this mostly comes from smokers' own testimonies. To date, little is known about how and whether people's behaviour changes because they have seen a warning, or about any long-term educational or cultural effects of risk warnings.

In financial products, the role of risk warnings is more nuanced. For products such as tobacco, it is relatively unambiguous that any consumption is unnecessary and harmful for the consumer. However, the majority of financial products provide an overall benefit for the consumer. Therefore the goal of risk warnings is not to reduce total consumption, but to ensure that consumers are properly informed of the risks associated with a product: consumers should have sufficient information (and understanding of that information) to select products that are suitable for their situation, and avoid those which may be unsuitable or harmful.

There is some existing evidence that risk warnings for financial products may not be as successful in this more nuanced goal. In an experiment on investment prospectuses, Cox and de Goeij (2016) found that risk warnings increased people's risk perception by about 5% but decreased the inclination to search for more information by 12%. At the same time, references to regulatory approval (in this case, automatic, rather than implying safety or quality) increased willingness to invest and decreased perceptions of risk.

To the authors' knowledge, there have been no experiments specifically investigating consumer preferences, understanding and search activity in response to risk warnings in social media adverts.

What comes first counts most: timing of information

We are strongly affected by context when interpreting adverts. Timing, colour, images, formatting, phrasing and framing all impact our attention to, and feelings towards, adverts, often more than the information they contain. In particular, the timing of information matters. Because people are likely to remember conjecture as truth (Henkel & Mattson, 2011; Searleman & Carter, 1988), and struggle to update their beliefs after hearing subsequent contradictory information, what comes first is disproportionately important (see Adams & Smart, 2017 for a summary). This is true even for relatively short time frames, with choices biased by the time that a decision maker spends looking at the first option they examine (Schotter, Berry, McKenzie, & Rayner, 2010; Krajbich, Armel, & Rangel, 2010).

The effect is often compounded by confirmation bias. In other words, we search for information to confirm our prior beliefs (Jonas, Schulz-Hardt, Frey, & Thelen, 2001). Early information influences subsequent information search after an initial preference is formed, such as decisions about which advert to click on. In addition, people making decisions often avoid being exposed to contradictory information by deliberately seeking out information to support their initial preference. This forms the rationale for many regulators' policies of standalone compliance: adverts must contain all required information on first contact, including relevant risks and other product information (e.g. exclusions), depending on what is triggered by the initial inducement language.

We hypothesised that the difficulty in assimilating late negative information might be exacerbated online. Making decisions onscreen rather than offline changes our behaviour in a number of ways, most notably in speeding up our thinking, making us more impulsive, and encouraging us to spend more (Benartzi & Lehrer, 2015).

Bridging the gap: methodological rationale

We wanted to bridge the gap between the academic literature on how people respond to positive and negative information and the day-to-day reality of how people make decisions about financial promotions. To do this, we carried out a series of escalating experiments, starting with single tweets and moving towards more lifelike scenarios using social media feeds. We slowly built evidence at each stage, designing each subsequent experiment to answer the questions raised in the last one, gradually adding complexity to simulate real-life decision making. Throughout, we investigated the effects of risk warnings and their design and timing on understanding, preference and search activity.

3 Experiments

Experimental set-up

We carried out two series of experiments. The first series investigated the **design** of risk warnings in social media and webpages and their effect on consumer understanding and preference. The second series investigated the **timing** of risk warnings to answer the question: should risk warnings be present from the very first social media post (standalone compliance) to achieve the best outcomes for consumers, or can they be provided later?

Ethical review and approval was provided by the University of Warwick Humanities and Social Sciences Research Ethics Committee. Experiment design, sample size, and analysis methods were planned prior to data collection.

Experimental environment

We chose an online experiment to investigate the research questions because of the high degree of control this gave us to manipulate tweet stimuli. It also allowed us to vary external validity throughout the course of the experiments; starting from a very controlled, simulated choice environment, involving ratings of single pieces of stimuli, and working towards a more lifelike choice environment, involving making decisions about purchases on the basis of scenario information. Our final experiments are closer to the "noisier" environment in which people make choices in real life. However, throughout, we use stimuli which are as closely aligned to real-life adverts as possible, even where this may introduce some noise, in order to make the findings as policy-relevant as possible.

Stimuli

All of our experiments used stimuli from the 10 product types in Table 3.1. These were chosen because they are frequently advertised via social media, they involve significant risk, and/or they are mass market products:

Table 3.1: Products used in the experiments

Product	Purpose	Explanation
Pensions	Savings	Tax-efficient way to save money to be used in retirement.
Exchange traded funds (ETFs)	Savings / investment	A type of investment fund traded on stock exchanges.
Crowdfunding	Investment	A way in which people and businesses (including start-ups) can try to raise money from the public to support a business, project, campaign or individual.
Mini-bonds	Investment	Way for individuals to lend money directly to businesses, usually for a set period of time.
Contracts for difference (CFDs)	Investment	Complex financial instruments which allow investors to speculate on the price of an asset. They are often offered

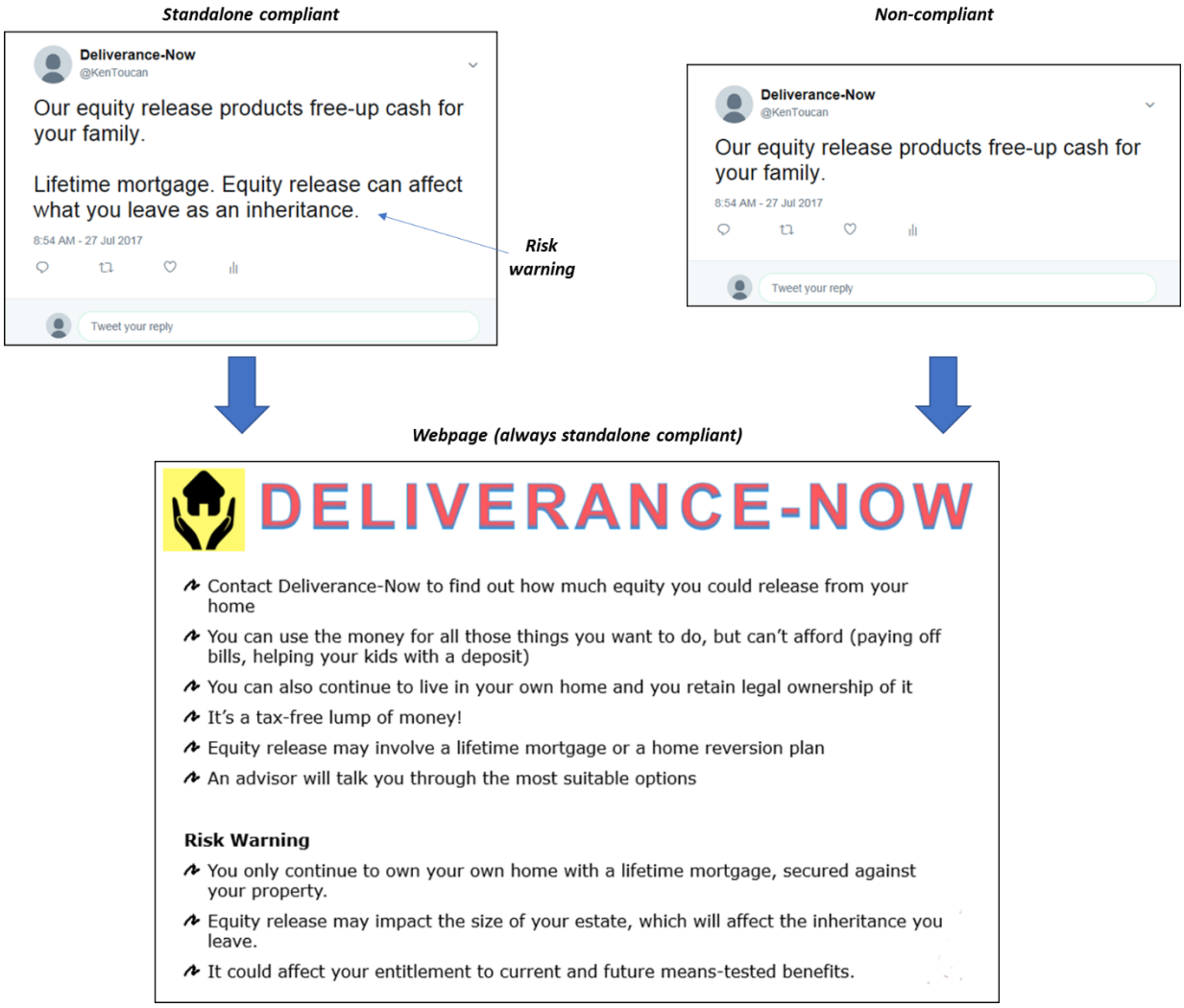
		through online platforms and typically with leverage which means investors only need to put down a portion of the investment's total value.
Logbook loans	Borrowing	Loan secured on a vehicle. The consumer can continue using the vehicle, but ownership of the vehicle transfers to the lender for the duration of the loan.
Guarantor loans	Borrowing	Loan where an individual other than the borrower or hirer provides a guarantee and/or indemnity if the borrower fails to repay.
Equity release	Borrowing	Loan which allows over 55s to access the equity (cash) tied up in their home. The money can be taken as a lump sum, in several smaller amounts or as a combination of both.
Payday loans	Borrowing	A relatively small amount of money lent at a high rate of interest on the agreement that it will be repaid when the borrower receives their next wages.
Car insurance	Insurance	An agreement to pay out a sum of money if the buyer's car is stolen, vandalised, catches on fire or is involved in an accident.

We created 100 fictional products (10 products for each of the 10 product categories) which were representative of existing products. For each of these products we created tweets, each with fewer than 140 characters.[7] Across experiments the tweet text remained the same, but the risk warning was included, removed or modified depending upon the specific question being addressed.[8] Each tweet was linked to a webpage which contained more information about the product as well a repeat of the risk warning, and additional risk information (see Figure 3.1 for an example of a standalone compliant and a non-compliant tweet and their associated webpage).

[7] Since completing these experiments, Twitter has increased the character limit for tweets to 280 characters.

[8] As mentioned in earlier sections, not all of the balancing information we tested would be categorised formally as a "risk warning". For example, the balancing information for crowdfunding products highlighted that the buyer would not be eligible for the Financial Services Compensation Scheme, which is not a mandated risk warning, according to the legal definition. For the purposes of this paper, "risk warning" can be understood in the non-legal sense to refer to any balancing information which informs the consumer of risks, costs or negative features of the product.

Figure 3.1: Example tweets and a webpage



The stimuli used in the experiments are stylised and simplified. It should not be assumed those appearing in this paper would necessarily be seen as fully compliant in the real world, as other factors may affect this.

Participants

In all experiments, we recruited participants from the website Prolific Academic. Participants carried out the experiment online. This approach provides increased validity, as the experiments address activity which would be performed online. These online samples have also been found to be more representative of the population than most traditional sampling methods used in behavioural research, and compare favourably with far more costly population level sampling approaches, which would be impractical for our purposes (Horton, Rand, & Zeckhauser, 2011; Mullinix, Leeper, Druckman, & Freese, 2015).

No participant took part in more than one experiment. Participants were recruited on the basis that they were over the age of 18, and currently living in the UK. Following the

experiments participants were excluded on the basis of the following criteria (which were defined in advance of the experiments):

- Participants who answered no to an "honesty" question that asked if they paid attention to the task.
- Participants who reported that they currently lived outside the UK (and thus were erroneously recruited), responses from non-unique IP addresses, and participants for whom the server logs showed errors or mismatches that resulted in missing data (though there were none in this latter category).
- Participants with the 2.5% fastest and slowest responses. This was to remove trials where participants responded unusually quickly or slowly, which is often an indication that a response button was pressed in error, or the participant became distracted.

Outcome measures

We measured behaviour in our experiments through a range of outcome measures summarised in Table 3.2 and discussed below.

Table 3.2: Summary of outcome measures

What was tested	**#**	**Preferences**			**Search**		**Understanding**	
		Tweet ratings	Product ratings	Choice	Clicks	Webpage time	Comprehension questions	Suitable product chosen
Readability	1A	√	√				√	
Pictures	1B	√	√				√	
FCA label	1C	√	√				√	
SC	2A	√	√					
	2B			√	√	√		
	2C			√	√	√		
	2D			√	√	√	√	
	2E				√	√		√
	2F				√	√		√

Preferences

We monitored participant preferences using three measures: subjective ratings of tweets, subjective ratings of products (tweet and webpage) and ultimate choice of product (where any product might have been suitable). We recognise that there is a difference between subjective ratings of the attractiveness of a product and ultimate choice of a product, which might involve other factors beyond preferences. There is a strong existing literature detailing numerous differences in behaviour between valuation and response (Lichtenstein and Slovic, 2006). However, one of the reasons for the interest in this literature, is that it documents the cases in which behaviour deviates from the generally very strong correlation between rating measures and choice measures. The specificity of these effects, and their relatively small impact compared to this overall correlation means it is unlikely that these different measures would yield significantly different preference estimates, and less likely still that they would interact in a meaningful way with our manipulations of risk warning presentation. To minimise any potential effects, instructions were framed to encourage participants to respond based upon how likely they would be to buy the product in real life.

Tweet ratings (used in 1A-C and 2A)
Participants were asked to rate individual tweets in response to the prompt: "How good do you think this product would be, if you were in the market for a product of this type." Participants responded on a 6-point scale presented below the tweet.

Product ratings (used in 1A-C and 2A)
After rating tweets, participants saw the webpage associated with the tweet and rated the product again on the same 6-point scale. We took ratings after seeing the tweet to ensure that participants were paying attention and we used product ratings as the main outcome measure, because participants had then been exposed to all of the available information and this would be the point in a real-life situation at which a purchase decision would be made.

Choice (used in 2B-D)
Participants viewed a social media feed with a number of adverts. To choose a product, the participant had to click on the tweet and then click "Choose" at the bottom of the associated product webpage. We measured which product participants chose.

Search
Clicks (used in 2B-E)
We measured all clicks that a participant made. Our outcome measures were:

- Which product was clicked on first
- Whether a product was never clicked on
- How many times each product was clicked on

Webpage time (used in 2B-E)
When clicking on a tweet, the participant could spend as long as they wanted looking at the associated webpage. They could then return to the social media feed and explore other tweets. Webpages could be visited multiple times.

We measured what effect standalone compliance had on the following outcome measures:

- The total time spent looking at a product webpage
- The average time spent looking at a product webpage per visit

Understanding
Comprehension questions (used in 1A-C and 2D)
We designed three sets of multiple choice comprehension questions: the first, used in 1A-C, the second used in 1A replication and the third, a composite set, used in 2D. Each set tested the understanding of the information contained in the tweet risk warning for each product. Each question had three possible answers, plus a "Don't know" option. All question sets are included in Annex 2.

In 1A-C, participants answered questions at the beginning of the experiment to generate a baseline and again at the end of the experiment. Because this might introduce priming (causing participants to attend or process information differently than if they were not primed with the issues addressed in the questions), in Experiment 2D, we presented the questions only at the end of the experiment. This still allows for between-subject comparisons of accuracy rates across conditions, but at the cost of statistical power as it does not allow us to control for an individual's baseline knowledge or accuracy.

Suitable product chosen (used in 2E & 2F)
In experiments 2E and 2F, participants were given a specific scenario and asked to find the most suitable product. The properties of each scenario meant that there was only one suitable product type based upon the associated risks of the available products. Therefore participants demonstrated their understanding of the risks by selecting an suitable over an unsuitable product.

Control measures

In every experiment, we recorded demographic data on the participants' age, gender, and time in the UK. We also measured other aspects of the participants' knowledge, experience and behaviour:

Financial literacy (measured in 2A)
At the end of the experiment, we asked three standard financial literacy questions, included in Annex 3. We found that, in line with existing literature, this was not predictive of any of our outcome measures, despite the task being very time consuming, and so we discontinued their use in later experiments to reduce the length of the experiment and the burden on participants.

Product familiarity (measured in all experiments)
At the end of the experiment, we asked participants to rate their familiarity with the products in the experiment on a 7-point scale. The distribution of familiarity ratings was relatively consistent across all experiments, indicating that any differences in behaviour between experiments is unlikely to be caused by between-sample differences in product familiarity.

Speed of responses (measured in all experiments)
Response times were primarily used to exclude very fast and very slow responses, as these are likely to represent times where the participant did not pay full attention before responding (or responded accidentally), and where the participant became distracted during their choice. In Experiment 2E and F, we found faster participants were more likely to choose suitable products, suggesting that this measure may be a proxy for intelligence and/or attention to the task.

Screen position and miscellaneous
Finally, we use, where appropriate, controls related to the specific experimental design. These include whether a product appeared at the top of the social media feed, and in 2E and F, whether one or two of the most suitable product types appeared at the top of the social media feed as well as the number of alternative products to choose from. These controls are introduced for each experiment.

Analysis

The specific analysis design is introduced with each experiment, but we give an overview here of the general approach employed. Primary analyses were centred around mixed model regressions. When predicting continuous data, a linear model was used and when predicting binary outcomes (e.g. choice, or question accuracy) a logistic model was used. These regressions included random intercepts and slopes for all predictors that had sufficient data points measured within each participant. For example, in Experiment 1A, these were included for main effects of warning readability upon risk comprehension accuracy, because each participant answered multiple comprehension questions after

seeing both high and low readability warnings for multiple different products. They were not included for product accuracy predictors because each participant answered only two questions for each product, meaning there was insufficient data to calculate participant level random effects.

Power calculations are noisy and the lack of previous literature examining these specific effects meant we had little existing data to base power calculations upon.[9] Therefore, we did not use a strict cut-off point or calculation when deciding upon a sample size. Instead, we took input from other behavioural experiments with similar general properties that affect statistical power (e.g. within vs between sample comparison, binary vs continuous outcome measures, number of trials per participant and per comparison), and used this to select a sample size that would be larger than that generally seen in the existing behavioural literature. Results are reported as statistically significant if $p < 0.05$, and full confidence intervals are reported in all appropriate analyses.

[9] http://datacolada.org/20

Experiment series 1: Design of risk warnings

1A: Readability

Background

Most risk warnings for financial products in the UK do not have prescribed wording, but instead must communicate a message. For example, firms must ensure that a financial promotion "for a product or service that places a client's capital at risk makes this clear", but firms may articulate this in their own words.[10] This means that some risk warnings may be more readable than others. We investigated the effect of risk warning readability on consumer understanding and preference.

Method

Participants

We recruited 215 participants, 10 of whom were excluded from the analysis. The mean age was 35.8 (SD = 11.2), with a gender distribution of 78% female, 22% male, and <1% other.

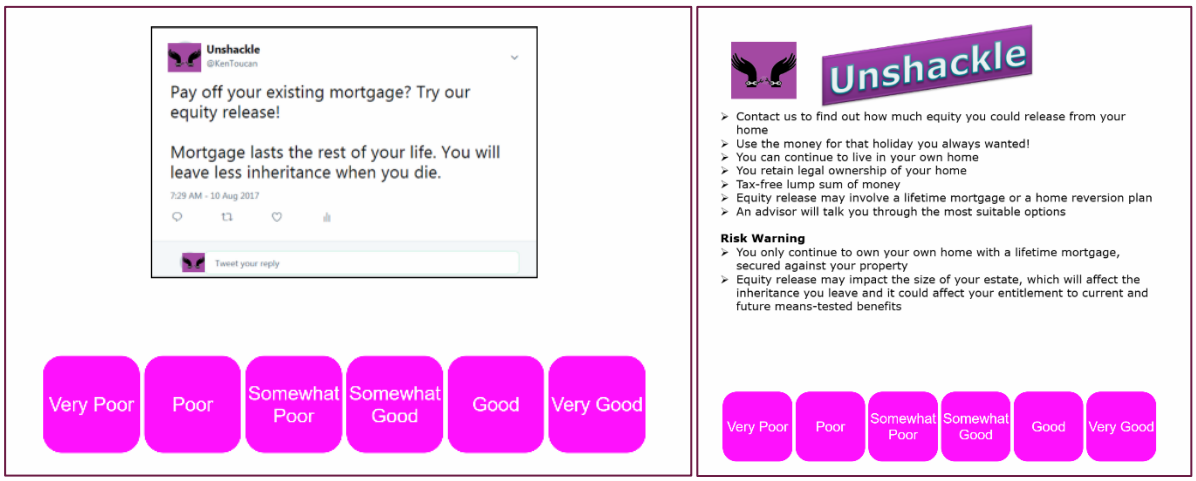
Stimuli

The stimuli comprised 40 different hypothetical financial products, 4 from each of the 10 categories. Each product had a tweet and an associated webpage and all stimuli contained a risk warning.

We used two versions of each risk warning: the original version, based upon the most prevalent wording seen in the marketplace, and a rewritten warning. Rewritten warnings were developed using insights from behavioural science to improve readability and to better communicate the risks involved (see Figure 3.2). The primary metrics used in designing the re-written warnings were psychological measures of concreteness, and word frequency (how often a word is used in everyday communication). Messages were written using words that scored highly on both of these metrics. Whole messages were also checked against commercial readability calculators to provide an additional measure of the improvement in readability, however, these tools must be used with caution. In some cases they provided implausible scores, seemingly because phrases such as "FSCS" did not exist in their dictionary, meaning such specific terms did not lower the readability score.

[10] FCA Handbook, COBS 4.2.4(G)

Figure 3.2: Example of tweet and webpage stimuli. The tweet contains a rewritten, high readability warning, the webpage contains an original, low-readability warning



Procedure

The experiment was a 2 (high vs low tweet readability) x2 (high vs low webpage readability) design, resulting in four conditions: 1. Low tweet, low webpage readability, 2. Low tweet, high webpage readability, 3. High tweet, low webpage readability, 4. High tweet, high webpage readability.

There were 5 stages or blocks in the experiment: Comprehension questions 1, product ratings 1, comprehension questions 2, product ratings 2, and finally, demographic and familiarity questions.

Before beginning any trials or seeing any of the stimuli, participants answered two multiple choice questions about each of the 10 types of financial products – 20 overall (see Comprehension Questions Set 1 in Annex 2). During the rating trials, participants were asked to imagine that they were in the market for the type of product presented in a given trial. In the first block of rating trials, they saw adverts for 10 products - one for each product type. In each trial, participants first saw the tweet and rated how good they thought the product was on a 6-point scale. They then saw the webpage for the product and again rated how good they believed the product to be. For each product, the readability of the tweet and webpage were randomly assigned. After completing the first block, participants then answered the comprehension questions again before rating a second block of 10 products in the same way. Finally, participants provided basic demographic information and rated their familiarity with each type of financial product used in the experiment using a seven-point scale.

Empirical methodology

We used a mixed model regression approach. When predicting accuracy on comprehension questions, we predicted responses in the second block and a logistic model was used to predict correct versus incorrect responses to each question. Two models were estimated: one to provide an estimate of the overall effect size across all product types, and one to estimate the effect sizes specific to each product type. For the model estimating the overall effect, we used dummies for:

- whether the correct answer had been given in Block 1 (the participant's accuracy before seeing any products),
- whether the tweet for that product type included a high readability risk warning,
- whether the webpage for that product type included a high readability risk warning, and an interaction dummy for tweet * webpage readability.

In addition, to account for the heterogeneity between comprehension accuracy for different product types (see below), dummy variables were included for each product type with pensions being used as the baseline category. For the model which estimated product specific effects, the same predictors were included, as well as interaction dummies between product type and our three predictors of interest: tweet readability, page readability, and tweet * page readability.

When predicting ratings, the same approach was employed of estimating a model to estimate the overall effect, and a model to estimate the product specific effects. The same predictors were used as in the models estimating comprehension accuracy, with one addition: the risk comprehension accuracy for the second block of comprehension questions. The main differences in analysis approach were that a linear model was estimated, and that models were estimated for ratings in both the first and second block.

Given the heterogeneity of baseline accuracy and baseline ratings across different product types (see below), we describe the model's mean effect predicted change in comprehension from a baseline of 50%. Compared to reporting mean response proportions, this approach helps to ameliorate the effects of this heterogeneity on the overall effect size, reducing the potential over and underweighting of different product types.

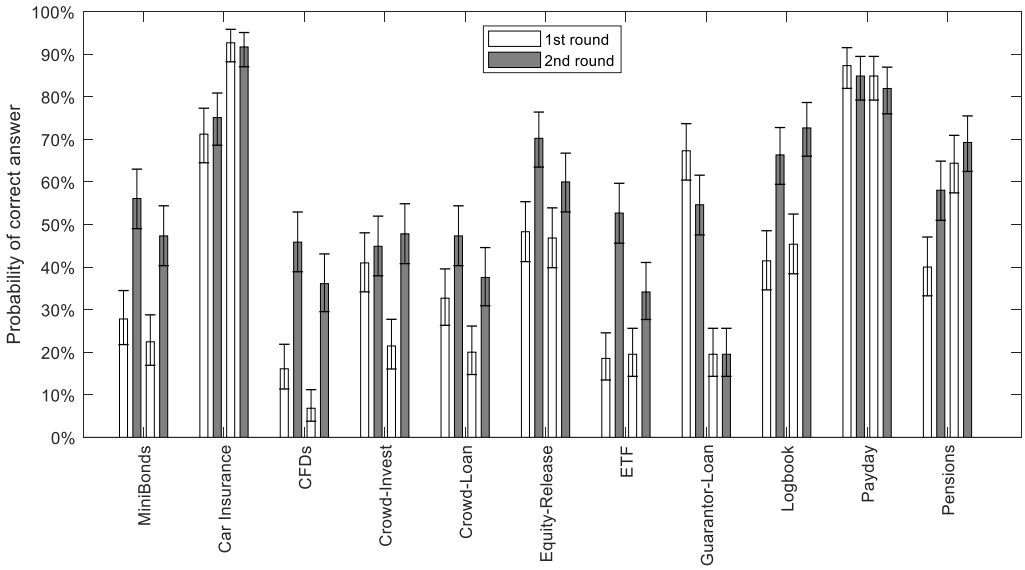
Results

Understanding (comprehension questions)

Baseline comprehension

When collapsing across conditions, average comprehension improves from the baseline of 42.5% to 57% when answering questions for the second time. Similar to the baseline, this includes significant heterogeneity, ranging from 19.5% for one guarantor loan question to 91.7% for one about car insurance. Figure 3.3 shows baseline comprehension.

Figure 3.3: Mean comprehension (%) for comprehension questions (two per product type) before and after seeing the first block of rating trials



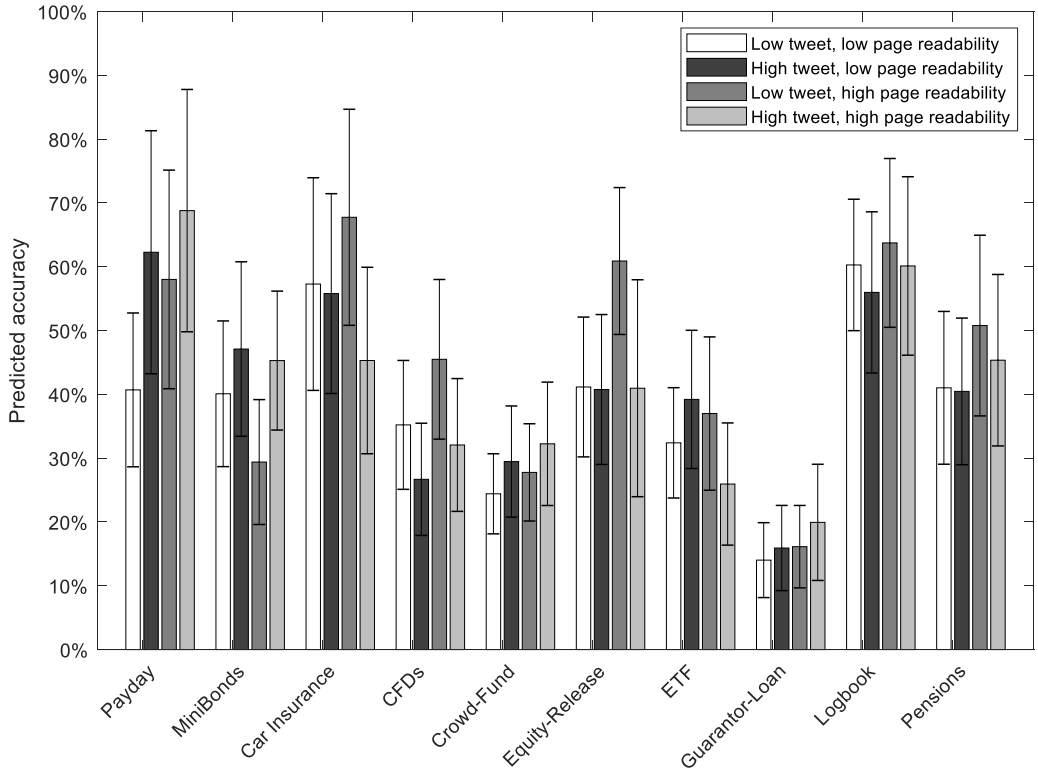
While participants showed improved comprehension of some products after exposure to the first set of adverts and webpages, other products, such as car insurance, saw no improvement.

<u>Comprehension</u>

The strongest predictor of comprehension in Block 2 comprehension questions was Block 1 comprehension: if a participant answered correctly the first time they saw the question, it was likely that they did so again the second time they saw it, confirming a strong effect of existing knowledge.

There is a significant main effect of webpage readability: a hypothetical individual who had a 50% chance of answering correctly when readability was low would have a 58.3% chance when readability was high (see Table 1A.1 in the Appendix). There was no effect of tweet readability on comprehension. While we saw effects on average across products, there were no strong product-specific effects of tweet or webpage readability changes (see Figure 3.4, Table 1A.2 in the Appendix).

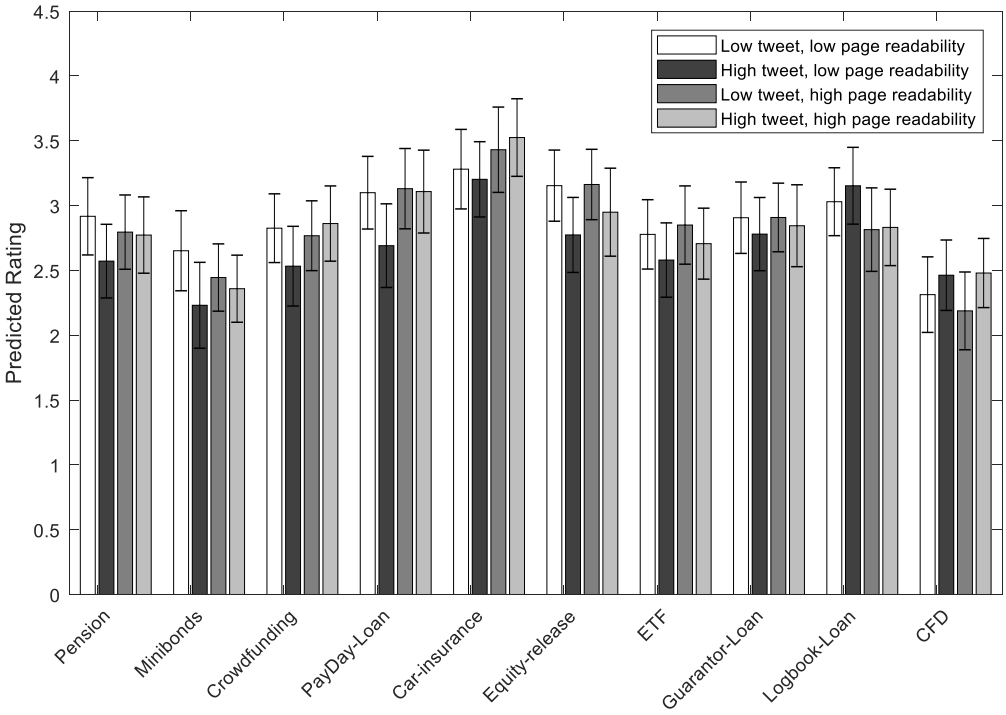
Figure 3.4: Predicted comprehension (%) by product and treatment for the second set of comprehension questions, if the question was previously answered incorrectly.



Preferences (tweet and product ratings)

In the first block, there was a negative main effect for tweet readability on product ratings (preference), and a positive interaction of a similar magnitude between tweet and site readability. This would suggest that the condition which results in the lowest overall ratings is a tweet with a highly readable risk warning followed by a webpage with a low readability one (see Figure 3.5 and Table 1A.2 in the Appendix). There are no strong product-specific effects.

Figure 3.5: Predicted ratings by product and treatment (Block 1)



There were no effects of readability on ratings in Block 2 (Table 1A.3 in the Appendix).

Discussion

The results show a modest effect of risk warning readability upon risk comprehension, but only for the risk warning on the webpage. No change was found when the tweet risk warning is made more readable. This may indicate that viewers pay more attention to website text than tweet text, or that the necessarily short message attached to the tweet does not contain enough information to affect comprehension. Irrespective of the underlying cause, this result suggests that the information on the web page has greater influence on risk comprehension (this is consistent with findings in Experiment series 2).

There is also a small effect on product ratings, with ratings being slightly lower on average for products with highly readable tweets, and low readability websites. However, the effect was only present for ratings in the first block, and the effect size is notably smaller than that for comprehension. Combining this observation with results from the next experiment suggests that the effect on ratings is likely to be a chance finding.

The design of the experiment involved exposing participants to the comprehension questions before seeing any tweets in order to establish baseline knowledge (this is the same in Experiments 1B and C). This introduces the possibility that participants may be primed to identify this same information in risk warnings and therefore may pay more attention to it than they would otherwise. However, if anything, this should reduce the impact of experimental manipulations we tested, so the results we find are likely to be conservative estimates.

1A replication

Background

This experiment tests an alternative set of comprehension questions (one per product), designed to reduce the ceiling and floor effects observed with Comprehension Question Set 1 (see Annex 2). These new questions were developed using feedback from a sample of FCA and University of Warwick employees who provided input about which questions they found hard and easy and why. The individual questions were designed to be a more complete test of the understanding of the risks associated with a product type. However, in being more encompassing this only allowed scope for one question for each product type (as opposed to two in the previous experiment). These new questions were then used in a direct replication of the previous experiment.

Method

Participants

We recruited 150 participants and removed 7 according to the exclusion criteria. The final sample after exclusions was 143, of whom 61% were female and 39% were male. The average age was 36.0 (SD, 10.6).

Stimuli

The stimuli were the same as those used in Experiment 1A, but the alternative set of 10 comprehension questions, one per product, was used (see Annex 2).

Procedure

The procedure was the same as in Experiment 1A.

Empirical methodology

The methodology was the same as in experiment 1A.

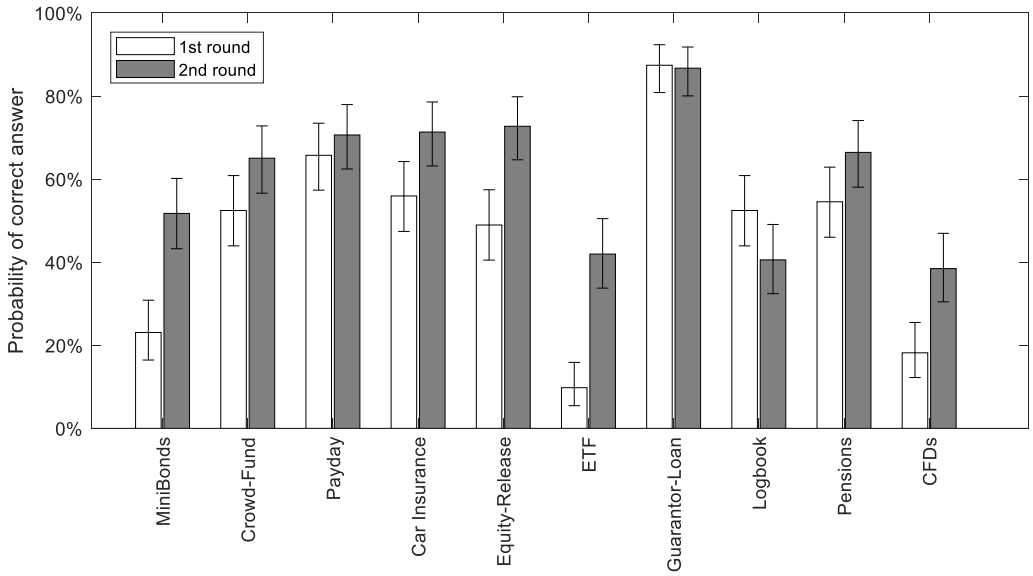
Results

Understanding (comprehension questions)

Baseline comprehension

Participants' accuracy when answering the questions for the second time improved from an average of 46.9% to 60.6% (see Figure 3.6). Similar to Comprehension Question Set 1, the largest gains in accuracy from one block to the next were mini-bonds, CFDs and ETFs.

Figure 3.6: Mean comprehension in % (one question per product type) before and after seeing the first block of rating trials



Effect of readability on comprehension

First we predict accuracy on the comprehension questions the second time they were answered. The estimated effect size for webpage readability (an increase of 7.5% from a 50% baseline) was similar to that of experiment 1A (8.3%), but the confidence intervals were wider, suggesting that this experiment had reduced statistical power, or increased variance in responses (see Table 1A.6 in the Appendix). As a result, this effect is not statistically significant. When a combined analysis is performed using data from both experiments, the confidence intervals are reduced, and readability of the webpage warning is a significant predictor across the experiments. However, there is no main effect of tweet readability, and no interaction effect (see Table 1A.7 in the Appendix). While this should be interpreted with caution as we did not plan this analysis in advance, it suggests that the lack of significant effect in the replication is due to reduced power.

Preferences (tweet and product ratings)

There was no significant effect on ratings (see Table 1A.8 in the Appendix). There were no strong product-wise effects.

The results show no effect of readability on ratings in Block 2. Ratings were higher when comprehension on the second set of questions was higher (see Table 1A.9 in the Appendix).

Discussion

The results for the current experiment show reduced statistical power, but do not refute the finding that improved readability of webpage risk warnings increases understanding. There is little or no effect of improving tweet warnings.

This experiment did not corroborate the finding in Experiment 1A that high readability tweets and low readability websites reduced ratings. This may be taken as further evidence that this was a chance finding and not robust to replication.

The results from this experiment suggested that the set of comprehension questions used in the prior study provided greater statistical power. This is in part because their design allowed for two questions per product type, whereas the questions in this new set were designed to avoid floor and ceiling effects, but several did so by combining elements of understanding from both questions used in the original set. These results demonstrate that this trade-off was not fully successful and therefore we designed a composite set of questions from the two sets to be used in Experiment 2D.

Overall, the results support the conclusion that using insights from behavioural science to improve the readability of risk warnings on product web pages results in a small but statistically significant improvement in participants' understanding of the associated risks. However, there is no such effect when risk warnings on tweets are improved in a similar manner.

1B: Images

Background

Risk warnings for some consumer credit products must inform the reader of the Money Advice Service (MAS). Firms can do this by providing the MAS logo rather than the website address. Some regulators also require firms to accompany their risk warnings with a set image that depicts the overall risk of the product type. It is plausible that these images could attract attention to the risk warning part of the tweet, thus improving understanding, or that they could attract attention away from the actual text of the risk warning, potentially harming understanding. Following other work showing an effect of images upon attention and consumer perception, we investigated the impact of a MAS picture and logo on consumer understanding and ratings in the context of social media.

Method

Participants

We recruited 160 participants and removed 13 according to our exclusion criteria. The sample had a mean age of 36.2 (SD 10.8), with 62% female and 38% male.

Stimuli

The stimuli were 20 different hypothetical financial products, 2 each from the 10 different categories of financial product. Each product had a tweet and a webpage, both of which contained risk warnings. The product webpage always contained a text-only risk warning. This is consistent with most real-life webpages. For the tweet, the risk warning could be from one of two conditions:

- Original risk warning
- Risk warning with the same text but also a link to the MAS website. This included the MAS logo, and the cartoon image header taken from the MAS webpage for that particular financial product (see Figure 3.7).

Figure 3.7: Example tweets with MAS pictures

Images varied for each financial product but were similar in content and design, usually including a cartoon figure with paperwork (for credit products) or with money (for investment products).

The inclusion of MAS pictures and logo for all products goes beyond the FCA's rules and guidance and should be seen simply as an example of images which are linked to risk warnings for the purpose of testing.

Procedure

The procedure was the same as in Experiment 1A.

Empirical methodology

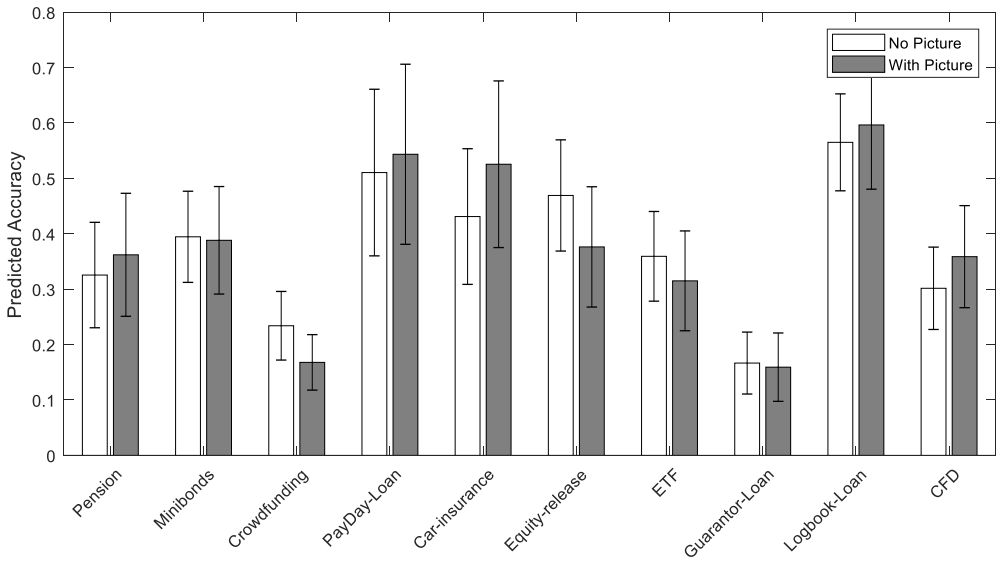
The same analysis approach was used as in experiment 1A. The only difference was that the experiment manipulated just one property of the stimuli (picture presence on tweets), as opposed to two (risk warning readability for both tweets and for web page).

Results

Understanding (comprehension questions)

There is no significant effect of the picture upon comprehension, suggesting that its presence does not distract attention, nor orient it towards the risk warning. There are no notable product specific effects (see Figure 3.8 and Table 1B.1 in the Appendix).

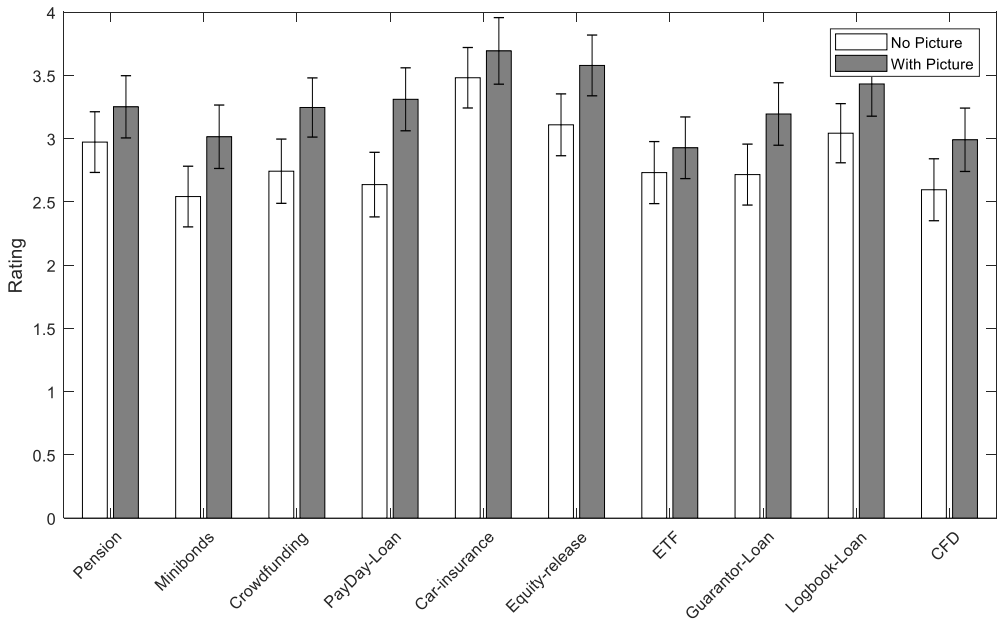
Figure 3.8: Effect of picture inclusion on risk comprehension by product: the predicted likelihood of correctly answering risk understanding questions after seeing the adverts if the question had been answered incorrectly beforehand



Preferences (tweet and product ratings)

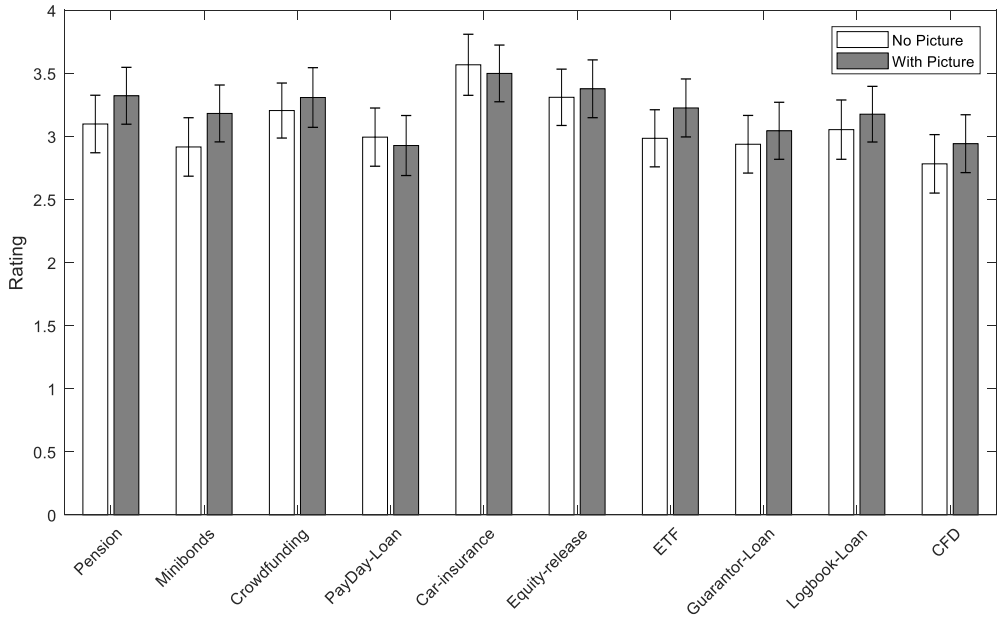
We see a large and statistically robust effect of the picture upon preferences in Block 1. Product ratings are significantly higher when the MAS picture is included in the risk warnings: moving from an average rating of 2.98 out of 6 without pictures to 3.22 with pictures. After standardising for the distribution of ratings, we see a difference of 0.25 standard deviations (see Table 1B.2 in the Appendix). Figure 3.9 shows the effect on ratings across product types. All 10 product types show a directional trend to receive higher ratings when the picture is included. The largest effect is found for payday loans, where the effect is strong enough that it is statistically significant within the single product type. The same is true for crowdfunding, guarantor loans and equity release.

Figure 3.9: Effect of picture inclusion on ratings by product (Block 1)



In Block 2 there is a significant overall effect of the picture on ratings, though this is smaller than that found in Block 1 (see Table 1B.3 in the Appendix and Figure 3.10). The average ratings were 3.17 for adverts containing pictures, and 3.05 for those without, a difference of 0.10 standard deviations.

Figure 3.10: Effect of picture inclusion on product rating by product (Block 2)



Discussion

Friendly, money-related pictures increase consumer perceptions of product attractiveness, but have no discernible effect on risk understanding. It is possible that the picture has a diminished effect in Block 2 because participants had already seen multiple examples of the MAS picture and became habituated to it.

It is not immediately clear what the implications are of the increase in ratings for a tweet risk warning that contains a picture. One suggestion may be that the content of the picture itself changes participants overall perception of the risk, through some halo effect or by distracting attention from the content of the risk warning. However, there was no detectable effect upon risk comprehension accuracy. Another possibility is that the specific content of the image conveys a particular message e.g. one containing pictures of money implies a strong return for investment products. The limited scope of alternatives that could be tested in this one experiment mean that we cannot address this question in much detail. However, the results across different product types do not suggest that there is a particular difference between borrowing and investment products, despite only the latter generally being paired with MAS images including cash. In general though, the positive effect of images in risk warnings upon product ratings should be considered when evaluating advertisements and advertising guidelines.

1C: FCA warning label

Background

The FCA does not require firms to label risk warnings as being mandated by the FCA, even if they are. However other studies have shown that labelling information could alter the perception of the information (Cox & de Goeij, 2016). There are three possible hypotheses for how consumers might respond to a FCA label, if the FCA were to require this. It could:

a) lend the warning credibility and attract the attention of consumers,
b) be ignored, perhaps because consumers see it as a requirement rather than information that is relevant to them, or
c) imply that the FCA approves or recommends the product (which would be in conflict with other requirements).

In this experiment, we tested the impact of adding the words "FCA Mandated Risk Warning" before tweet risk warnings on preferences and understanding.

Method

Participants

A total of 160 participants were recruited and we removed 11 according to the exclusion criteria. Of the 149 remaining, 60% were female, 40% were male, and the mean age was 33.0 (SD 10.3).

Stimuli

The stimuli were the same 40 different hypothetical financial products used in experiment 1A. We produced two versions of each of the tweets: the original and an alternative that clearly labelled the risk warning as coming from the FCA ("FCA Mandated Risk Warning", see Figure 3.11).

Figure 3.11: Tweet with 'FCA Mandated Risk Warning' label

Procedure

The procedure was the same as in Experiment 1B.

Empirical methodology

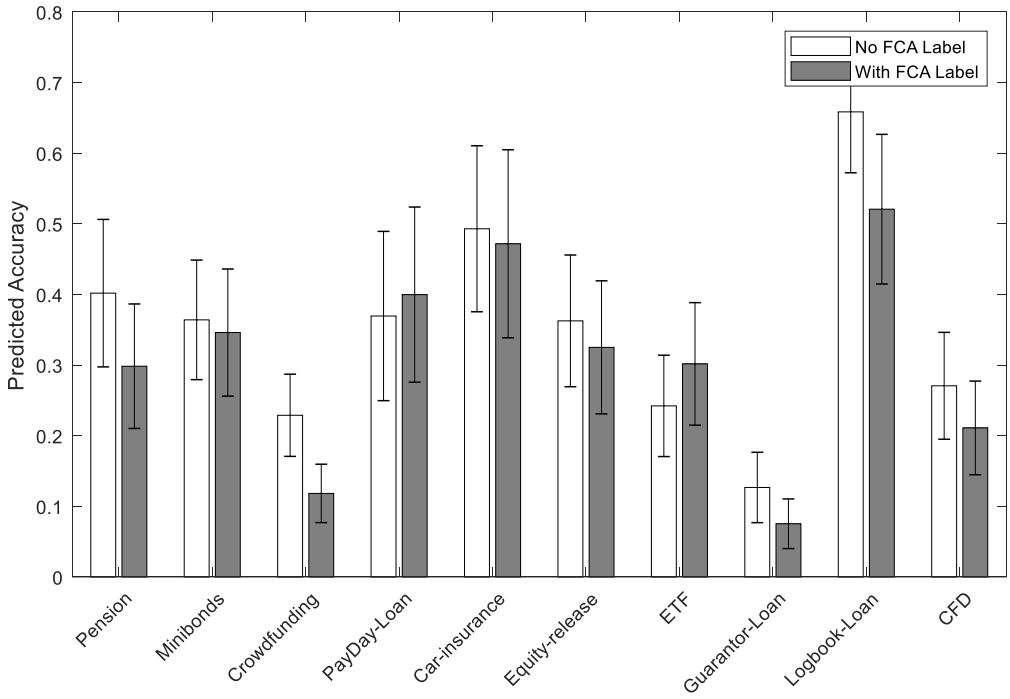
The same analysis approach was used as in Experiment 1B.

Results

Understanding (comprehension questions)

When the FCA label was present, comprehension was significantly lower, reduced by 7.0% for an individual who otherwise had a 50% chance of answering questions correctly (see Table 1C.1 in the Appendix). Marginal effects show that the direction of the effect was negative for 8 of the 10 product types, but when looking within product types, we can identify a product-specific effect for crowdfunding (see Figure 3.12).

Figure 3.12: Effect of FCA label on risk comprehension by product



Preferences (tweet and product ratings)

Despite reducing comprehension, the FCA label had no discernible effect on ratings in Block 1 or Block 2 (see Tables 1C.2 and 1C.3 in the Appendix). There were no significant product-specific marginal effects.

In Block 2, products received a higher rating if the participant correctly answered the relevant comprehension questions at the beginning of the experiment.

Discussion

Labelling risk warnings as originating from the FCA decreases understanding of the message. This might be because participants tune out formal-looking messages, similar to the way people tend to focus attention away from adverts when they are concentrating on online content (known as 'banner blindness'). It is also possible that participants were not familiar with the FCA or the word "mandated" and this led them to screen this message out. Alternatively, the mention of the FCA could be inferred as approval or a recommendation, which could have the perverse consequence of reducing attention to the risks.

Irrespective of the underlying mechanism, we can conclude that emphasizing that the risk warning was required by the FCA will not have a positive impact upon individuals' understanding of the risks conveyed within that message.

Discussion: Design of risk warnings

The design of risk warnings can affect consumer understanding and attractiveness ratings. Our results show that more highly readable warning text within webpages delivers modest improvements in consumers' understanding of associated risks.

However, this effect is only seen for risk messages on product webpages, not for short warnings presented as part of social media posts. Relevant, friendly pictures designed to draw attention to the risk warning actually improve product ratings (even though the message they contain is still overtly negative) and an "FCA MANDATED WARNING" label reduces understanding of risk warnings.

It is perhaps understandable that we should see a greater and more stable effect of very prominent changes, such as including pictures. Since risk warnings are present on large numbers of adverts, it is possible that consumers become accustomed to them and no longer read them in full unless something dramatic has changed. Attention to, and reading of, different types of risk warnings would be a fruitful avenue of future research, particularly eye tracking, in which we can observe where a consumer focuses their gaze.

Experiment series 2: Timing of risk warnings

We carried out six escalating experiments designed to investigate how the timing of risk warnings affect behaviour in a simulated real-life setting. Specifically, we compared situations where participants see a risk warning when they initially see a tweet, to situations where they see it for the first time after clicking through to a product webpage. The experiments used the same stimuli as Experiment series 1, but applied them across a wider range of scenarios and tasks. The mix of tasks employed allowed us to measure aspects of preference, information search, and understanding. Experiments and the outcomes we measured are shown in Table 3.3.

Table 3.3: Summary of outcome measures for Experiment series 2

What was tested	**Preferences**			**Search**		**Understanding**	
	Tweet ratings	Product ratings	Choice	Clicks	Webpage time	Comprehension questions	Suitable product chosen
2A) Rating single tweets.	√	√					
2B) Social media feed, best of single product type			√	√	√		
2C) Social media feed, simple scenarios, best of multiple product types			√	√	√		
2D) Social media feed with large number of alternatives			√	√	√	√	
2E) Social media feed, mixed compliance, specific scenarios, identify suitable product type based on risks				√	√		√
2F) Social media feed, standalone compliant environment vs non-compliant environment, specific scenarios, identify suitable				√	√		√

product type based on risks							

2A: Rating single tweets

Background

Information that comes first tends to receive more attention than information that comes later, especially if the reader is already invested in a decision to buy. In fact, confirmation bias can cause individuals to seek out confirming information and interpret information as consistent with prior beliefs. Individuals may also actively avoid information that conflicts with these early beliefs (Golman, Hadmann & Loewenstein, 2017). Experiment 2A tests whether this is true of risk warnings presented within tweets: Do risk warnings have reduced impact on judgements of attractiveness if this information is presented late in the process, i.e. after clicking through to the product webpage?

Method

Participants

We recruited 400 participants and removed 36 on the basis of the exclusion criteria. Of the remaining 364 participants, 57% were female and 43% were male. The mean age was 36.5 (S.D. 12.0).

Procedure

The procedure was similar to the rating blocks in the Experiment series 1: Participants saw a tweet and rated how good they thought the product was on a six-point scale ("tweet ratings"). This was to ensure that they fully engaged with the information at this early stage and did not simply click through immediately to see the product webpage. They then saw the corresponding webpage and rated the product again ("product ratings").

Stimuli were 20 products: two from each of the 10 product categories. These were presented to participants as two blocks of 10. Each block contained one product from each product category. For each participant, 5 product categories were randomly selected to be presented as standalone compliant in the first block, then not standalone compliant in the second block. For the other 5, the ordering was reversed. By using only two examples of each product type it was possible to get a rating from each participant for each of the 20 products. Within each product type one of the products would have been presented as standalone compliant and the other not. Therefore, we can directly analyse the difference in ratings between the two (with statistical controls to account for potential underlying differences in attractiveness between the two individual products).

At the end of the experiment, in addition to providing basic demographic information, participants rated their familiarity with the 10 product types using a 7-point scale and answered 5 standard financial literacy questions (see Annex 3).

Empirical methodology

The design of the experiment allows us to directly predict the within-subject difference in ratings between a standalone compliant and non-standalone compliant example of a

particular product type. Thus, it is this difference that is used as the outcome measure in mixed model regressions. Separate models are estimated for tweet ratings and product ratings.

Similar to the analysis approach employed in Experiment series 1, two types of model are estimated: one that estimates the overall effect of standalone compliance across all product types, and one that estimates the mean effects within each specific product type. The first of these includes only the intercept term (which captures the difference in rating), and dummy variables for each of the product types which indicates which of the two products was presented as standalone compliant. The second model also includes main effect dummy variables for each product type.

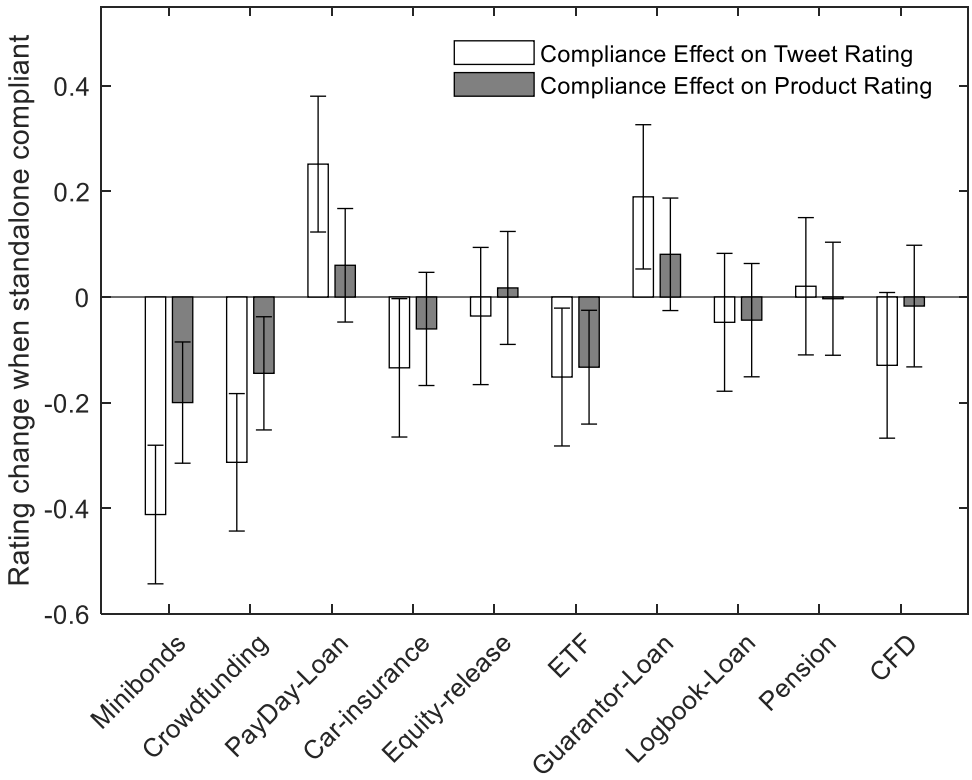
Results

Preferences (tweet and product ratings)

When only the tweet has been seen, tweet ratings for standalone compliant tweets are significantly lower than for compliant tweets, by 0.12 ratings points or 0.038 standard deviations (see Table 2A.1 in the Appendix). A similar, but smaller effect is present for product ratings given after the full webpage has been seen, with standalone compliant tweets receiving a rating 0.06 points lower on average, or 0.018 standard deviations (see Table 2A.2 in the Appendix). This suggests that although the later risk information on the webpage reduces the impact of missing risk warnings in tweets, participants' behaviour is still impacted by the presence or absence of early risk information. However, in both cases the effect is small, and only statistically significant because the within-subject experiment design provides particularly good statistical power.

We also examined the impact of standalone compliance across different product types. This revealed heterogeneity between products. Figure 3.13 shows that for payday loans in particular, standalone compliant tweets tended to be rated more favourably both in tweet and product ratings.

Figure 3.13: Effect of standalone compliance on tweet and product ratings by product and advert medium



Discussion

While on average standalone compliance reduces participants' ratings of products, it appears to affect different product types differently. It is possible that standalone compliance improves participants' judgements of payday loans because pictures are included as part of the risk warning through the mandated MAS link, as investigated in Experiment 1B. It is also possible that standalone compliance improves participants' judgements of certain products because the risk warnings are less severe than expected i.e. people's existing impression of these products is so low that a somewhat negative risk warning improves their perception (baseline ratings for payday product types are low on average).

Participants' accuracy on financial literacy questions was similar to that of the general population. However, there is strong existing evidence that general financial literacy is a poor predictor of individuals' choices of financial products (Fernandes, Lynch & Netemeyer, 2014), and these questions took an unexpectedly long time for participants to complete. Therefore, they were not collected in subsequent experiments in order to save time and allow for more flexibility in other areas of experiment design.

The effect of standalone compliance is greater upon tweet ratings. This is potentially important because this is the point at which individuals must decide whether or not to click through. So even if the effect is ameliorated after individuals see the full webpage, differences in perception of the tweet might alter clicking decisions, and therefore change information search. The following experiments investigate search activity in addition to preference and understanding.

2B: Choice between four products

Background

Consumers do not normally view one tweet at a time. We therefore simulated an environment in which participants saw a social media feed with multiple tweets about the same product category. We then measured their behaviour as they explored the adverts to select what they thought was the best product within each product type.

Method

Participants

A total of 199 participants were recruited and 17 were excluded. Of the remaining 182 participants, 70% were female, 30% male, and <1% other. The mean age was 25.0 (S.D. = 4.2).

Procedure

Participants saw a page that was designed to look like a social media feed (see Figure 3.14). The feed displayed four tweet-style adverts for similar products from a single product type (e.g. payday loans). Participants could click on the tweets to reveal a full webpage which contained additional information. They could then either click to choose that product (analogous to buying the product from the provider's website), or click back to return to the social media feed.

Each participant saw 10 sets of 4 tweets, one set for each product type, and had to choose one product from each set.

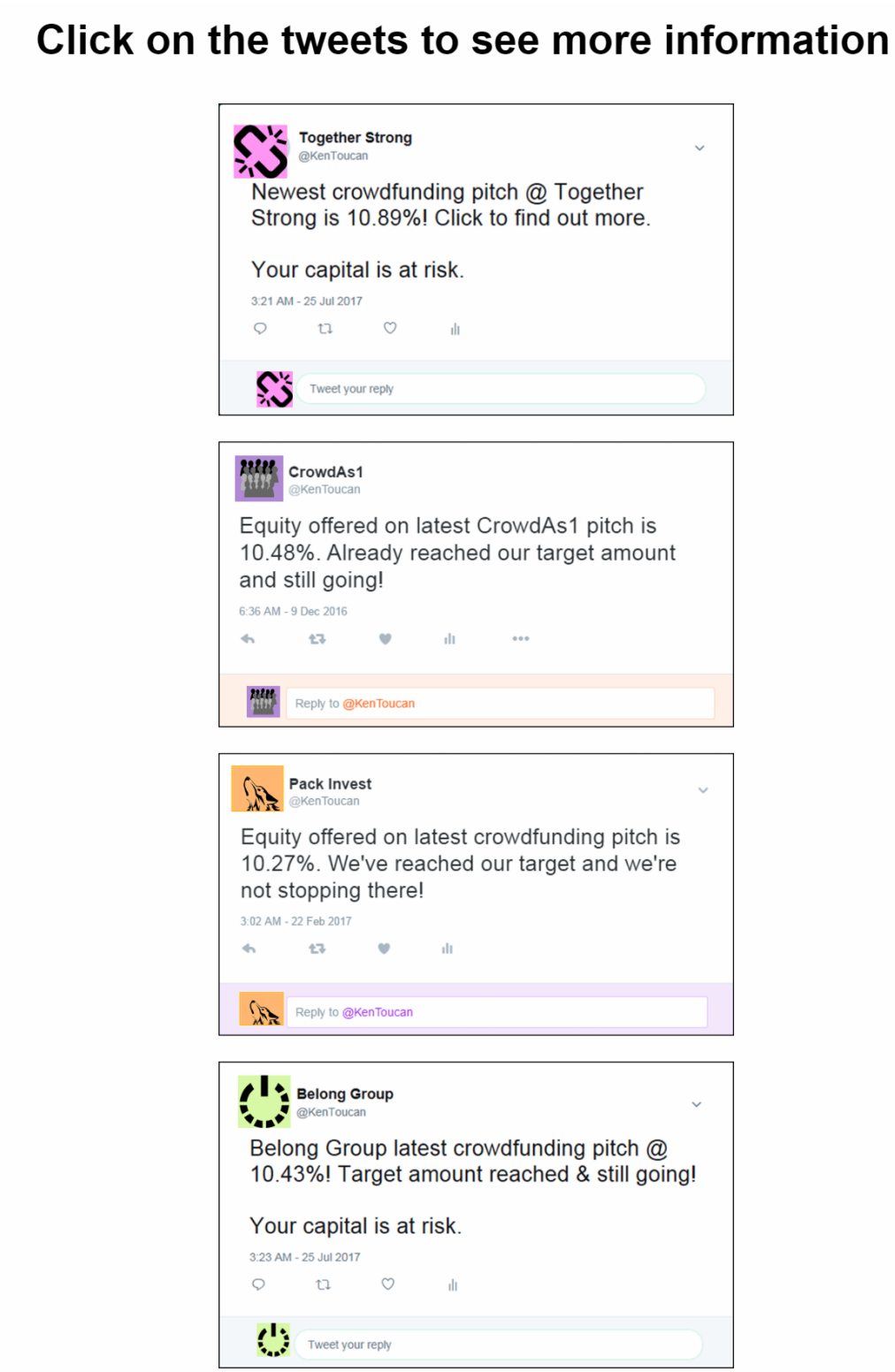
There were three conditions:

- **Compliant:** All of the tweets were standalone compliant.
- **Non-Compliant:** None of the tweets were standalone compliant.
- **Mixed Compliance:** Two of the four tweets were standalone compliant, the other two were not.

For each participant, three of the product types were allocated as standalone compliant, three were allocated as non-compliant, and four were allocated as mixed compliance.

During the choice task, all clicks were recorded, allowing us to measure, not just what product was chosen, but also which items were clicked on, in what order, and how long the webpage information was viewed. In addition, the size of the browser window was recorded so that it could be determined what tweets were presented in the screen as the choice began, and which could only be seen after scrolling down.

Figure 3.14: Example of a social media feed (Mixed Compliance)



Empirical methodology

The results were again analysed with mixed model regressions. Separate regressions were run for each outcome variable: choice, whether a tweet was the first one clicked upon, whether a tweet was never clicked at all, the number of times a tweet was clicked on, the total duration spent viewing the product's web page, and the average duration of each visit to a product's web page. As in Experiment series 1, one set of models was

designed to estimate the overall effect of the different compliance conditions, and one set was designed to estimate the product specific effects.

Predictors for the overall effect models were dummies for:

- non-compliant condition, mixed compliance condition (with the compliant condition as the baseline category),
- compliant tweets in the mixed compliance condition,
- whether each tweet began on the screen,
- each product type (with pensions as the baseline category)
- 3 of the 4 different products within each product type (with the remaining used as the baseline category). This last set of dummies controls for the relative differences in attractiveness of the different products that are unrelated to compliance condition.

Note that for the outcome variables choice and first click, the product type dummies were not included. This is because each individual choice trial contained products from a single product type, and all trials necessarily had only one product chosen, and one product that was clicked first. Therefore there is no variation, and no possibility for product type-wise differences.

In models which estimated product-specific effects, interaction terms were also included between the product type dummies and the three compliance condition dummies (non-compliant, mixed compliant, and compliant tweet within a mixed compliance trial).

Results

Preferences (choice of product)

When looking at product preferences, a direct comparison between standalone compliance and non-compliance can only be made within the Mixed Compliance condition, since exactly one product must always be chosen. There was no significant effect of standalone compliance upon the likelihood of a product being selected (see Table 2B.1 in the Appendix).

Search (clicks and webpage time)

When looking at indicators of information search, most outcome measures allow for different baseline levels in different choice trials (e.g. different number of overall clicks per trial) and can therefore be compared between compliance conditions. Overall the results show that participants searched less extensively in the Compliant condition compared to both the Non-Compliant and Mixed Compliance conditions.

Compared to the Compliant condition, there were significantly more clicks in the Non-Compliant condition (11.0% relative increase) and the Mixed Compliance condition (7.8% increase; see Table 2B.2). In the compliant condition, it was significantly more likely that a product was never clicked on at all compared to both the non-compliant (14.6% more likely) and the mixed compliance (5.5% more likely) condition (see Table 2B.3). However, when looking at the compliance of individual tweets within the Mixed Compliance condition, there were no significant effects for either of these measures of search, nor was there an effect upon whether a tweet was the first one clicked upon (see Table 2B.4).

The effect upon web page viewing times was more equivocal. There was no significant effect of standalone compliance on the overall time spent viewing a product's webpage,

nor on the average time spent on the page per visit (see Tables 2B.5 and 2B.6). The only significant effect was that in the Mixed Compliance condition, participants spent less time viewing webpages of compliant tweets than non-compliant tweets.

Independent of compliance, the position on screen of content (which was randomised) had a significant effect on all outcome measures. Tweets which began on the screen and did not require the participant to scroll down to reveal them were more likely to be clicked first, clicked on more often, less likely never to be clicked, and participants spent longer looking at the webpage, both in total, and per visit. However, they were less likely to be chosen.

There were no obvious differences across different product types for either the Compliant vs Non-Compliant conditions or between compliant and non-compliant tweets in the Mixed Compliance condition.

Discussion

This experiment examined participants' behaviour when choosing between products of the same type, offered in a stream of tweets. Some of these results are strongly in line with previous literature, and intuitive expectations: such as the finding that tweets are more likely to be clicked and examined if they begin on the page and don't require any scrolling down. However, others are more surprising, and particularly relevant to our research questions. There was significantly reduced search behaviour in an environment where all tweets were standalone compliant, with participants examining and comparing fewer products before making a choice. This could potentially indicate a reduced tendency to shop around for the best available product, which could lead to worse consumer outcomes.

In an environment where only some tweets are standalone compliant, we find no significant difference in participants' likelihood of clicking on the advert in order to see the relevant web page. However, if the tweet was standalone compliant, slightly less time is spent examining the more detailed information on the webpage. This, again, could have negative consequences if the less detailed early warning reduces the attention that a consumer gives to later, more specific or directly relevant information.

Though these results are interesting, presenting products of only one type in each choice brings a potential weakness. This approach was taken for this first choice experiment as it allows for the strongest experimental control. However, it also means that all of the alternative products on display contained the same risk warning information. Therefore, this information did not distinguish between the products, and potentially implied to the participants that all of the products were more similar than they were. Since this is not representative of most real world situations, the next experiment used different product types within each social media feed.

2C: Four adverts, different product types

Background

In real life consumers may start a product search with a problem to be solved rather than a specific type of product in mind. In searching for suitable financial products, they will encounter adverts for a variety of different product types. Each of these will have its own associated risks, and therefore there will also be variety in risk warnings between

reasonable alternatives. In Experiment 2C, we investigated search activity and preference among products of different types when participants were given an explicit scenario and goal, and then presented with a number of different product types to choose between.

Method

Participants

A total of 300 participants were recruited. Nine participants were excluded from the analysis. Of the remaining 290 participants, 67% were female, 33% male, and <1% other. The mean age was 36.4 (S.D. = 11.1).

Procedure

At the beginning of each trial, participants were presented with a brief scenario and were told they should select the most appropriate product. Four options were then presented as part of a simulated social media stream, in the same manner as Experiment 2B. The four options were from different product categories, all of which were plausible products to fulfil a given scenario.

Each participant responded to a total of seven questions, from three different scenario types:

Borrowing: *"You need to borrow £(250 / 500 / 2,500 / 5,000) in the next 10 days and are unable to do so from friends and family. You should choose which of the available financial products/companies would be the best way to get the funds."* (Available product types were payday loans, equity-release, logbook loans, and guarantor loans.)[11]

House deposit: "You are saving for a deposit to buy a house and currently have £(1,000 / 5,000). You should choose which of the available products/companies would be the best way to invest it." (Available product types were CFDs, ETFs, crowdfunding, and mini-bonds.)

Windfall investment: *"You have had an unexpected windfall of £(1,500 / 15,000). You don't have anything you want to spend it on at the moment so you want to invest it to try and get the best return possible, or so it is most useful to you in the future. You should choose which of the available products/companies would be the best way to invest it."* (As there were 5 applicable product types for this scenario, pensions were always included and the remaining 3 were picked randomly from CFDs, ETFs, crowdfunding, and mini-bonds.)

For each participant, half the product types in borrowing scenarios, and half the product types in investment scenarios were randomly assigned as non-compliant, and the others as compliant. This meant that for an individual participant, the adverts for a given product type would always be standalone compliant (in every trial, for the entire duration of the experiment) or always non-compliant. But overall, the tweets in each scenario were always a mixture of compliant and non-compliant tweets (i.e. the Mixed Compliance condition from experiment 2B).

[11] It has since come to our attention that equity release may not be a suitable product for borrowing small amounts in the short-term. However, this was felt to be the most suitable of the possible categories for this product and we are confident that its inclusion does not affect the results.

Empirical methodology

The analysis strategy was similar to that used in Experiment 2B. The predictors included in the model estimating the overall effect of compliance were dummies for:

- house deposit scenario, windfall scenario (the borrowing scenario was used as the baseline category),
- whether the tweet began on the screen,
- whether the tweet was standalone compliant,
- each product type (pensions used as the baseline category),
- 3 of each of the products within each product type (with one of the 4 picked randomly to serve as the baseline category).

Results

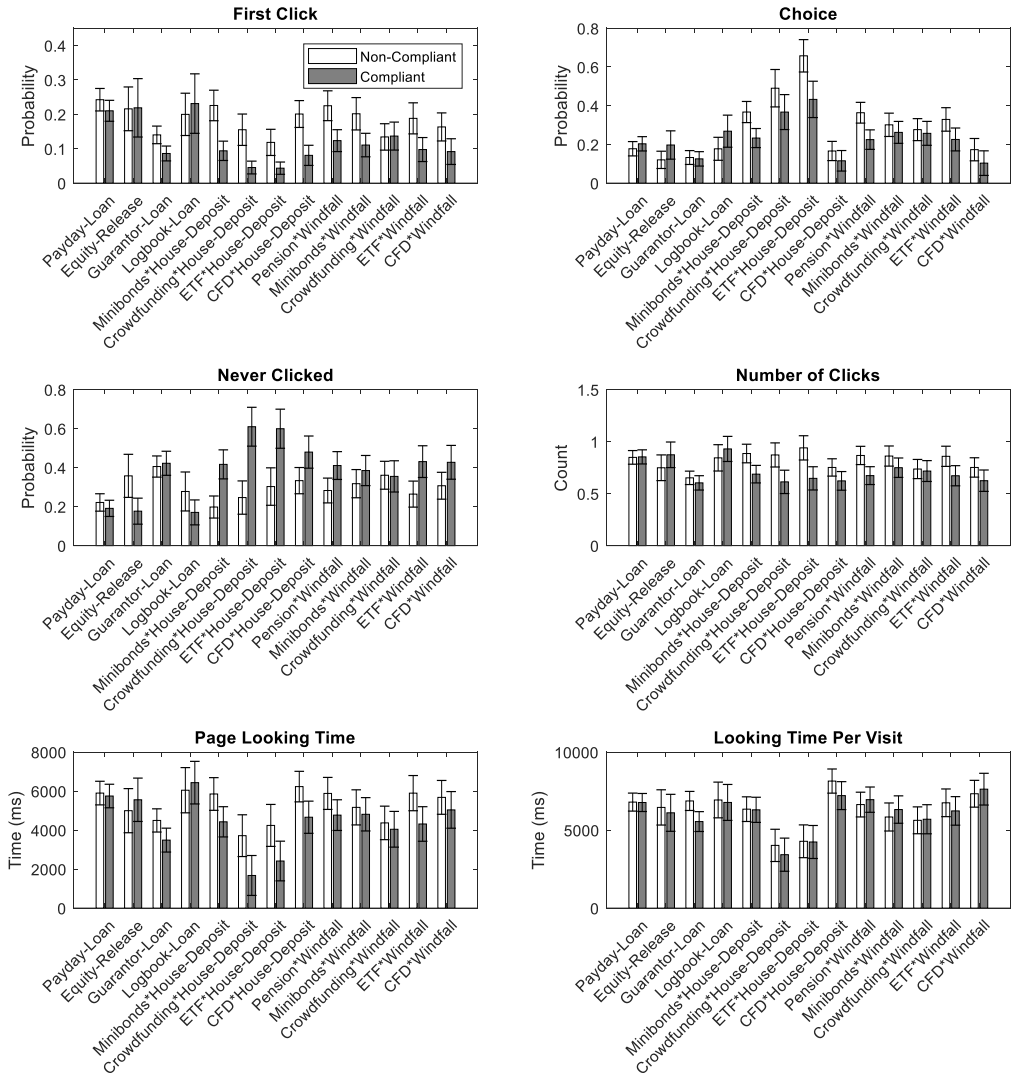
Preferences (choice of product)

Products advertised using standalone compliant tweets were significantly less likely to be chosen (a proportional reduction of 9.5% - see Table 2C.1 in the Appendix).

Search (clicks and webpage time)

Standalone compliant tweets were significantly less likely to be clicked on first (29.8% reduction relative to the baseline, Table 2C.2), received fewer clicks (9.5% reduction, Table 2C.3), and were significantly more likely to never be clicked on (relative increase of 16.3%, Table 2C.4). Participants also spent significantly less time on the webpages of products advertised using standalone compliant tweets. This was true for total looking time (a reduction of 13.9%, Table 2C.5) and time per visit (a reduction of 4.5%, Table 2C.6). Results for each outcome measure are shown separately by product type in Figure 3.15. There is some variance between product types, but there is no obvious pattern to these effects.

Figure 3.15: Effect of standalone compliance on all outcome measures by product



Discussion

When there are significant differences between the products being advertised (as would be expected in the real world), standalone compliance had a significant effect upon preference and information search. Products advertised through standalone compliant tweets were less likely to be chosen compared to non-compliant alternatives. Furthermore, standalone compliant tweets were less likely to be clicked on, and were clicked on later in the search. When they were clicked, participants then spent less time looking at the full information on the webpage.

One potential concern in this experiment is that the number of alternatives is relatively small: participants only saw four tweets. This makes it relatively easy to investigate all the options. As this would not be the case in the real world, we repeated this experiment with more tweets in Experiment 2D.

2D: Ten adverts, different product types

Background

In real life, participants would be likely to see a social media feed with more than four items and they would not usually be able to investigate all of the options in one place or sitting. In this experiment, we increased the number of adverts participants saw. This was intended to make the task harder and give us more reliable estimates of real-life search activity where there are more products than one individual can reasonably research and compare. The experiment was also designed to better examine the effect of standalone compliance upon risk comprehension, and how this interacts with patterns of behaviour. To provide a measure of understanding, at the end of the experiment, participants answered a composite set of comprehension questions from Experiment series 1.

Method

Participants

A total of 320 participants were recruited. Twelve participants were excluded from the analysis. Of the remaining 308 participants, 60% were female, 40% male, and <1% other. The mean age was 34.6 (S.D. = 10.8).

Procedure

This experiment followed the same procedure as Experiment 2C, but participants saw 10 tweet adverts instead of 4.

For the windfall scenario, where there were five relevant product types, there were two adverts from each product type included in the social media stream. For the other scenarios, where there were four relevant product types, two adverts each were shown from two of the relevant product types and three adverts each were shown from the remaining two product types. This meant that participants always saw 10 unique products in each social media stream.

Comprehension questions were included at the end of the task. These are a composite set of those used in Experiment series 1 and can be found in Annex 2 (Comprehension Question Set 3).

Empirical methodology

The analysis approach was similar to that used in Experiment 2C. However, a participant's mean accuracy for risk comprehension questions was calculated for each product type and used as a predictor in all analyses. There was an additional set of mixed model logistic regressions estimated to predict the probability of answering comprehension questions correctly. The first was estimated to predict the overall effect of standalone compliance upon comprehension. The predictors were a dummy variable to indicate whether that product had been displayed using tweets that were standalone compliant, and separate dummies for each comprehension question, to control for the different mean levels of accuracy for each question. The second model included process measures as a more exploratory approach. The predictors were those of the previous model, as well as the participant's total number of clicks on products of that type, their mean reaction time, the amount of time the participant spent on webpages for products of that type, the proportion of time overall that they spent looking at webpages vs the twitter feed, and interactions between these 4 new predictors and the compliance predictor.

Results

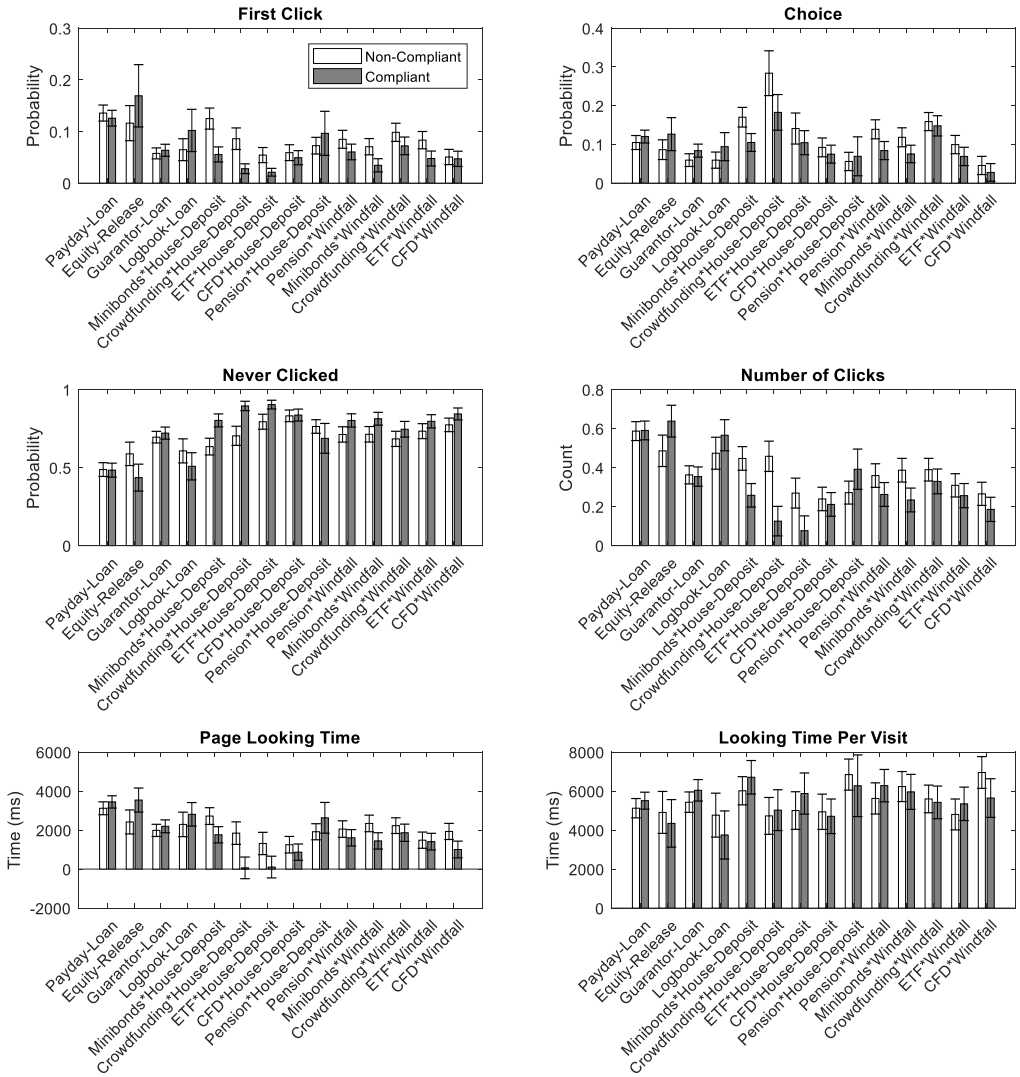
Preferences (choice of product)

Standalone compliant tweets were significantly less likely to be chosen (a relative reduction of 8.6% - see Table 2D.1), indicating a reduction in attractiveness.

Search (clicks and webpage time)

There was reduced search for standalone compliant tweets: they were significantly less likely to be clicked upon first (relative reduction of 19.8%, Table 2D.2), significantly more likely to never be clicked on (relative increase of 7.1%, Table 2D.3) and clicked on significantly less often (relative reduction of 14.2%, Table 2D.4). Participants also spent significantly less time looking at the webpages of standalone compliant adverts (relative decrease of 21.7%, Table 2D.5), but this appears to be driven by the number of times an advert is clicked (i.e. number of visits), as there was no effect of compliance on time per visit (Table 2D.6). There were some notable differences between separate product types, as shown in Figure 3.16. For example, standalone compliance had a particular effect upon the propensity of participants to consider crowdfunding and exchange traded funds as investments for a housing deposit.

Figure 3.16: Effect of standalone compliance on all outcome measures by product



Understanding (risk comprehension)

There was no significant effect of standalone compliance on risk comprehension. Participants were no more likely to answer comprehension questions accurately if the product had been shown with standalone compliant tweet adverts (Table 2D.7).

To identify potential patterns, some exploratory analyses were conducted. These revealed two small, but statistically significant effects. When tweets were standalone compliant, risk understanding was subsequently better if participants spent more of their time looking at websites than tweets, and if they spent longer overall before reaching a decision (Table 2D.8). In addition, we also report some exploratory analyses looking at the link between participants' demographic characteristics and their mean accuracy across all comprehension questions. Individual level correlations revealed that participants were more accurate if they were older ($r = 0.265$, $p<0.001$), and had indicated a higher level of familiarity with the product ($r = .265$, $p = <0.001$). A t-test

also showed that participants who had been living in the UK for a longer period of time also had higher mean accuracy (more than 10 years; $t(306) = 2.904$, $p = 0.004$).

Discussion

This experiment examined behaviour when the number of available alternative products was large, and some product types were standalone compliant, whilst others were not. There was reduced search for products that were standalone compliant, and participants were less likely to prefer them. However, there was no significant effect upon risk comprehension, as measured by the questions developed for Experiment series 1. The finding that participants were more accurate if they indicated greater familiarity with the product is not surprising. However, the finding that both age and length of time in the UK is predictive hints at a potential role for risk warnings in the wider world, to educate those with less exposure to, and experience of, such products.

Other exploratory analyses gave less clear indications of the interaction between comprehension and properties of information search. Only two aspects of the task predicted comprehension. The first of these is the interaction between reaction time and compliance. This is intuitive, as there is more information presented in the standalone compliant condition, and thus participants who take longer with the task are more likely to read it and improve their understanding. However, the other effect found a positive interaction effect between compliance and the proportion of time spent on the webpage. This is counterintuitive as standalone compliance only adds information to the social media feed, and yet this result suggests a beneficial effect of spending less time (proportionally) attending to that feed. Given the lower statistical strength of this result, it may well be due to chance. The subsequent experiments were designed specifically to make better sense of the interaction between comprehension, information search and preference simultaneously.

2E: Suitability in mixed compliance environments

Background

So far, the experiments have shown us whether non-compliance changes choices, but we have not been able to judge whether such changes are good or bad. This means that we cannot estimate potential harm from compliance or non-compliance. To investigate this, we created scenarios in which only one product was suitable for the circumstances given.

Method

Participants

A total of 361 participants were recruited, and 12 were excluded from the analysis. In this sample, there were 63% females and 37% males. The average age was 34.0 (SD 11.2).

Procedure

Participants saw eight scenarios, and had to choose between six products. Each scenario described a financial goal, such as borrowing or investing money, that the participant should aim to meet, as well as other relevant and irrelevant factors, such as their car ownership, mortgage status or risk preferences. The products were presented and searched in the same way as in Experiments 2B to 2D, with participants viewing a social

media feed and clicking through to see the product webpage (see Figure 3.17 for an example screenshot).

Figure 3.17: example screenshot (top part showing: scrollable)

Broadly, all six available products could potentially apply to the scenario (e.g. only investment products were offered when the scenario involved investing money), but the properties of the scenario and the risks for each product meant that only one product type was fully appropriate. Crucially, the information that determined whether a product was appropriate was, as far as possible, contained within the risk warnings, as shown in the example in Table 3.4 below.[12]

[12] Because of the real-world design of Experiments 2E and F, occasionally products were unsuitable for the given scenario due to features of the product itself, rather than information in the risk warning. In such cases, we rely on participants' rudimentary knowledge about how a product works. For example, in the scenario above, a pension would be an unsuitable product because it ties up money for too long, a fact that is not represented in the risk warning. We analysed the effects of removing the two scenarios which included examples of this and this made no difference to the overall results.

Table 3.4: Example scenario, possible answers and rationale for correct answer

"You are in your thirties and have just won £5,000 on premium bonds. You would like to re-invest it into a product which offers the possibility of high returns. You are mortgage-free and have savings and a stable income, so you could cope if you lost this money, but you would not want to lose more than this. You do not want to tie up the money for more than 5 years. Which of the following would you choose for your situation?"

(Shown to participants)	*(Not shown to participants)*	
Advert	***Right answer?***	***Why?***
Mini-bond 1 (higher return)	*Yes*	*Higher return and correct level of risk*
Mini-bond 2 (lower return)	*No*	*Correct level of risk but lower return*
CFD 1 (3-year)	*No*	*Could lose more than put in*
CFD 2 (5-year)	*No*	*Could lose more than put in*
Pension 1	*No*	*Ties up money for too long*
Pension 2	*No*	*Ties up money for too long*
Leave it in the bank	*No*	*Low returns*

Each social media feed included 3 product types, with two adverts for each and an option to do nothing, the "status quo option" (e.g. leave it in the bank). This meant that participants needed to choose the right product type and then the better product from the two available options. Each of the available product options were ranked, such that one was better than the other (for example, had a lower Annual Percentage Rate (APR), or a more suitable term length).

The stimuli and their presentation were the same as in Experiment 2C. As in Experiment 2C, all trials were Mixed Compliance, with product types randomly assigned as either standalone compliant or non-compliant for the duration of the experiment, for each participant.

Empirical methodology

Unlike earlier experiments, it was not feasible to estimate product specific effects. This is because the large number of combinations of scenario and product types make such an analysis impractical. Instead emphasis was placed upon designing a suitably varied and balanced array of scenarios, and the analysis approach focuses upon the overall effect of standalone compliance.

The first group of models predict aspects of the choice outcome. The outcome measures are whether a product of the appropriate type was chosen, and whether the best overall product was chosen (the product of the correct type that had e.g. the best interest rate). The predictors were whether the correct product type was standalone compliant, the number of alternative products that were standalone compliant in that choice trial, a dummy indicating if at least one product of the appropriate type began on the screen (prior to scrolling), a dummy indicating if both of the products of the appropriate type

began on the screen, dummies for each scenario to control for differing levels of accuracy between scenarios, and an interaction term between number of alternative products that were standalone compliant, and, finally, whether the appropriate choice for the scenario was the status quo.

The second group of models predict process measures and aspects of information search. The predictors were the same as those for the models predicting aspects of choice, but in addition contained a dummy indicating whether a product was of the appropriate product type, a dummy indicating whether it was the best product within its own product type, and an interaction between these two dummies to identify the product that was the best overall. In addition, there were interaction terms between these three predictors and whether the product was standalone compliant.

Finally, some more exploratory analyses were conducted examining the relationship between process measures and choice outcome. For these, the outcome measures were whether the correct product type was chosen, and whether the best overall product was chosen. The predictors were those used in the models above that predicted choice outcome, but in addition were predictors for the reaction time on the trial, the total number of clicks on the trial, the number of products never clicked on, the ratio of time spent looking at the webpages vs the twitter feed, and interactions between compliance and each of these predictors.

Results

Understanding (suitable product chosen)

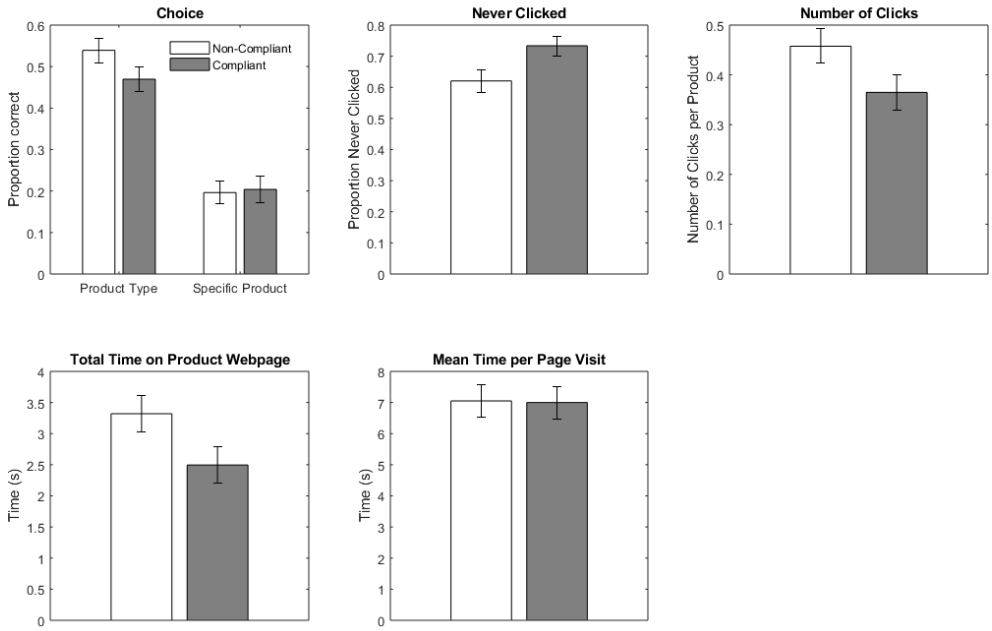
Around half of the participants overall were able to choose the most suitable product type (48.2%), and fewer were able to select the best single option available (39.1%).

If the correct product type was standalone compliant, participants were significantly less likely to choose it (relative reduction of 7.0% - see Table 2E.1 in the Appendix). If predicting instead whether participants selected the best overall product (i.e. the appropriate type, and the highest ranked), then there was no significant effect of standalone compliance (see Table 2E.2). This shows that standalone compliance affects participants' likelihood of choosing the most suitable product type.

Search (clicks and webpage time)

Standalone compliance reduced search in a similar way to earlier experiments: compliant tweets were less likely to be clicked on first (relative reduction of 29.4%, Table 2E.3), more likely to never be clicked on (relative increase 11.3%, Table 2E.4), received fewer clicks (relative reduction of 19.0%, Table 2E.5), and less time was spent looking at the webpages of compliant tweets (relative reduction 21.7%, Table 2E.6). However, there was no effect on looking time per visit (Table 2E.7). See Figure 3.17 for a summary.

Figure 3.17: Effect of standalone compliance on all outcome measures



One of the most interesting questions is whether standalone compliance specifically changes how information search is directed towards the correct product types. For example, if participants clicked fewer products overall then this may be considered a negative outcome as it reduces the information gathered and the degree of shopping around. However, if the reduction in overall number of clicks were due to participants reliably clicking only on suitable products, this may be considered a positive outcome. It is possible that the additional risk warning information in standalone compliant tweets may enable participants to prioritise search for the correct product type more easily. This is not observed in our results. When comparing search towards the correct tweet when it was compliant vs non-compliant, we find that the reduction in search is not as large as compared to the average across all products. But the overall effect is still negative for most outcome measures. For correct products, compliant tweets were less likely to be clicked on first (relative reduction of 20.4%) and more likely never to be clicked on (relative increase of 9.1%). There was no significant difference between compliant and non-compliant tweets for number of clicks, total time spent looking at webpages and time per visit.

To examine what properties of information search may affect the likelihood of selecting an appropriate product type, some exploratory analyses were conducted. These show that the most relevant aspect of information search is the number of clicks, with more clicks being related to an increased chance of selecting an appropriate product type. There is also a smaller effect of how the participant spends their time; spending a higher proportion of time looking at the social media feed versus webpages is associated with making more appropriate choices (Table 2E.8). Surprisingly, there is a small negative effect of overall response time: the longer spent deliberating on a trial, the less likely it was that the most suitable product type was chosen. When looking at the likelihood of choosing the most suitable overall product (i.e. the better ranked product within the appropriate product type), the number of clicks made is a significantly stronger predictor,

but the proportion of time spent on the social media feed versus webpages is not significant.

Discussion

If the most suitable product is standalone compliant, in this task environment participants are less likely to identify it and select it as the most suitable response for the scenario. The results show that reduced search activity plays a role in this, with less searching aimed at standalone compliant products, even when they are the most suitable. This is not necessarily causal. Searching less may lead to poorer choices, most likely because of reduced understanding. However, it could be that good test-takers, or those with better prior knowledge are more likely to expend more effort on the task, searching more and longer. This is complicated by the finding that longer response times are associated with worse accuracy. This suggests that the most successful participants were those who switched quickly between information on multiple products, and did not spend long periods of time deliberating on a small number of product webpages.

This task examined a Mixed Compliance environment where some products are standalone compliant and some are not. It means that much of the difference in search and choice could be driven by participants preferring non-compliant adverts to compliant ones. This means their behaviour could be very different in an environment where there are no such differences, and everything is either standalone compliant, or non-compliant. It is also possible that participants could use the information in the risk warnings of standalone compliant tweets and extrapolate this to non-compliant tweets. This might make them better at identifying the most suitable product type in a Mixed Compliance environment, than they would be when no tweets are compliant. (However, note we found no effect of the number of other products in a trial which were compliant.) Experiment 2F compares a non-compliant environment with a fully compliant one, which allows us to investigate these hypotheses further.

It is also notable that only around half of the participants overall were able to choose the most suitable product type (48.2%), and fewer were able to select the best single option available (39.1%). In fact, in many ways, the experiment environment is likely to be much more conducive to good decisions than real life: participants are paying attention to the task and more educated than the UK average (Stewart, Chandler & Paolacci, 2017), webpages are simplified and the consumer journey is easy, there are a limited number of options and no extraneous social media posts to distract attention. Notwithstanding other opportunities to get information, this means that decisions in real life are likely to be even worse and suggests that substantial harm could come from poor decisions made on the basis of social media adverts. However, this must be balanced against the fact that this experiment was hypothetical and participants did not benefit financially from making a better decision as they would in the real world.

2F: Suitability in standalone and non-compliant environments

Background

In Experiment 2E we showed that participants were less likely to select the most appropriate type of product if the tweet was standalone compliant. However, this was within a Mixed Compliance environment. Therefore, if an individual prefers non-compliant over compliant adverts, or sees the presence of a risk warning in only some products as

a signal, then they would always have the opportunity to select a non-compliant product. Although this is an important test of how standalone compliance affects information search, this mixed compliance situation is unlikely to exist in the real world, at least not for comparable consumer products that could conceivably be relevant to similar situations. The FCA's current policy is full standalone compliance. Therefore, this study is a between-subjects experiment comparing environments with full standalone compliance and no compliance. That is, each participant completed the task in an experiment universe which was either fully standalone compliant, or entirely non-compliant.

Method

Participants

A total of 412 participants were recruited, and 12 were excluded from the analysis. There were 58% females, 42% males and <1% other. The average age was 34.1 (SD 12.4).

Procedure

The procedure was the same as Experiment 2E, and used the same scenarios. The only difference was that participants were randomly assigned to one of two conditions: compliant or non-compliant. This meant that throughout the entire experiment, each participant either saw adverts that were all standalone compliant, or all non-compliant.

Empirical methodology

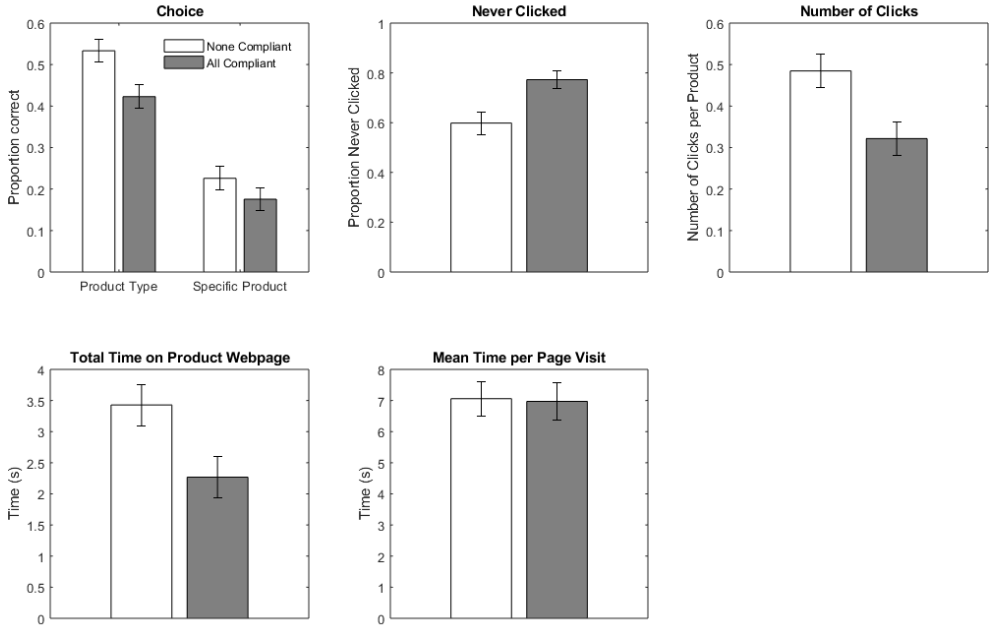
The analysis approach was the same as that in experiment 2E. The only difference is that since the compliance manipulation is now between-subject, not within, participant level random effects could not be estimated for this predictor.

Results

Understanding (suitable product chosen)

We begin by examining the effect of standalone compliance upon participants' accuracy in identifying the most suitable product type for each scenario. The results show that, after controlling for other variables, participants are significantly less likely to select the most suitable product type when adverts are standalone compliant (a relative reduction of 14.0%. See Table 2F.1 in the Appendix). They are also significantly worse at identifying the best overall product (the better ranked product with the appropriate product type; a relative reduction of 22.8%, Table 2F.2). See Figure 3.18 for a summary.

Figure 3.18: Effect of standalone compliance on all outcome measures[13]



Search (clicks and webpage time)

This experiment echoes findings in Experiments 2D and E. Overall, compliance was associated with an increase in the likelihood of a tweet never being clicked (relative increase of 22.9%, Table 2F.3), a reduction in the number of clicks (relative reduction 40.3%, Table 2F.4), and a decrease in the total time spent looking at webpage information (relative decrease of 41.6%, Table 2F.5).

Again, suitable product types were more likely to be clicked first compared to unsuitable product types (relative increase of 109.9%), less likely never to be clicked (relative reduction of 24.4%), clicked upon more often (relative increase of 44.4%) and longer was spent looking at their webpages (relative increase of 97.6%). This indicates that participants were more likely to prioritise their search towards the most plausible options.

When looking at the interaction between compliance and whether an advert was for the correct type of product, there was a small but statistically significant increase in the proportion of tweets never clicked (relative increase of 8.0%), and a reduction in the number of clicks (relative decrease of 17.6%), but there was no significant effect on whether the correct product type was clicked on first, or on the time spent looking at webpages. There was no effect on average looking time per visit for any measures (see Table 2F.6).

The exploratory analyses from Experiment 2E were also performed here to identify the potential effects of information search patterns upon choice accuracy. These largely replicate the results from the mixed environment of 2E, and generally with larger effect sizes.

[13] *Note that since this is a between-subject manipulation, the differences between Non-compliant and All compliant are not a relative preference within an individual for non-compliant products. Instead, this demonstrates an absolute reduction in choice accuracy and information search when in a standalone compliant universe compared to a non-compliant universe.*

The lack of response time effect was the most notable difference. A greater number of clicks during a trial is associated with a greater chance of selecting the most suitable product type, as is spending proportionally more time looking at the social media feed compared to webpages (see Table 2F.7). These effects are substantially stronger in the standalone compliant environment. In fact, in this analysis, these process predictors explain such a large proportion of the variance in choice that standalone compliance itself is no longer a significant predictor. When predicting the likelihood of selecting the most suitable overall product, the number of clicks is still a significant predictor, and this effect is larger within the standalone compliant environment (see Table 2F.8). However, time spent looking at tweets versus webpages is no longer significant, and the main effect of the compliance condition is still a weak predictor, with standalone compliance being related to a lower chance of selecting the best product in this task.

Discussion

The results show that within this task, participants are less accurate at selecting the correct product type when in a standalone compliant environment compared to a non-compliant environment. The primary cause is reduced information search in the standalone compliant environment, with fewer product clicks (a proxy for the number of times a different product is investigated and considered) being the strongest predictor of sub-optimal choices.

Discussion: Timing of risk warnings

In the context of the simplified social media stream we developed here, standalone compliance reduces search activity and results in lower understanding which causes poorer decisions. Participants also find compliant adverts less attractive and pay less attention to them (though note some product specific effects).

One explanation is that the presence of risk warnings and negative information in tweets causes participants to lower their valuation or perception of the product, making them less likely to explore it further. This means they click on these adverts less and then are less likely to read the more detailed information on the product's webpage. This makes them less likely to get the information necessary to understand the product, buy it or choose a suitable alternative. This is supported by the finding in Experiment 2F that participants were more likely to choose the status quo options in a standalone compliant environment than a non-compliant environment (20.8% versus 16.8% of trials).

Of course, if the advertised product is risky, unsuitable, or poor value, this outcome could be a good one for consumers. However, in real life consumers may be unable or unwilling to take a status quo option, and often these products are beneficial to consumers despite their associated risks (so long as these are properly understood). Therefore, anything that reduces their information search before purchase may result in poorer outcomes.

Since compliant adverts are less attractive than non-compliant ones, it can be argued that all adverts should be either compliant or non-compliant to avoid some companies having an unfair advantage. However, our results show that these effects are not simply driven by a relative preference for non-compliant adverts at the expense of considering compliant ones. Even in a fully-compliant environment, participants were still less likely

to search for more information and ultimately selected less suitable products when compared to behaviour in a fully non-compliant environment.

The webpage on which the product can be purchased plays an important part in consumers' understanding of the risks and the likelihood they will make a good decision. The results in this second series of experiments support the results in Experiment series 1 that demonstrated an effect of webpage risk warnings upon understanding, but not tweet risk warnings. In Experiment 2D, participants performed no worse on risk comprehension questions when there were no tweet warnings presented. Further, in Experiments 2E and 2F, participants' choices demonstrated worse understanding when these tweet warnings were included. Because understanding in these situations seems to be driven primarily by the webpage risk warnings, it is important that this information is clear, fair and not misleading, just as other adverts are required to be.

To ensure good control of the experimental stimuli, there were aspects of our task environment which had to be simplified. For example, the product webpages used no images or significant colour contrast and primarily contained short, bullet-pointed information. Obviously this does not fully represent the variety of pages present in the real world. Therefore, improving understanding of the properties of webpage design and risk warning presentation that most affect consumers' ability to extract and understand the associated risks could be useful.

Another aspect of the design worth highlighting was that the length of the tweets (and therefore their size on the screen) could not be standardised across conditions. Standalone compliant tweets were always longer because they included information that the non-compliant tweets did not. This could mean that length, rather than standalone compliance is responsible for the effects we see. For example, participants might click on fewer tweets because when tweets were longer, fewer initially appeared on the screen in the tweet feed. Viewing the same number of tweets within the feed required additional scrolling when tweets were longer. We explored ways to control for length, but this would involve either introducing new information or repeating old information in the non-compliant tweets only and this would also make it difficult for us to disentangle causation. Ultimately, given that participants did frequently explore and select tweets that required scrolling, and that the statistical analyses included this as an explicit control, this is unlikely to be the cause of these results. Even if it were found to play a role, it would still be of interest, as the same pattern will be found in the real world: standalone compliant tweets will tend to be longer.

4 Discussion, implications and conclusions

We examined the effect of standalone compliance (SC) on three key aspects of consumer choice: preference, information search, and understanding. The results show that variations in the **design** and **timing** of risk warnings can affect them all. In particular, we find that standalone compliance in a character-limited social media environment reduces preference ratings, search activity (such as shopping around), and understanding of risks (as evidenced by the choice of less suitable products).

Design of risk warnings

Modifying risk warnings has a modest effect on individuals' understanding of the associated risks and their preferences for and against products:

- Writing risk warnings using insights from behavioural science and objective readability metrics did improve understanding, but only if they appeared on product webpages (not on social media posts).
- The inclusion of pictures within the tweet risk warning increased preference for the product, but did not affect understanding.
- Labelling tweet risk warnings with 'FCA Mandated Warning' decreased understanding and did not affect preferences.

Discussion

Our findings on risk warning **design** emphasise their importance for financial products, but also demonstrate the difficulty in designing appropriate frameworks and maximising their effectiveness. For example, the inclusion of images and logos from the Money Advice Service (MAS) within the risk warnings increased participants' ratings of the products, even though the risk warnings were still imparting objectively negative information. However, the cause of this effect is not entirely clear. Participants' comprehension was not affected, so the image does not appear to be distracting attention away from the risk information.

The explicit labelling of a risk warning as being 'FCA Mandated', rather than attracting more attention, actually resulted in reduced understanding. These results support the existing practice of not referencing the regulator, as this appears to be a distraction, or worse, is seen as an endorsement of the product.

Perhaps not surprisingly, when risk warning text was redeveloped using insights from behavioural science and with language that scored highly on objective measures of readability (including concreteness and word frequency), participants' understanding improved (though this effect was only found for the more detailed risk warnings presented on webpages).

Implications

The results suggest that advertisers and consumers could benefit from experimentation on the design and presentation of risk warnings. However, incentives may not be

aligned: firms which put considerable effort into designing risk warnings may be at a competitive disadvantage if others do not communicate their risks so clearly (though they may benefit from a better long-term relationship with customers and potential customers). There may also be an argument for experimenting with other methods of improving consumer understanding, which go beyond risk warnings and providing information.

Timing of risk warnings (standalone compliance)

We find a strong and consistent effect of standalone compliance on consumer preference, search activity and understanding.

- For **preference**, we find that individuals have a more positive impression of adverts which are not standalone compliant. These receive higher ratings and are more likely to be chosen.
- For **search activity**, we find that individuals are less likely to search for additional information about standalone compliant tweets. In mixed compliance environments (i.e. where some adverts are compliant and some are not), search is biased towards non-compliant products. In standalone compliant environments compared with non-compliant environments, there is less search in general, with fewer comparisons between products.
- For **understanding**, we find that standalone compliant tweets mean that participants are less likely to identify suitable products for a given scenario. In mixed compliance environments participants are less likely to choose the most suitable product if it is standalone compliant, and in fully standalone compliant environments, participants are less likely to choose the most suitable product in general. Standalone compliance did not appear to affect participants' accuracy in answering risk comprehension questions.

Discussion

The results from these experiments are consistent: within the specific task environment tested, standalone compliance has a negative impact on participants' understanding of associated risks and it increases the likelihood of individuals selecting less suitable financial products.

Of course, this is only within the task structure we investigated. There are several properties of real-world search that could not be fully replicated. The design choices and trade-offs made for these experiments were deliberate and were designed to allow sufficient control within the experiment. For example, there were far fewer alternatives than in the real world, options were presented with equal prominence, and all had the same low cost of acquiring additional information. In the real world, additional search may be more difficult and less productive. There is rarely a simply presented list of all the available options, and heavily promoted alternatives may dominate search attempts. In this kind of environment standalone compliance may have quite different effects upon search.

The task employed here also forces participants to choose immediately. In real-world decisions, search and deliberation often occur across several sessions, and individuals may delay choice until a better alternative becomes available. Some of our experiments

did give a status quo option, but the one-shot choice environment implies a permanence to that choice which would not exist in the real world.

Another important consideration is that the tweets and webpages used were simplistic. They contained few images or emotional content, and did not contain any familiar brands. Their main content was factual information about the product, and the risk warnings formed a large proportion of the webpage text. This is rarely the case in the real world, particularly on product webpages. However, given how successful these messages were, particularly the comparatively large (and positive) impact of webpage risk warnings against the (potentially even negative) effects of tweeted risk warnings, it is crucial to understand the properties of these risk warnings that make them most effective. Future research could identify the properties that maximise the effectiveness of risk warnings on product landing pages, and webpages more generally. This could include positioning, size, language, inclusion of pictures, consistency (or repeated novelty) of appearance and formatting, persistence on the page throughout the purchasing process, to name but a few. Better understanding of these properties could help to develop policies which protect and inform consumers.

There are two primary hypotheses for why standalone compliance reduced search and resulted in less appropriate choices and each have different implications. One is that risk warnings make all products seem like poor options. Individuals do not want to waste time examining poor options, so they reduce search activity. The second is that as risk warnings are very similar on all products of the same type, this could make all products seem very similar. There is little point expending time and energy on search if one expects all options to be similar.

Both of these hypotheses have their merits. Although our data does not strongly support one over the other, there are suggestions that the former is more likely. This is because ratings of individual products were lower when they were presented as standalone compliant, and participants were more likely to select the status quo option when in a standalone compliant environment (essentially rejecting all options). But this does not rule out the effects of similarity, and future research could attempt to identify the relative importance of these hypotheses to encourage more information search, product comparison, and ultimately more competition.

Finally, a remaining question is what role standalone compliance might play in educating the consumer when they are not actively looking for products. It is rare for members of the public to receive any explicit education in specific financial products, and yet there is an existing baseline of knowledge. Our results show that some of this is related to an individual's self-reported familiarity with the type of product, suggesting that those who have dealt with such products understand the associated risks better. However, there are also effects for age, and length of time living in the UK. Given the prevalence of adverts for financial products, and thus the associated risk warnings, it is possible that individuals are educated by this repeated exposure.

It seems unlikely that the relatively brief and often unspecific risk warnings presented on tweets are having a significant educational effect. It is perhaps more intuitively plausible that individuals simply engage with more products as they age, and by doing so would see detailed risk warnings on web pages and other information sources, as they progress through the purchasing process. However, we have insufficient data to answer this

definitively, and it may be helpful to consider the source(s) of consumers' knowledge when developing policy.

Implications

If products were to be advertised without risks warnings on social media posts, the results indicate that there would be no detectable decrease in consumers' risk comprehension, provided suitable risk warning information is prominently presented on product webpages. However, it is still not known if passive exposure to these messages can educate consumers even if they are not currently actively considering such a product.

Given the specificity of the scenarios examined here, the results cannot be assumed to extend to other types of media or advertisement. There are many properties of short, online and social media adverts that are unique, and will likely result in very different effects of standalone compliance, for example, the immediacy of clicking on the advert and seeing risk warnings on the webpage, or the ease of accessing more information to compare across multiple products at once. We also do not know the effects (if any) of Twitter's recent character limit change to 280 per tweet.

We identify two potential underlying causes for the reduced information search (though both could be acting simultaneously). The implications for each are slightly different. If the primary mechanism of action is a reduction in overall attractiveness for all options (and our results suggest this is a significant component), then little can be done to change behaviour through a change in risk warnings. The alternative explanation is that risk warnings give the false impression that all products are the same, and thus the consumer believes that there is no point in shopping around. If this were true, then some of the negative effects could be ameliorated by including more specific risk warnings that are tailored to each individual product.

One may consider that for some types of products, the risks are substantially greater, and therefore standalone compliance should remain for these types of products. Based upon the results here, this would likely reduce consumers' interest in these products compared to the less risky products which may no longer be standalone compliant. Though this may help to dissuade uninformed consumers from considering such products, we have little evidence one way or the other. However, one implication of the results is that should a consumer then make a decision which means they are only considering products that are still required to be standalone compliant, it is likely that they will then search these products less, and acquire less information.

Conclusions

The results show that in a simulated social media environment, standalone compliance in adverts reduces individuals' information search and leads to less suitable product choices. Compared to an environment where risk warnings are presented on the product webpage but not on the social media advert, individuals explored fewer alternative products and were more likely to make product choices that were unsuitable for a given scenario.

The results also highlight the relative importance of risk warnings on a product's webpage. They show that individuals' understanding of the associated risks can be improved when risk warnings are formulated using insights from behavioural science and

readability metrics. Also, we highlight that the effects of risk warnings can be skewed by otherwise irrelevant presentational choices such as the inclusion of advice service logos, or explicitly referencing that the risk warning is FCA mandated.

There is potential to incorporate these findings into policy, but this must be done with caution and monitoring of any effects.

Annex 1: References

Adams, P., & Smart, L. (2017). From advert to action: behavioural insights into the advertising of financial products. *Financial Conduct Authority Occasional Paper 26*.

Benartzi, S., & Lehrer, J. (2015). *The smarter screen: surprising ways to influence and improve online behavior*. Hachette UK.

Cox, R., & de Goeij, P. (2016). The Effects of Risk Framing, Balanced Information and a Regulatory Seal of Approval on Investor Perception and Behavior. *Working paper.*

Edwards, D., & Potter, J. (1992). *Inquiries in social construction. Discursive Psychology.* Thousand Oaks, CA, US: Sage Publications, Inc.

Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*, *60*(8), 1861–1883.

Financial Conduct Authority. (2015). Social media and customer communications*. Finalised Guidance FG15/4.*

Financial Conduct Authority. (2016). Smarter Consumer Communications, *Feedback Statement FS16/10.*

Fong, G. T., Hammond, D., & Hitchman, S. C. (2009). The impact of pictures on the effectiveness of tobacco warnings. *Bulletin of the World Health Organization, 87*, 640-643.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*(4), 650–669. https://doi.org/10.1037/0033-295X.103.4.650.

Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information avoidance. *Journal of Economic Literature, 55*(1), 96-135.Hammond, D., Fong, G. T., McNeill, A., Borland, R., & Cummings, K. M. (2006). Effectiveness of cigarette warning labels in informing smokers about the risks of smoking: findings from the International Tobacco Control (ITC) Four Country Survey. *Tobacco control, 15*(3), iii19-iii25.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature, 42*, 1009-1055.

Henkel, L. A., & Mattson, M. E. (2011). Reading is believing: the truth effect and source credibility. *Consciousness and cognition, 20*(4), 1705-1721.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: conducting experiments in a real labor market. Experimental Economics, 14(3), 399–425.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of*

Personality and Social Psychology, *80*(4), 557–571. https://doi.org/10.1037/0022-3514.80.4.557

Kahneman, D., & Tversky, A. (1982). The Psychology of Preferences. *Scientific American*, *246*(1), 160–173. https://doi.org/10.1038/scientificamerican0182-160

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. https://doi.org/10.1038/nn.2635.

Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. Cambridge University Press.

Mullinix, K., Leeper, T., Druckman, J., & Freese, J. (2015). The Generalizability of Survey Experiments. Journal of Experimental Political Science, 2(2), 109-138. doi:10.1017/XPS.2015.19

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: Some foundational issues. *Learning and Individual Differences*, *7*(2), 145–162. https://doi.org/10.1016/1041-6080(95)90028-4

Schotter, E. R., Berry, R. W., McKenzie, C. R. M., & Rayner, K. (2010). Gaze bias: Selective encoding and liking effects. *Visual Cognition*, *18*(8), 1113–1132. https://doi.org/10.1080/13506281003668900

Searleman, A., & Carter, H. (1988). The effectiveness of different types of pragmatic implications found in commercials to mislead subjects. *Applied Cognitive Psychology, 2*(4), 265-272.

Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing Samples in Cognitive Science. *Trends in Cognitive Sciences*, *21*(10), 736–748. https://doi.org/10.1016/j.tics.2017.06.007

Annex 2: Comprehension questions

Comprehension Question Set 1 (1A-C)

Which of the following statements about mini-bonds is most true? Choose one answer.

If the firm fails completely, your initial investment is protected

It is possible to lose all of your initial investment

If the firm fails to make any promised interest payments then you can demand they refund your initial investment

Don't know

Which of the following statements about mini-bonds is most true? Choose one answer.

It is possible you will not see any money in the form of interest payments

You don't need to hold them for the full term to get your initial investment back

You can sell the bond to someone else if you want to (ie they are transferrable)

Don't know

Which of the following statements about car insurance is most true? Choose one answer.

If a company or broker advertises an average cover price then you will be offered at least one policy at that price

Insurance companies reward loyalty, so auto-renewal car insurance will always provide effective, value for money cover

A company or broker can advertise an example cover price even if only a small fraction of customers get such a good offer

Don't know

Which of the following statements about car insurance is most true? Choose one answer.

The firm will pay out for any claims you make if there is incorrect information in your application

The amount the insurer pays for your car accident will depend on the level of cover and excess you choose

Your age and occupation will not affect your car insurance quote

Don’t know

Which of the following statements about Contract for Difference products is most true? Choose one answer.

You can never end up with less money than you originally invested

You can lose more than your initial investment

It is possible you may lose some of your original investment, but you are unlikely to lose all of it

Don’t know

Which of the following statements about Contract for Difference products is most true? Choose one answer.

You’re likely to make money, they are a low risk investment product

About equal proportions of people make and lose money, they are a somewhat risky investment product

You’re likely to lose money, they are a high risk investment product

Don’t know

Which of the following statements about investment-based crowdfunding is most true? Choose one answer.

You are unlikely to receive less money on exit than originally invested

It is likely income in the form of dividends will be paid

If the company you invest in defaults or there is fraud your money will not be repaid and/or dividends will not be paid

Don’t know

Which of the following statements about investment-based crowdfunding is most true? Choose one answer.

It is very likely you will lose all your money, as most start-up businesses fail

Most platforms have a way you can cash in your investment

If the crowd-funding platform goes bust the return is still guaranteed

Don’t know

Which of the following statements about loan-based crowdfunding is most true? Choose one answer.

You can reasonably quickly sell/refund your shares for the amount you originally invested

You cannot lose all of the money you invest

It is a higher long-term risk than keeping money in tracker funds

Don’t know

Which of the following statements about loan-based crowdfunding is most true? Choose one answer.

There is generally no formal, legal mechanism for reclaiming funds from a borrower after a loan has defaulted

You can cash your investment in for as much money as you originally paid

You are always guaranteed some compensation should the firm fail but this is likely to be much less than you invested

Don’t know

Which of the following statements about equity release is most true? Choose one answer.

It tends to be less expensive than an ordinary mortgage

It can reduce the value of your estate (inheritance you leave)

It still allows you to use the value of your property to cover the costs of long term care in the future

Don’t know

Which of the following statements about equity release is most true? Choose one answer.

It may affect your entitlement to state benefits

It will not reduce the value of your estate

Your family will inherit the entire property but only following your death

Don’t know

Which of the following statements about ETFs (exchange traded funds) is most true? Choose one answer.

If the market starts to move against you, you can always withdraw your investment before the ETF begins to lose value

You directly own and have direct claim to the underlying investments in the fund

You may get back less than you invest

Don’t know

Which of the following statements about ETFs (Exchange Traded Funds) is most true? Choose one answer.

They are highly susceptible to market risk as you alone must react if the market goes down

They are actively managed to reduce susceptibility to market risk

They are not susceptible to short term market risk

Don’t know

Which of the following statements about Guarantor Loans is most true? Choose one answer.

The loan provider can only take you to court, your guarantor must help them find you

The loan provider can only take your guarantor to court as they are the senior signatory

The loan provider can take both you and your guarantor to court

Don’t know

Which of the following statements about Guarantor Loans is most true? Choose one answer.

The guarantor can be anyone

If you fail to repay then the guarantor must repay the capital, but cannot be forced to pay the interest

If you miss repayments the guarantor must help the loan company to communicate with you and provide access to you/your belongings

Don’t know

Which of the following statements about Logbook loans is most true? Choose one answer.

If you fail to make repayments the lender can reclaim any of your belongings to repay the loan

If you fail to make repayments then the lender can take the car used to get the loan

If you fail to make repayments then the lender can take ownership of any cars you own

Don’t know

Which of the following statements about Logbook loans is most true? Choose one answer.

The lender owns the car until you have made the final payment and the lender can take your car if you fail to make repayments

You still own your car, but have to make repayments to the lender

In the event that you are not willing to repay your loan, the lender can prevent you from legally owning other cars

Don’t know

Which of the following statements about Payday Loans is most true? Choose one answer.

The amount you have to pay back will be high: up to double the amount you borrow in the first place

The amount you have to pay back will be equal to the amount you borrow in the first place

The cost of payday loans is similar to that of other short-term credit such as arranged overdrafts

Don’t know

Which of the following statements about Payday Loans is most true? Choose one answer.

Taking out a payday loan will not affect your ability to get credit in the future (eg a mortgage)

Companies are not allowed to take past payday loans into account when deciding whether to lend to you

Late repayment can cause you serious money problems and make it difficult to borrow in the future

Don’t know

Which of the following statements about pension investments is most true? Choose one answer.

Pension investments are protected by the government so you will never get back less than you pay in

It is possible to get back less than you pay in

Pensions are always invested in low risk funds

Don’t know

Which of the following statements about pension investments is most true? Choose one answer.

The value of your pension may vary over time

The value of your pension will only increase over time

You can cash in your pension at any time

Don’t know

Comprehension Question Set 2 (1A-Replication)

The company you have bought mini-bonds from has been made bankrupt. Which of the following is likely to be the case?

You can apply to get your money back from the Financial Services Compensation Scheme but you may not be successful

You are likely to lose your money because the Financial Services Compensation Scheme does not apply

You may be able to transfer your investment to another company

Don't know

You invested in a pension twenty years ago and are now coming up to retirement. What should you expect from your pension?

Your total pension will be greater than the amount you put in

Your total pension will be smaller than the amount you put in

Your total pension could be smaller or greater than the amount you put in

Don't know

You start trading on a Contracts for Difference (CFD) platform with £100, but find that you make a lot of mistakes. By the end of the day, the smallest amount you could have in your account is:

£0

-£200 (i.e. you owe the provider double your investment)

You could owe the provider an unlimited amount

Don't know

You are retired and want to release money from your home to pay for home improvements. You don't receive any government benefits and you intend to leave your house to your children when you die. Why might a lifetime mortgage (a type of equity release) not be suitable for you?

Because you would not retain ownership of your home

Because it may affect the inheritance you are able to leave

Because it is a product intended for those who receive government benefits

Don't know

You see an advert offering car insurance "from £200". What percentage of customers are eligible to buy insurance costing £200?

At least 90%

At least 50%

At least 10%

Don't know

You miss a number of payments on your logbook loan. Which of the following could happen?

Your car could be taken away, as you do not own it until you have made the final payment

Another company may make contact with you to recover the money owed, since the company you got the loan from only arranged the loan, rather than provided the money.

Both of the above

Don't know

Which of the following is true about Exchange traded funds (ETFs)?

ETFs don't give you access to the whole market

Your original investment might go down

Both of the above

Don't know

You are a student and take out a guarantor loan to pay for travel abroad. This means that:

If you fail to make a repayment, your guarantor has to do so

If you fail to make a repayment, your guarantor must allow debt collectors access to your possessions

Neither of the above

Don’t know

Crowdfunding offers which of the following benefits?

There are safeguards to help you recover your money if the start-up company fails

You can predict what returns you are likely to get by looking at previous returns

Neither of the above

Don’t know

You fail to make a payment on your payday loan. What might the consequences be?

You are unlikely to become bankrupt due to consumer protection rules

The Money Advice Service will guarantee your loan so that you don’t end up in financial difficulty

You may end up in financial difficulty so severe that you may have to declare bankruptcy

Don’t know

Comprehension Question Set 3 (2D)

The company you have bought mini-bonds from has been made bankrupt. Which of the following is likely to be the case?

You can apply to get your money back from the Financial Services Compensation Scheme but you may not be successful

You are likely to lose your money

You may be able to transfer your investment to another company

Don't know

Which of the following statements about mini-bonds is most true? Choose one answer.

It is possible you will not receive any interest payments

You can cash them in at any point up to the full term

You can sell the bond to someone else if you want to (ie they are transferrable)

Don't know

Which of the following is most true about private defined contribution pensions?

The amount you get paid by a pension when you retire is defined by how much you paid into the pension

The amount you get paid by a pension depends on how much you paid in, and whether or not you were auto-enrolled

The amount you get paid by a pension depends on how much you paid in and how it was invested by the pension managers

Don't know

Which of the following statements about pension investments is most true? Choose one answer.

Pension investments are protected by the Government so you will never get back less than you pay in

It is possible to get back less than you pay in

Your employer is required to provide a minimum pension payment dependant on how long you have worked for them when you retire

Don't know

You started trading on a Contract for difference (CFD) platform after practising in a risk-free environment. You start trading with £100, but find that you make a lot of mistakes. By the end of the day, the least money you could have in your account is:

£0

-£200 (ie you owe the provider double your initial investment)

You could owe the provider an additional unlimited amount

Don't know

Which of the following statements about CFD products is most true? Choose one answer.

Most people make some money, they are a low risk investment product

About equal proportions of people make and lose money, they are a somewhat risky investment product

You're likely to lose money, they are a high risk investment product

Don't know

You are retired and want to release money from your home to pay for home improvements. You don't receive any government benefits and you intend to leave your house to your children when you die. Why might equity release not be suitable for you?

Because you would not retain ownership of your home

Because it may affect the inheritance you are able to leave

Because it is a product intended for those who receive government benefits

Don’t know

Which of the following statements about equity release is most true? Choose one answer.

It may affect your entitlement to state benefits

It will not reduce the value of your estate

Your family will inherit the entire property but can only do so after your death

Don’t know

You see an advert offering car insurance ‘from £200’. What percentage of customers are eligible to buy insurance costing £200?

At least 90%

At least 50%

At least 10%

Don’t know

Which of the following statements about car insurance is most true? Choose one answer.

If a company or broker advertises an average cover price then you will be offered at least one policy at that price

At least half of the company’s customers must be able to achieve an advertised cover price

A company or broker can advertise an example cover price even if only a small fraction of customers get such a good offer

Don’t know

You miss a number of payments on your logbook loan. Which of the following could happen?

Your car could be taken away, as you do not own it until you have made the final payment

Another company may make contact with you to recover the money owed, since the company you got the loan from only arranged the loan, rather than provided the money.

Both of the above

Don’t know

Which of the following statements about logbook loans is most true? Choose one answer.

If you fail to make repayments the lender can reclaim any of your belongings to repay the loan

If you fail to make repayments then the lender can only take the car used to get the loan

If you fail to make repayments then the lender can take ownership of any cars you own

Don’t know

Which of the following is not true about exchange traded funds (ETFs)?

ETFs don’t give you access to the whole market

Your original investment might go down

They are liquid

Don’t know

Which of the following statements about ETFs (Exchange Traded Funds) is most true? Choose one answer.

If the market starts to move against you, you can always withdraw your investment before the ETF begins to lose value

You directly own and have direct claim to the underlying investments in the fund

You may get back less than you invest

Don’t know

Which of the following statements about Guarantor Loans is most true? Choose one answer.

If you fail to make a repayment, your guarantor has to pay it

If you fail to make a repayment, your guarantor must help the loan company to communicate with you and provide access to you/your belongings

If you fail to make a repayment, this is recorded on your guarantor's credit history instead of yours

Don't know

Which of the following statements about Guarantor Loans is most true? Choose one answer.

A guarantor must be a close relative, or someone you have shared an address with

A guarantor must be someone of sufficient standing to act as character witness

A guarantor must be someone with sufficient funds to cover the repayments

Don't know

Which of the following is most true of crowdfunding?

There are safeguards to help you recover your money if the start-up company fails

You can predict what returns you are likely to get by looking at previous returns

It is legal for companies to advertise crowdfunding even if they believe the venture will fail

Don't know

Which of the following statements about investment-based crowdfunding is most true? Choose one answer.

It is very likely you will lose all your money, as most start-up businesses fail

Most platforms have a way you can cash in your investment

If the crowd-funding platform goes bust the return is still guaranteed

Don't know

You fail to make a payment on your payday loan. What might the consequences be?

Consumer protection rules mean that you will not become bankrupt

The Money Advice Service will guarantee your loan so that you don't end up in financial difficulty

You may end up in financial difficulty so severe that you may have to declare bankruptcy

Don't know

Which of the following statements about payday loans is most true? Choose one answer.

Companies are allowed to take past payday loans into account when deciding whether to approve you for another.

Late repayments can make it difficult to borrow money in any way, long into the future eg for future mortgage applications

Both of the above

Don't know

Annex 3: Financial literacy questions

Suppose you have £100 in a savings account and the interest rate is 20% per year and you never withdraw money or interest payments. After 5 years, how much would you have in this account in total?

More than £200

Exactly £200

Less than £200

Don't know

Which of the following statements is correct?

If someone buys a bond of firm A:

They own a part of firm A

They have lent money to firm A

They are liable for firm A's debt

None of the above

Don't know

Do you think this statement is true or false?

"Historically, bonds have been riskier than stocks."

True

False

Don't know

In the past, which asset described below has displayed the highest fluctuations over time?

Savings accounts

Stocks

Bonds

Don't know

Do you think this statement is true or false?

"After age 65, you have to convert all your pensions to provide a lifetime income."

True

False

Don't know

FCA FINANCIAL CONDUCT AUTHORITY