European Commission

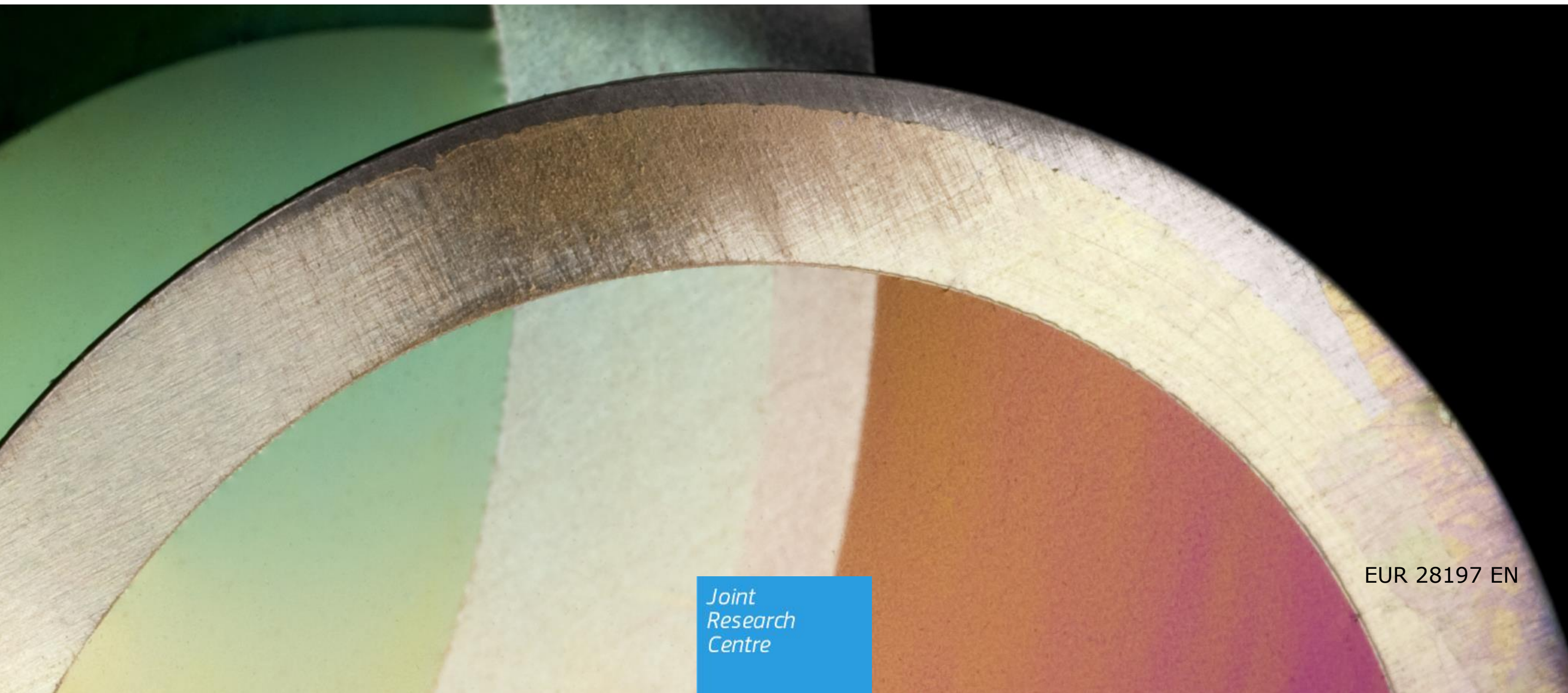
JRC TECHNICAL REPORTS

Nudging Online Security Behaviour with Warning Messages

Results from an online experiment

René van Bavel
Nuria Rodríguez-Priego

2016



EUR 28197 EN

Joint Research Centre

This publication is a Technical Report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Contact information
Address: Unit B6, Joint Research Centre, Edificio Expo, C/ Inca Garcilaso 3, Sevilla E-41092, Spain
E-mail: b06-sec@jrc.ec.europa.eu
Tel.: +34 954 488 211

JRC Science Hub
https://ec.europa.eu/jrc

JRC103223

EUR 28197 EN

PDF ISBN 978-92-79-63487-1 ISSN 1831-9424 doi:10.2791/2476

Luxembourg: Publications Office of the European Union, 2016

How to cite: van Bavel, & R. Rodríguez-Priego, N. (2016). Nudging Online Security Behaviour with Warning Messages: Results from an online experiment. *JRC Technical Reports.* EUR 28197 EN; doi:10.2791/2476

Table of contents

Acknowledgements

We are grateful to Jose Vila, Pam Briggs, Néstor Duch-Brown, Ioannis Maghiros, Rebeca Parra, Ignacio Alastrué, and Patricia Farrer for their help and support. The experiment was conducted by the Laboratory for Research in Experimental Economics (LINEEX), Unviersity of Valencia, in the frame of the project *Behavioural Insights on Cybersecurity* (JRC/SVQ/2014/J.3/0039/RC-AMI). Results were presented at the 2016 SABE/IAREP Conference: *Behavioural Insights in Research and Policy Making*, 8-10 July in Wageningen, the Netherlands. The views expressed in this article are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

Abstract

This study is part of a larger effort to better understand online behaviour. We tested the effect on people’s security behaviour of different ways of warning them about cybersecurity threats with an online experiment (n=5,065) in Germany, Sweden, Poland, the UK and Spain. Participants had to make a purchase in a mock online store, and their behaviour was observed through four behavioural measures. Results show that making users aware of the steps they can take to minimise their exposure to risk is effective in generating more secure behaviour, as suggested by protection motivation theory. Gain and loss-framed messages, and a message with a male anthropomorphic character, also had some effect on behaviour compared to the control group. The study also included a questionnaire. Results showed that more risk-averse participants exhibited more cautious behaviour. Finally, although they influenced behaviour itself, warning messages based on behavioural insights did not affect participants' self-reported knowledge of how to prevent cyberattacks.

Executive summary

This study is part of a larger effort to better understand online behaviour. It follows a recent trend where empirical findings about human behaviour are increasingly taken into consideration by policy-makers worldwide.

It tested the effect on people's security behaviour of different ways of warning them about cybersecurity threats. In the language of behavioural economics, it explored the role of changes to the choice architecture, or *nudges,* on online decision-making. By demonstrating their efectiveness, it makes a case for their consideration as an additional policy tool.

An online experiment (n=5,065) was conducted across five EU Member States: Germany, Sweden, Poland, the UK and Spain. Participants had to make a purchase in a mock online store, and their behaviour was observed through four behavioural measures: whether they connected to a server safely, whether they bought the product through a 'trusted vendor', whether they used secure passwords, and whether they remembered to log out.

The nudges were in the form of warning messages, which reminded consumers to navigate safely and appeared at the beginning of the shopping exercise. These messages were altered slightly, according to different behavioural insights.

Results show that making users aware of the steps they can take to minimise their exposure to risk is effective in generating more secure behaviour, as posited by protection motivation theory (PMT). These results challenge claims that users don't care or are somehow dismissive of the risks of cybersecurity. A more plausible explanation is that they are simply not aware of what they need to do. Warning messages should seek to directly remedy this.

The report also presents a number of other findings. Gain and loss-framed messages had an effect on behaviour compared to the control group, but the magnitude of the loss-framed was no greater than that of the gain-framed message. A male anthropomorphic character had an effect, but a female character did not. Finally, there was no difference between presenting a risk as having a small probability but a large impact or, conversely, as having a large probability but a small impact.

Of the various behavioural measures used, the choice of selecting a trusted vendor (for a fee) vs choosing an untrusted but free vendor is the one that had the greatest variability, across treatments and across countries. This behaviour, therefore, emerged as the one most likely to be affected by *nudges*. Policies to improve security could consider this behaviour as a starting point when piloting their initiatives.

The study, taking advantage of its large sample size, also included a number of questionnaire items. Their analysis revealed that more risk-averse participants will exhibit more cautious behaviour on a number of measures and that, on the whole, nudges did not affect their knowledge of how to prevent cyberattacks. Instead, nudges seem to work better on automatic rather than thought-out behaviour. The exception to this was knowledge of the benefits of logging out, which was affected by two PMT-inspired conditions.

One of the strengths of this report is the use of actual, observed online behaviour. Many studies have used either intention as the output variable or self-reported accounts of behaviour, both of which have limitations.

These results are useful for policy-makers interested in warning users about potential cybersecurity risks. Warning messages based on behavioural insights might not increase consumers' knowledge, but they can help improve their browsing experience and build online trust.

1 Introduction

While digital technology has enabled innovation, greater connectedness, economic growth and productivity, it has also given rise to a new threat: cybercrime. Users are aware of this threat and can shy away from certain activities over the Internet (especially those that involve the disclosure of personal information or economic transactions). In Europe, users still worry about the lack of security in online payments which prevent them from using the Internet for e-commerce (27% of respondents, according to a recent survey). They also have concerns about returning goods or making complaints (19%), and some still claim they lack the necessary skills to make transactions online (13%; EC, 2016)[1]. This means that the full potential of digital technology to empower consumers and drive economic growth remains unrealised.

Reinforcing trust and security in the online environment, therefore, has become a political priority. In Europe, the European Commission has stressed the need to make the European digital economy more trusted and secure, so that citizens and businesses can fully reap its benefits. This objective is recognised in the Digital Single Market strategy[2] (a top priority for President Jean-Claude Juncker), and by the Digital Agenda for Europe (DAE), the European Union flagship initiative on all ICT-related activities.

Increased trust must go hand-in-hand with increased security. If trust increases in an insecure environment, more people will disclose sensitive information and be vulnerable to cyberattacks. A balance needs to be struck.

Researchers and security experts are increasingly aware that security provided by stronger digital barriers has limits. No matter how perfect your security system is, it will always depend on people's behaviour (e.g. not clicking on malicious links and keeping secure passwords). Human error is still one of the weakest links in the cybersecurity chain, and is responsible for nearly one-quarter of all cybersecurity failures (Waldrop, 2016).

People make mistakes because they are human, because they do not have sufficient information, because their online behaviour can become habitual (leading them to pay less attention and attribute less importance to the decision making process), or because they perceive the risks as low. Whatever the case, no matter how secure the technology, or how strong the legal system for persecuting cybercrime, cybercrime will continue to exist unless people navigate safely. And as long as cybercrime exists, people will not fully trust the online environment.

Making people aware of the risks involved and getting them to behave safely online, however, is no easy task. Traditionally, users have been informed through warning messages, intended to increase their awareness and reassure them about the risks involved. However, this approach assumes people are rational and well-informed, which they are not. To be effective, policy initiatives should not rely on users making very informed or rational decisions (Acquisti, Brandimarte & Lowenstein, 2015). Greater insights into how people behave online are required.

Hackers seem to have advanced knowledge on this. They send phishing emails from a sender that seems authoritative, at a time of the day when we are busy, increasing the chances that we will click where we should not. The institutions meant to defend users, on the other hand, are lagging behind. The excessive requests for authentication in an organisation (23 per day on average, according to a study by Stevens, 2014), drain people's time and mental energy. And recent evidence suggests that the guidelines for proper password management are misguided (Waldrop, 2016).

[1] https://ec.europa.eu/digital-single-market/en/news/europes-digital-progress-report-2016

[2] http://ec.europa.eu/priorities/digital-single-market/

But while the focus has traditionally been on cybersecurity, from the perspective of computer science and not psychology, the tide seems to be turning. Recently President Obama proposed spending more than $19 billion on federal cybersecurity funding, including a research and development plan that makes human-factors research an explicit priority. There is a similar trend in the UK, which focuses on, for example, how criminals organise their business and how to help users with their passwords (Waldrop, 2016).

This report follows that trend. It is part of a larger initiative which explores the contribution of behavioural insights to cybersecurity. The aim is to observe whether changes in the design of web interfaces (i.e. the *choice architecture* according to the behavioural economics literature) lead to changes in online behaviour, and so merit attention as a policy tool. This approach builds on the premise that nudges, which are changes in the choice architecture to elicit certain behaviours, have been shown to be effective in other domains.

2 The insights behind the nudges

This study is based on an online experiment (n = 5,065) across five European countries: Germany, Sweden, Poland, Spain and the UK. We randomly assigned individuals to a control group or one of nine treatment groups, and then let them make a purchase in a mock e-commerce exercise. The nudges applied were subtle, and were embedded in a warning message reminding them to navigate safely. The way in which this message was framed and how it directed participants' attention differed across treatment groups. These nudges came from the relevant literature or from previous empirical work conducted in this field by the authors.

2.1 Protection motivation theory

One set of nudges was based on insights from protection motivation theory (PMT; Rogers, 1975, 1983), which seeks to clarify our concepts on the cognitive processes which mediate behaviour in the face of a threat. It posits that, when facing a threatening event, people conduct two appraisal processes: one focused on the threat itself and the other on the options they have to diminish it (*threat appraisal* and *coping appraisal*, respectively). This will affect their intention to take precautionary action and will result in adaptive or maladaptive behaviours vis-à-vis the threat.

In their threat appraisal, people will consider how negative the consequences of the threat are (*perceived severity*) and the likelihood of the threat materialising (*perceived vulnerability*). In their coping appraisal, people will assess whether undertaking a recommended course of action will remove the threat (*response efficacy*) and their level of confidence in being able to carry it out (*self-efficacy*) (Maddux & Rogers, 1983; Boer & Seydel, 1996).

PMT has been applied to online safety protection, specifically to virus protection behaviour (Lee, LaRose and Rifon, 2008), security behaviour among people who know how to protect their systems but fail to do so (Workman, Bommer & Straub, 2008), security behavioural intentions of home computer users (Anderson & Agarwal, 2010), convincing Internet users to protect themselves (Shillair et al, 2015), teenagers' willingness to provide information online (Youn, 2005), security behaviour in response to fear appeals by employers (Johnston & Warkentin, 2010), and employees' adherence to information security policies (Siponen, Mahmood & Pahnila, 2014).

Research in information systems has mainly used *intention to behave* as the main dependent variable when using PMT. This is probably because the initial formulation of PMT in social psychology used *intention to behave*. But of course there is a gap between intention and behaviour, and this is a limitation of the studies that have used PMT to explain security behaviour intention (Johnston and Warkentin, 2010, Liang & Xue 2010; Lee, 2011; Herath & Rao, 2009; Crossler, Long, Loraas & Trinkle, 2014).

Some studies have used behaviour as the dependent variable (Neuwirth, Dunwoody & Griffin, 2000; Woon et al, 2005; Workman et al. 2008). Quite clearly, behaviour works better than intention, as an outcome variable. Behaviour is more interesting, as protection of information resources does not happen when individuals intend to behave, but rather when they actually do so (Crossler et al. 2013). However, even in cases where behaviour is the dependent variable, self-reported behaviour is used (Crossler, 2014). This study, by contrast, relied on actual, observed behaviour. This feature constitutes one of its strengths and a distinct contribution to the field.

Three PMT-inspired warnings were tested:

- *Coping appraisal message*: facilitated individuals' coping appraisal by telling them it was easy to minimise the chances of a cyberattack and by indicating what steps they could take.
- *Threat appraisal message*: sought to heighten the individual's perception of the threat by telling individuals they could be subject to a virus attack.

- *Coping and threat appraisal message*: combined both elements described above into one warning message.

The threat appraisal message highlighted both the severity of the threat and the user's vulnerability at the same time, since prior research (outside security behaviour) suggests they jointly determine the likelihood of individuals performing adaptive behaviours (Neuwirth, Dunwoody & Griffin, 2000). We considered that the cost of breaking the message down into these two components, in terms of an additional experimental treatment, would outweigh the value of having information for both elements separately. Moreover, it would run the risk of giving too weak a warning to have any impact at all.

The following hypotheses guided the application of these insights to the study:

> *Hypothesis 1.* The group exposed to the coping appraisal message will show more secure online behaviour than the control group.
>
> *Hypothesis 2.* The group exposed to the threat appraisal message will show more secure online behaviour than the control group.
>
> *Hypothesis 3.* The group exposed to the threat and coping appraisal message will show more secure online behaviour than the control group.
>
> *Hypothesis 4.* The group exposed to the threat and coping appraisal message will show more secure online behaviour than the coping appraisal message group.
>
> *Hypothesis 5.* The group exposed to the threat and coping appraisal message will show more secure online behaviour than the threat appraisal message group.

2.2 Gain vs. loss framing

One of the main contributions of behavioural economics has been to show that the way in which information is presented, i.e. the *framing* of a message, can have an effect on behaviour. This is contrary to conventional economic thinking, which does not consider these changes to be significant. The rational consumer is expected to see through the fog and arrive at the relevant kernels of information.

One of the main ways that framing has an effect is by presenting situations in terms of gain or losses. According to principles such as loss aversion and the endowment effect, the pain of losing something is greater than the joy of getting it (Kahneman, 2011). Messages which highlight potential losses should therefore be more effective than those highlighting potential gains.

The following hypothesis guided the application of this insight to the study:

> *Hypothesis 6.* The group exposed to the loss-framed condition will show more secure online behaviour than the control group.
>
> *Hypothesis 7.* The group exposed to the gain-framed condition will show more secure online behaviour than the control group.
>
> *Hypothesis 8.* The group exposed to the loss-framed condition will show more secure online behaviour than the group exposed to the gain-framed condition.

2.3 Anthropomorphic characters

Human-like characters are increasingly used in eCommerce. They increase trust and perception of enjoyment, especially when they look like traditional salespersons offering a helping hand (Qiu & Benbasat, 2009). However, their effects on behaviour are less clear. Some researchers argue that an anthropomorphic character makes people feel observed, which leads them to be more careful in their disclosure of personal information (Groom & Calo, 2011; Moon, 200). Others argue that since they are unwittingly treated as trustworthy counterparts, they invite people to disclose greater amounts of personal information (Bente, Dratsch, Relibach, Reyl & Lushaj, 2014; Heckman & Wobbrock, 2000).

Previous empirical findings on nudges to security behaviour show that a male anthropomorphic character increases security behaviour. However, there is no such effect for a female character (Rodriguez-Priego & van Bavel, 2016). In order to follow-up these findings and investigate further the effect of anthropomorphic characters, we posited the following hypotheses:

> *Hypothesis 9.* The group exposed to the female anthropomorphic condition will show more secure online behaviour than the group exposed to control condition.
>
> *Hypothesis 10.* The group exposed to the male anthropomorphic condition will show more secure online behaviour than the group exposed to control condition.

2.4 Low-risk, high-impact vs high-risk, low-impact

Another insight from behavioural economics is that people will take risks when dealing with potential losses, but will avoid them when dealing with potential gains. For example, people prefer a 50% chance of losing 1000 euros to a certain loss of 500 euros, but they will prefer a certain gain of 500 euros over a 50% chance of winning 1000 euros (Kahneman & Tversky, 1979). This insight is based on loss aversion: people are willing to take a risk in order to avoid the pain of losing. But it is not captured in the loss vs. gain-framed conditions.

The loss vs. gain-framed conditions sought to test whether the mere framing of the impact of an attack as a gain or as a loss would lead to a change in behaviour. These conditions, on the other hand, sought to explore whether the prospect of having an almost-certain, albeit small, loss caused more dread (and therefore led to more secure behaviour) than an uncertain, albeit large, loss.

Security behaviour can be characterised as being low-risk, but with a potentially devastating effect if a breach occurs. In this scenario, people take risks. However, if the risk were higher, but the impact less severe, people would behave more securely, following prospect theory. Faced with choosing between low-risk, high-impact vs. high-risk, low-impact (*ceteris paribus*), people should prefer the former. The latter situation would be less preferable and would therefore make people more careful in their online behaviour.

In the experiment, this insight translated into two conditions, both of which made reference to participants' variable fee. As described below, this fee varied depending on how securely participants navigated. One condition (*low-risk, high-impact*) said that 1 in 10 people who did not navigate safely in the website would lose 90% of their variable fee. The other (*high-risk, low-impact*) said that 9 out of 10 people who did not navigate safely would lose 10% of their variable fee. Since people want to avoid certain losses, the latter condition should display more secure behaviour than the former.

Since these conditions talked about a 1-in-10 chance of losing 90% of their points or a 9-in-10 chance of losing 10%, a comparison with the control condition (which gave no numbers of this kind) was not appropriate. They were only comparable with each other.

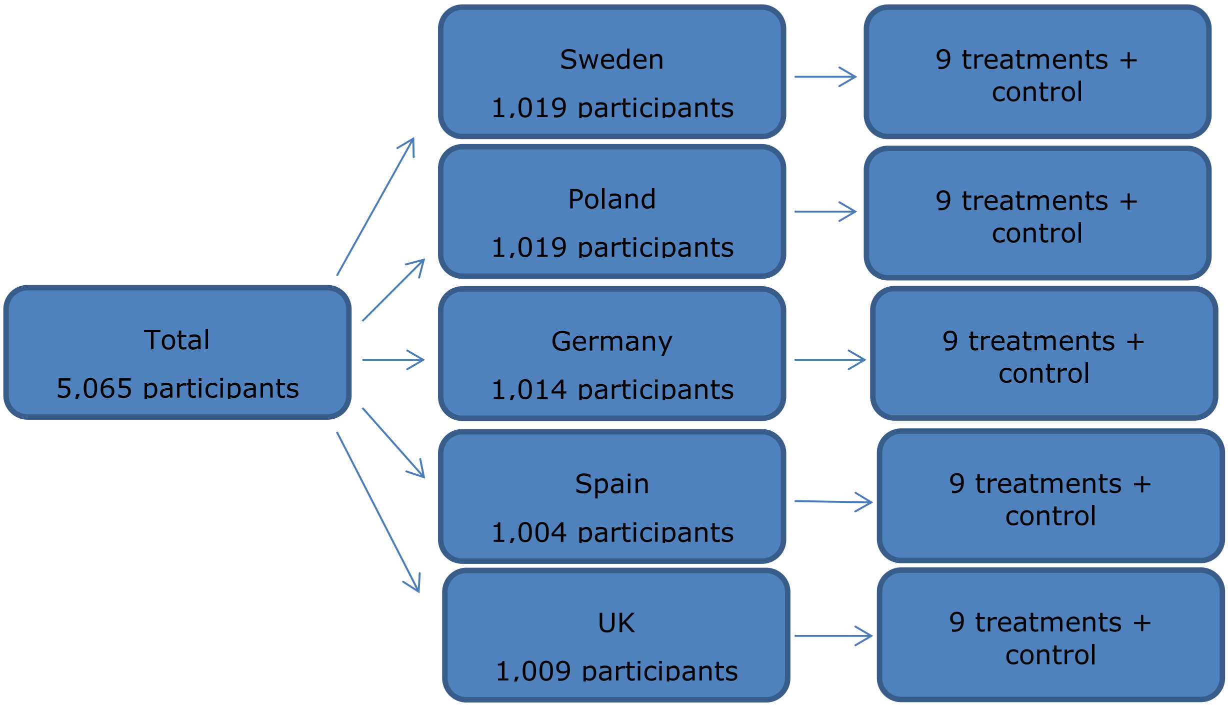
> *Hypothesis 11.* The group exposed to the *high–risk, low-impact* condition will show more secure online behaviour than the group exposed to the *low risk, high-impact* condition.

3 Methodology

Data was collected by the Laboratory for Research in Behavioural Experimental Economics (LINEEX), Centre for Research in Social and Economic Behaviour (ERI-CES), University of Valencia using panel data (i.e. subject pools). The sample consisted of 5,065 participants (50.56% females[3]) from five countries, representing different broad cultural areas of the EU: Sweden, Poland, Germany, Spain and the UK. Data on the profile of the population using the internet was taken from Eurostat. The sample was distributed evenly across countries (around 1,000 per country).

The Ethics Committee on Experimental Behavioural Economics at ERI-CES approved this experiment and confirmed that it adhered to its charter of ethics. The experiment was carried out between October and December 2015. The study tested a total of ten security messages based on the insights described in the previous section, and targeted around 100 subjects per experimental treatment, balanced according to age and gender.

Figure 1: Division of sample across countries and treatments



3.1 Experimental conditions

In the purchasing process, participants were asked to buy a real product (desktop wallpaper). While doing so, they had to make several decisions that affected their security. The ten security messages appeared as pop-ups in the centre of the screen before the purchasing process began. Participants had to close the pop-up window to continue with the experiment. The message was then placed in the upper part of the screen and remained there throughout.

In the control condition for this experiment, the pop-up message simply reminded the participant to navigate safely. There was the option of having a control condition with no security message at all (rather than a simple security message). However, a control condition of this kind would have limited the value of a potential result in any of the experimental treatments. In particular, it would not have been possible to determine

[3] Further information on socio-demographics can be found in Table 1.

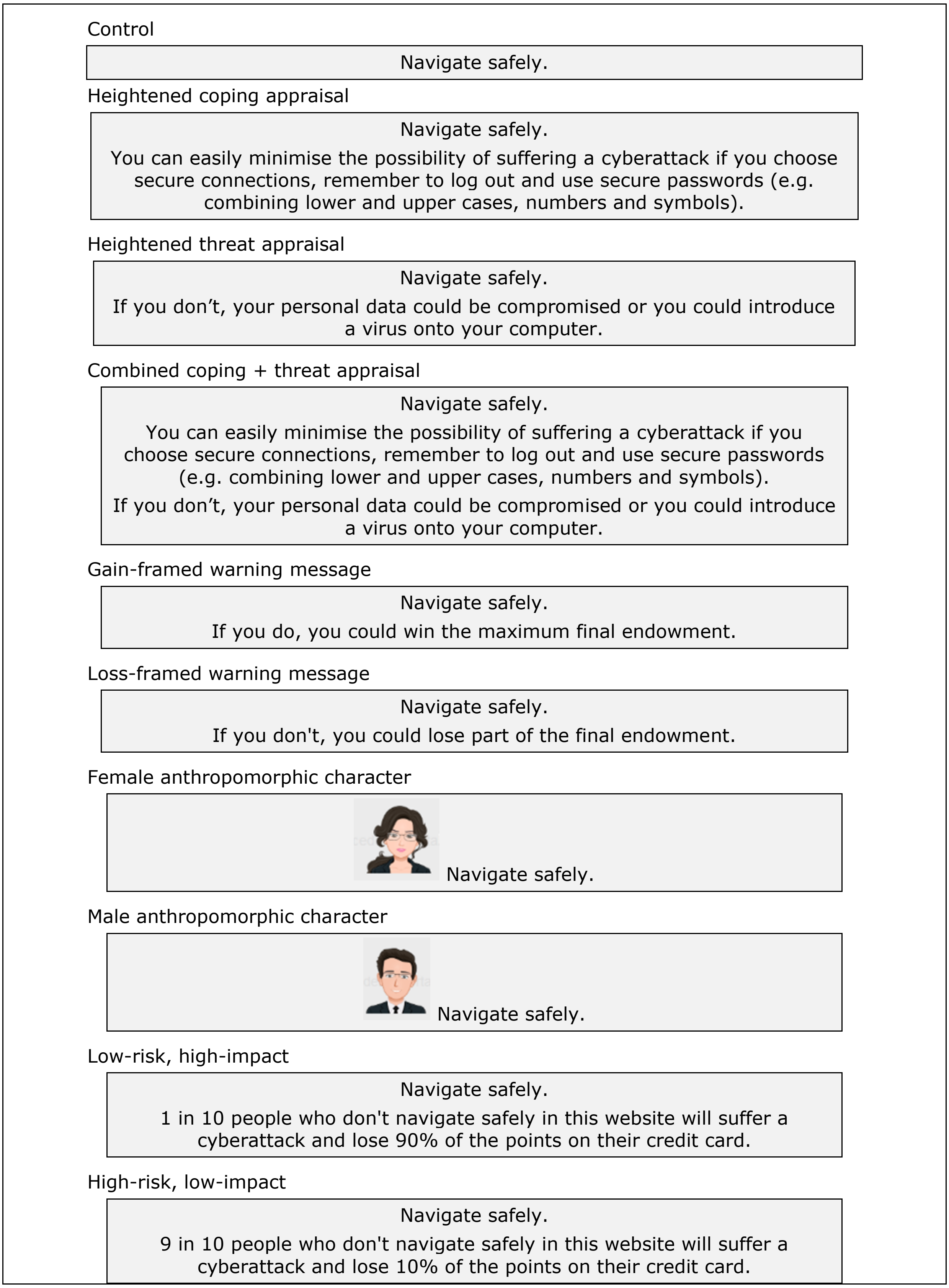
whether the effect of a treatment was due to the nature of the message itself (and hence the importance of the behavioural insights that fed into its design) or due to the fact that one condition had a message and the other had none at all. Warning messages on their own, irrespective of their design, are presumed to have an impact on behaviour.

The experimental conditions were:

1. Heightened coping appraisal: this condition was based on protection motivation theory (PMT), and sought to heighten self-efficacy and response efficacy, both components of "coping appraisals". The security message told participants they had the ability to easily make decisions on how safely they navigated, and gave indications on how they could protect themselves.
2. Heightened threat appraisal: this condition was also based on PMT. It sought to heighten participants' perception of the cybersecurity threat. The security message warned that if they did not navigate safely, their personal data could be compromised or a virus could be introduced onto their computers.
3. Combined coping + threat appraisal: this condition included both elements of the coping appraisal and threat appraisal messages. The aim was to test whether providing both types of appraisals would increase the effectiveness of the security message.
4. Gain-framed warning message: this condition added a gain-framed message to the reminder to navigate safely. It highlighted how much variable income participants could *win* if they navigated safely.
5. Loss-framed warning message: this condition added a loss-framed message to the reminder to navigate safely. It highlighted how much variable income participants could *lose* if they failed to navigate safely.
6. Female anthropomorphic character: this condition presented the same reminder to navigate safely as the control condition, but added a female anthropomorphic character inside the pop-up message.
7. Male anthropomorphic character: this condition presented the same reminder to navigate safely as the control condition, but added a male anthropomorphic character inside the pop-up message.
8. Low-risk, high-impact condition: in addition to the reminder to navigate safely (same as in the control condition), this condition stated that 1 in 10 people who did not navigate safely in this website would suffer a cyberattack and lose 90% of their variable incentive.
9. High–risk, low-impact condition: this was the counterpart to the low-risk, high-impact condition. It stated, in addition to the reminder to navigate safely, that 9 in 10 people who did not navigate safely in this website would suffer a cyberattack and lose 10% of their variable incentive.

Conditions 9 and 10 were only comparable with each other, and not with the control condition. They differed from the control condition on two counts. Not only were their security messages different in form, but they were also different in substance. They mentioned specific probabilities, whereas the control condition (and other experimental conditions, for that matter) did not.

Figure 2: Ten security messages



Control

Navigate safely.

Heightened coping appraisal

Navigate safely.

You can easily minimise the possibility of suffering a cyberattack if you choose secure connections, remember to log out and use secure passwords (e.g. combining lower and upper cases, numbers and symbols).

Heightened threat appraisal

Navigate safely.

If you don’t, your personal data could be compromised or you could introduce a virus onto your computer.

Combined coping + threat appraisal

Navigate safely.

You can easily minimise the possibility of suffering a cyberattack if you choose secure connections, remember to log out and use secure passwords (e.g. combining lower and upper cases, numbers and symbols).

If you don’t, your personal data could be compromised or you could introduce a virus onto your computer.

Gain-framed warning message

Navigate safely.

If you do, you could win the maximum final endowment.

Loss-framed warning message

Navigate safely.

If you don't, you could lose part of the final endowment.

Female anthropomorphic character

Navigate safely.

Male anthropomorphic character

Navigate safely.

Low-risk, high-impact

Navigate safely.

1 in 10 people who don't navigate safely in this website will suffer a cyberattack and lose 90% of the points on their credit card.

High-risk, low-impact

Navigate safely.

9 in 10 people who don't navigate safely in this website will suffer a cyberattack and lose 10% of the points on their credit card.

3.2 Behavioural measures

The experiment measured four behavioural outcomes that are considered necessary for users to maintain cybersecurity (Coventry, Briggs, Jeske & van Moorsel, 2014). Although the list of such behaviours is longer, the experiment focussed on those related to online purchasing processes which could be feasibly tested in an experiment. These are:

3.2.1 Secure connection

This behavioural measure was designed to reflect the real world costs of stringent cybersecurity behaviour. It sought to evoke the *compliance budget* that users resort to when making a decision (Beautement, Sasse & Wonham, 2009). Participants had to choose between decreasing the risk of suffering a cyberattack, by spending extra time when connecting to a simulated intranet, or selecting an immediate connection, which increased their risk of a cyberattack.

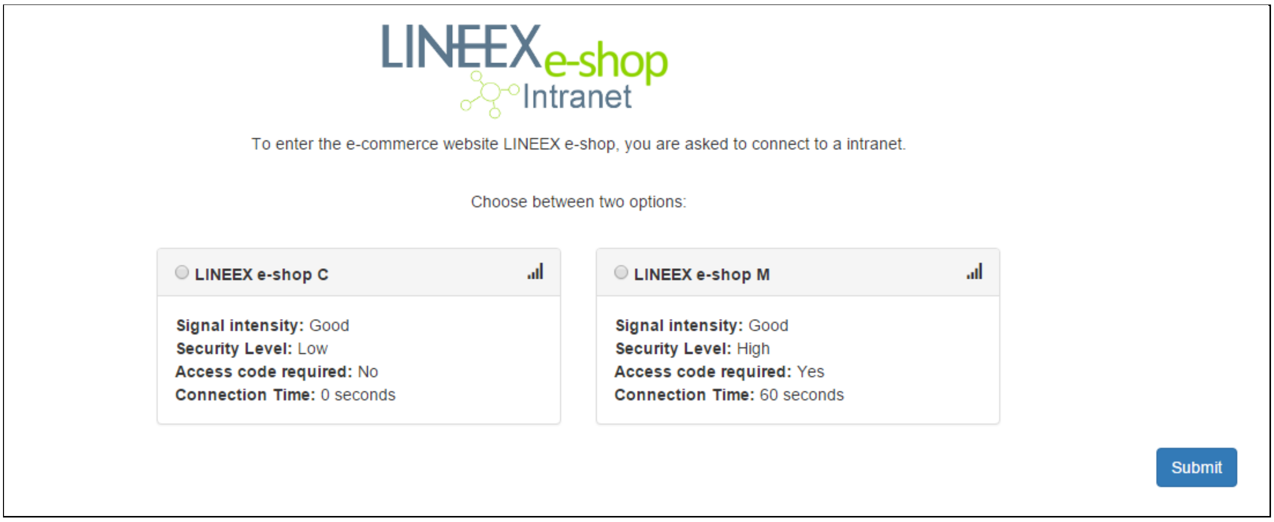
A connection to the simulated intranet was necessary before entering the eCommerce website. Participants could choose between two options: a secure vs. an unsecured connection. The variable "secure connection" was binary. It scored 0 if subjects chose to behave unsafely and selected the unsecured option; and 1 if they made the secure choice. The options appeared randomly on the left or right-hand side of the screen to avoid location having an effect on participants' decisions.

The unsecured connection meant an instant connection to a simulated intranet. Participants did not have to wait for the connection, and it did not require a password (see Figure 3).

For the secure connection, participants had to wait 60 seconds and they had to type in a complicated access code provided on the screen, which included a random combination of 12 upper- and lower-case letters and numbers (i.e. it required an additional cognitive effort, see Figure 4). Participants were made aware that it would take 60 seconds to connect, but that the connection was secure.

The next screen displayed a processing bar during the connection. Below the bar, participants could see a button that allowed them to switch to the unsecured but immediate connection if they did not want to wait the entire 60 seconds. By including this possibility, participants could change their minds as they do in the real world.

Figure 3: Intranet connection screen with two options



To enter the e-commerce website LINEEX e-shop, you are asked to connect to a intranet.

Choose between two options:

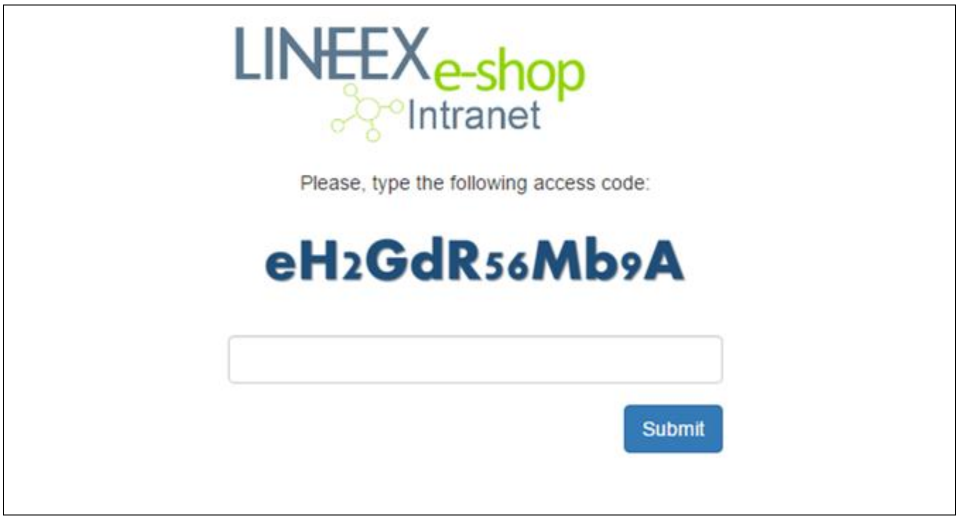
○ **LINEEX e-shop C**

Signal intensity: Good
Security Level: Low
Access code required: No
Connection Time: 0 seconds

○ **LINEEX e-shop M**

Signal intensity: Good
Security Level: High
Access code required: Yes
Connection Time: 60 seconds

Figure 4: Requirement to enter access code



Please, type the following access code:

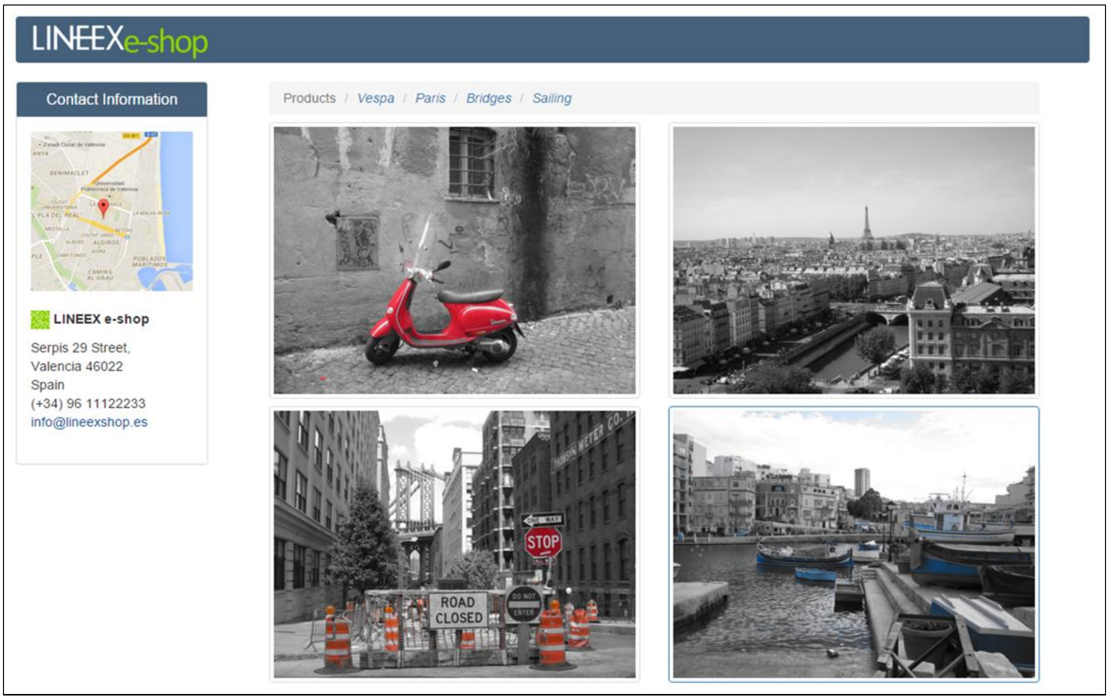
eH2GdR56Mb9A

3.2.2 Trusted vendor

This behavioural measure sought to reflect how, in the real world, access to free products often implies a security risk, which is effectively eliminated when a product is purchased through a trusted vendor.

Once subjects connected to the intranet, they were able to see the eCommerce website. The home page contained the company name and logo. In the bottom left-hand corner, there was a link to the terms and conditions. The link contained information about how the data would be managed, used and stored; the rights of the user, and copyright information. All this information followed the Data Protection Directive 95/46/EC. Participants had to accept the terms and conditions during the sign-up process by clicking the button "I agree to the Terms and Conditions". The home page was the entry point for subjects to choose the products (see Figure 5).

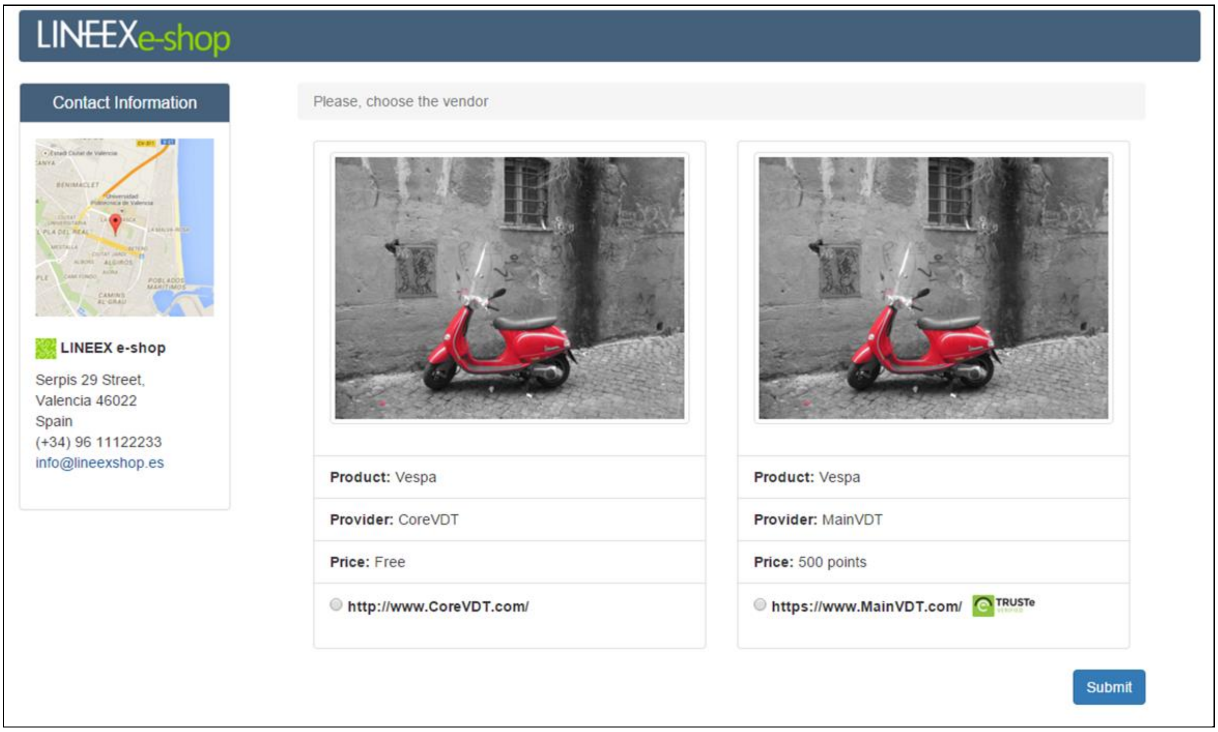
Figure 5: Home page



LINEEX e-shop
Serpis 29 Street,
Valencia 46022
Spain
(+34) 96 11122233
info@lineexshop.es

When a subject clicked on a product, a detailed page for that product opened. On this page, the subject had to choose between two vendors (their order altered randomly; Figure 6). One vendor offered the product for free. In this case, the link to download the product had no security signs (i.e. no image for an e-trusted site appeared). The simulated link for this supplier was http (Hypertext Transfer Protocol). The other vendor offered the product for €2, but the link to download it was https (Hypertext Transfer Protocol *Secure*) and appeared next to an image indicating it was an e-trusted site. The measure scored zero if participants chose the unsecured option and one if they chose the secure option (i.e. the trusted vendor).

Figure 6: Choice of vendor page



LINEEXe-shop

Contact Information

LINEEX e-shop
Serpis 29 Street,
Valencia 46022
Spain
(+34) 96 11122233
info@lineexshop.es

Please, choose the vendor

Product: Vespa
Provider: CoreVDT
Price: Free
○ http://www.CoreVDT.com/

Product: Vespa
Provider: MainVDT
Price: 500 points
○ https://www.MainVDT.com/ TRUSTe

Submit

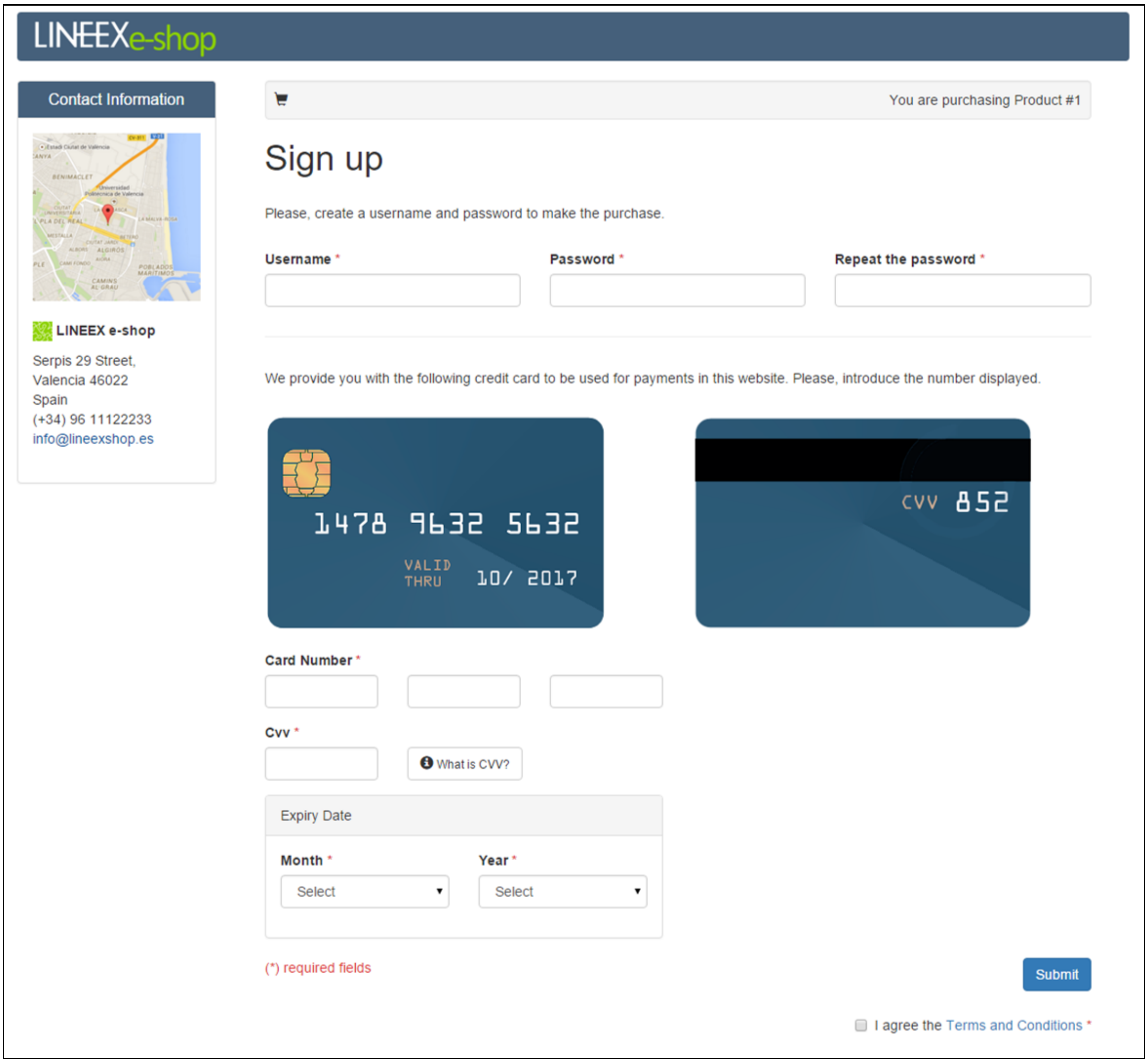
3.2.3 Password strength

After selecting one of the vendors, participants had to register by creating a username and password. On the same screen, they were also asked to introduce a credit card number, CVV and expiry date. A simulated credit card was provided to participants on the screen (see Figure 7).

The behavioural measure 'password strength' established the strength of the chosen password. It was measured according to six common security parameters and scored between zero (if subjects did not meet any of the parameters) and six (if they met them all). These parameters were the following:

- Minimum number of characters: 8
- Minimum number of lower case characters: 2
- Minimum number of upper case characters: 2
- Minimum number of numeric digit characters: 2
- Minimum number of special characters: 2
- Boolean check whether password contains the username

Figure 7: Page requesting password



LINEEXe-shop

Contact Information

LINEEX e-shop

Serpis 29 Street,
Valencia 46022
Spain
(+34) 96 11122233
info@lineexshop.es

You are purchasing Product #1

Sign up

Please, create a username and password to make the purchase.

Username *

Password *

Repeat the password *

We provide you with the following credit card to be used for payments in this website. Please, introduce the number displayed.

1478 9632 5632

VALID THRU 10/ 2017

CVV 852

Card Number *

Cvv *

What is CVV?

Expiry Date

Month * Select

Year * Select

(*) required fields

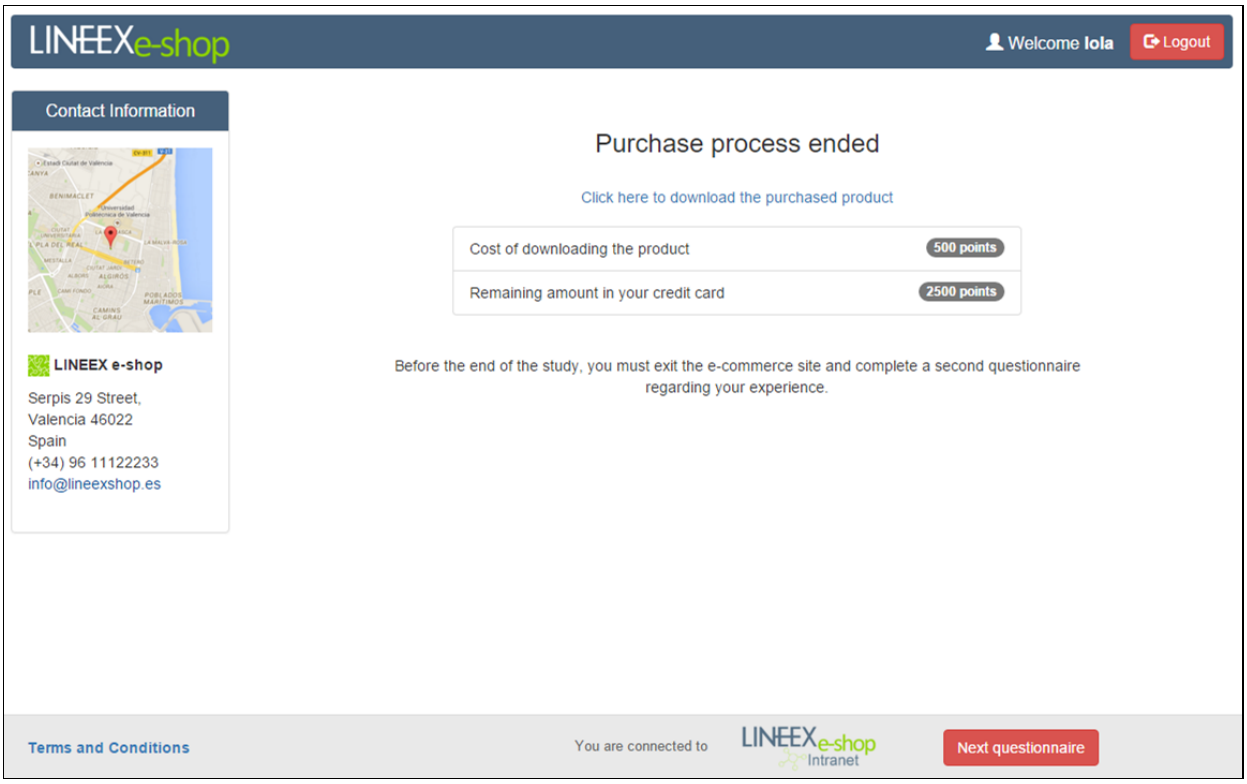
Submit

□ I agree the Terms and Conditions *

3.2.4 Log-out

This behavioural measure sought to document whether users were attentive and logged out of their eCommerce session, or whether they continued to navigate without logging out. Once subjects had completed the purchasing process, a screen displayed information about the cost of the product purchased and how much they had left on their credit cards. A new button appeared at the bottom right-hand side of this screen which led participants to the 'next questionnaire'. However, they had the option to log-out before doing so, by clicking on a button in the top right-hand corner (see Figure 8). Participants were not directly guided to log-out. They were simply asked to exit the eCommerce site and complete the second questionnaire. The behavioural measure 'log-out' scored zero if they just clicked on the 'next questionnaire' button and one if they chose the safe option and logged-out first.

Figure 8: Log-out page



Purchase process ended

Click here to download the purchased product

Cost of downloading the product	500 points
Remaining amount in your credit card	2500 points

Before the end of the study, you must exit the e-commerce site and complete a second questionnaire regarding your experience.

LINEEX e-shop
Serpis 29 Street,
Valencia 46022
Spain
(+34) 96 11122233
info@lineexshop.es

4 Results

The final sample of 5,065 was segmented according to gender and age (see Table 1 for a breakdown). With regard to the education of participants, most participants had either finished high school or had a university degree (see Table 2 for a breakdown).

Table 1: Sample socio-demographics by country

Gender	Age					
	<35's		35+'s		*Total*	
	n	%	n	%	*n*	*% of total*
Spain						
Men	198	38.18	320	61.82	*518*	*51.59*
Women	208	42.84	278	57.16	*486*	*48.41*
Total	*406*	*40.44*	*598*	*59.56*	*1004*	*100*
UK						
Men	149	30.98	332	69.02	*481*	*47.67*
Women	231	43.75	297	56.25	*528*	*52.33*
Total	*380*	*37.66*	*629*	*62.34*	*1009*	*100*
Germany						
Men	184	35.87	329	64.13	*513*	*50.59*
Women	182	36.33	319	63.67	*501*	*49.41*
Total	*366*	*36.09*	*648*	*63.91*	*1014*	*100*
Sweden						
Men	156	30.65	353	69.35	*509*	*49.95*
Women	208	40.78	302	59.22	*510*	*50.05*
Total	364	35.72	655	64.28	1019	*100*
Poland						
Men	245	49.10	254	50.90	*499*	*48.97*
Women	271	52.12	249	47.88	*520*	*51.03*
Total	*516*	*50.64*	*503*	*49.36*	*1019*	*100*
Total sample						
Men	932	36.98	1588	63.08	*2520*	*49.75*
Women	1100	43.22	1445	56.78	*2545*	*50.25*
Total	*2032*	*40.12*	*3033*	*59.88*	*5065*	*100*

Table 2: Education of participants in the total sample

Education level	**n**	**%**
No studies	17	0.34
Primary or secondary education	674	13.31
High school or technical education	2,094	41.34
University graduate	1,686	33.29
Postgraduate	483	9.54
PhD	87	1.72
No answer	24	0.47

A total of 26,991 participants clicked on the email that gave access to the experiment, but only 17,700 accessed the experiment. Out of these, 5,322 completed the experiment. However, 257 of these were classified as 'speeders', i.e. they completed one of the questionnaires or the full experiment in less than one third of the median time allocated by participants in a given country. See Table 3 for a breakdown of data by country.

Table 3: Breakdown of participants by country

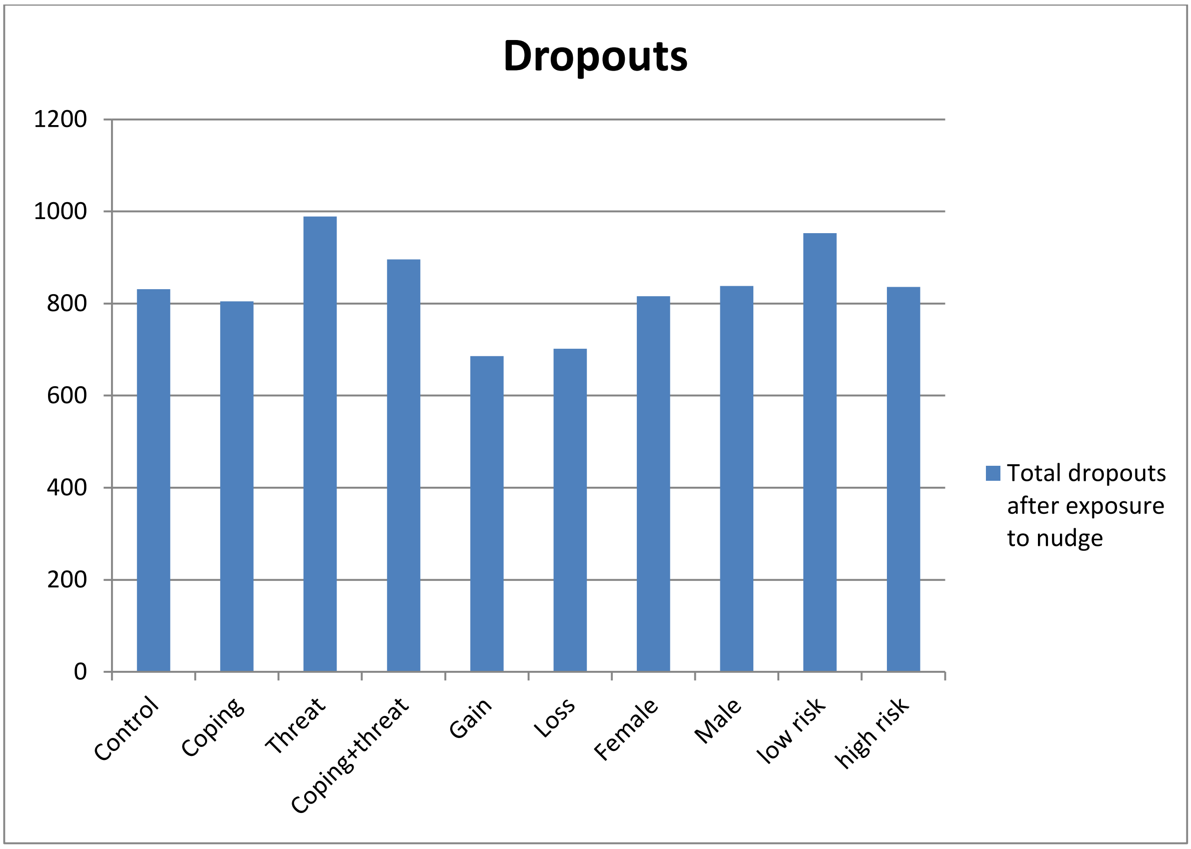
	Spain	**UK**	**Germany**	**Sweden**	**Poland**	**Total**
Total subjects click the email	4180	5577	6041	6003	5190	26991
Total subjects access the experiment	2799	3629	3736	3867	3669	17700
Total subjects complete the experiment	1064	1062	1061	1052	1083	5322
Total 'speeders'	60	53	47	33	64	257
Total subjects	*1004*	*1009*	*1014*	*1019*	*1019*	*5065*

The number of dropouts (i.e. those who accessed the experiment but decided to not to complete it) merits close attention, as it might have a bearing on the results presented in this section. In particular, the warning messages about cybersecurity threats (particularly the *threat appraisal* condition, which made reference to a virus) might have had an effect on the decision to dropout out of the experiment. Indeed, most dropouts occurred at the stage of the purchasing process where this message appeared.

When we look at the number of dropouts that occurred *only after exposure to the nudge*, it emerges that the threat appraisal condition witnessed the largest number of dropouts (66.2%), followed by the low-risk, high-impact condition (65.1%) and the combined coping + threat appraisal condition (63.8%). The control condition showed a dropout rate of 62.1%. The low-risk, high impact condition is not comparable to the control condition. But when we compare the threat appraisal condition to the control, the difference is statistically significant (t = 2.29, p<0.05). Not so the coping + threat condition. It appears that the threat of a virus getting onto the participant's computer might have been too strong, and may have led to an increased number of dropouts.

Also of note are the gain-framed and loss-framed conditions, which had noticeably lower dropout rates than the control condition (57.7% and 58.0% respectively, vs 62.1%). These differences were statistically significant (t = 2.36, p<0.05 and t = 2.13, p<0.05, respectively). It appears that the reminder to participants that they stood to gain a reward for navigating safely (however this was framed) was enough of a nudge to make them keep participating in the experiment. This analysis suggests that the dropout rate, its implications for the composition of the sample and its possible inclusion as an output measure, merits further attention.

Figure 9: Total number of dropouts after exposure to the warning message, per experimental condition



4.1 Effect of experimental conditions

Next we present the results of the impact of the experimental conditions on the behavioural measures. For each set of behavioural insights, we provide information on the distribution of the decisions made by the participants in this experiment over the four behavioural measures under consideration, plus a composite indicator.

Two-tailed t-tests were conducted to test the hypotheses. They compared the means of groups under observation (i.e. control condition vs. treatment, or treatment vs. treatment). Information regarding the subsamples – i.e. mean, standard deviation, and minimum and maximum score – is also provided.

4.1.1 PMT-inspired conditions

Results from the PMT-inspired conditions show that the *heightened coping appraisal* condition had an effect on all four behavioural measures. In all of them, participants exposed to a message that said they could easily minimise the possibility of suffering a cyberattack by choosing connections, remembering to log out and using secure passwords, exhibited more secure behaviour than participants in the control condition. Hypothesis 1 is supported in all four behavioural measures.

The *heightened threat appraisal* condition, on the other hand, was less effective. Warning users that failing to navigate safely could lead to their personal data being compromised or a virus getting onto their computer only had an effect on the 'trusted vendor' measure. Hypothesis 2 is only supported in one behavioural measure, but not in the remaining three. However, we must always consider the possibility that the increased dropout rate for *heightened threat appraisal* resulted in a sample that was slightly different to the other samples. It could be that those who did not drop out were less likely to navigate safely (since those who were prone to navigating safely might have been scared at the first sight of the nudge and dropped out). It is possible that the effects of the threat appraisal condition are understated.

With regard to the combination of both PMT elements (coping and threat appraisal) into one condition, results show that it was effective in generating more secure behaviour in three behavioural measures: 'trusted vendor', 'password strength' and 'log-out'. Hypothesis 3 is therefore supported in these measures (but not in 'secure connection').

There were no statistically significant differences in the means of all four behavioural measures between *combined coping + threat appraisal* and *heightened coping appraisal.* Hypothesis 4 is rejected. The analysis showed differences between *heightened threat appraisal* and *combined coping + threat appraisal*, but only for two of the four behavioural measures: 'password strength' and 'log-out'. Hypothesis 5 is supported, but only in these two measures.

Table 4: Results of hypotheses testing on PMT treatments for each behavioural measure

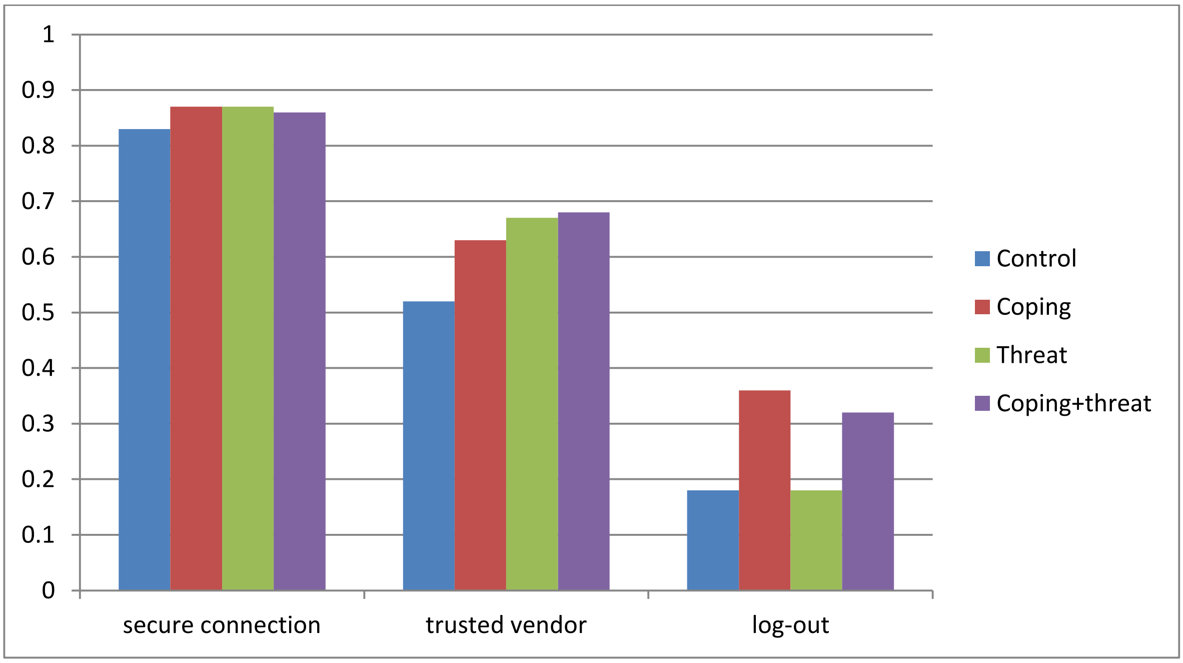
Conditions	**n**	**Mean**	**SD**	**Min - Max**	**t-test**# Treatment vs. control	**t-test**# Treatment vs. treatment
Secure connection:						
Control	507	0.83	0.38	0 – 1	NA	NA
Coping	505	0.87	0.33	0 - 1	**0.0452****	-
Threat	504	0.87	0.34	0 - 1	0.0873	-
Coping + threat	508	0.86	0.35	0 - 1	0.2231	-
Coping vs. coping+threat	-	-	-	-	-	0.4303
Threat vs. coping+threat	-	-	-	-	-	0.6204
Trusted vendor:						
Control	507	0.52	0.50	0 – 1	NA	NA
Cope	505	0.63	0.48	0 – 1	**0.0006*****	-
Threat	504	0.67	0.47	0 – 1	**0.0000*****	-
Coping + threat	508	0.68	0.47	0 – 1	**0.0000*****	-
Coping vs. coping+threat	-	-	-	-	-	0.0861
Threat vs. coping+threat	-	-	-	-	-	0.7224
Password strength:						
Control	507	3.03	1.06	0 – 6	NA	NA
Coping	505	3.44	1.11	0 – 6	**0.0000*****	-

Threat	504	3.08	1.10	1 – 6	0.4463	-
Coping + threat	508	3.51	1.06	1 – 6	**0.0000*** **	-
Coping vs. coping+threat	-	-	-	-	-	0.3305
Threat vs. coping+threat	-	-	-	-	-	**0.0000*** **
Log-out:						
Control	507	0.18	0.38	0 – 1	NA	NA
Coping	505	0.36	0.48	0 – 1	**0.0000*** **	-
Threat	504	0.18	0.39	0 – 1	0.8998	-
Coping + threat	508	0.32	0.47	0 – 1	**0.0000*** **	-
Coping vs. coping+threat	-	-	-	-	-	0.1857
Threat vs. coping+threat	-	-	-	-	-	**0.0000*** **

$^{\#}$ *p*-value

*** *p*<0.01, ** *p*<0.05

Figure 10: Mean scores for the three binary behavioural measures, by PMT condition



4.1.2 Gain vs loss-framed conditions

A gain-framed message was more effective than the control condition in generating secure behaviour, but only for the 'trusted vendor' behavioural measure. Hypothesis 6 is supported only in this measure; in the three remaining measures, there was no statistically significant difference.

The loss-framed message had the same result: it was more effective than the control condition, but only for the 'trusted vendor' condition. Hypothesis 7 is supported when using this behavioural measure, but not the remaining three, where there was no statistically significant difference. Finally, there was no difference between loss and gain-framed conditions for any of the behavioural messages. Hypothesis 8 is not supported.

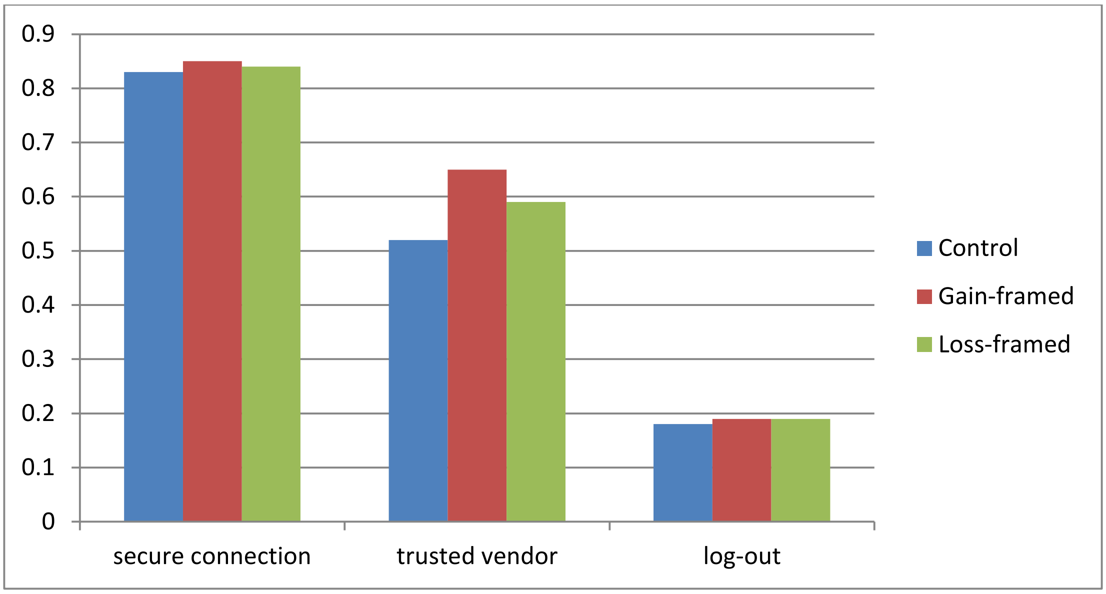
Table 5: Results of hypotheses testing on gain vs loss-framed treatments for each behavioural measure

Conditions	**n**	**Mean**	**SD**	**Min - Max**	**t-test**$^{\#}$ Treatment vs. Control	**t-test**$^{\#}$ Treatment vs. Treatment
Secure connection:						
Control	507	0.83	0.38	0 – 1	NA	NA
Gain-framed	507	0.85	0.36	0 – 1	0.4443	-
Loss-framed	509	0.84	0.36	0 - 1	0.5355	-
Gain vs loss-framed	-	-	-	-	-	0.8839
Trusted vendor:						
Control	507	0.52	0.50	0 – 1	NA	NA
Gain-framed	507	0.65	0.48	0 – 1	**0.0000*****	-
Loss-framed	509	0.59	0.49	0 – 1	**0.0235****	-
Gain vs loss-framed	-	-	-	-	-	0.0591
Password strength:						
Control	507	3.03	1.06	0 – 6	NA	NA
Gain-framed	507	3.10	1.04	0 – 6	0.2942	-
Loss-framed	509	3.04	1.05	1 – 6	0.8363	-
Gain vs loss-framed	-	-	-	-	-	0.3978
Log-out:						
Control	507	0.18	0.38	0 – 1	NA	NA
Gain-framed	507	0.19	0.40	0 –1	0.5185	-
Loss-framed	509	0.19	0.40	0 – 1	0.4872	-
Gain vs loss-framed	-	-	-	-	-	0.9613

$^{\#}$ *p*-value

*** *p*<0.01, ** *p*<0.05

Figure 11: Mean scores for the three binary behavioural measures, by gain or loss-framed condition



4.1.3 Female and male anthropomorphic characters

A female anthropomorphic character had no effect on secure behaviour compared to the control condition, for any of the four behavioural measures. Hypothesis 9 is not supported. Regarding the male anthropomorphic character, the only significant result involved the 'trusted vendor' behavioural measure: a message with a male anthropomorphic character was more effective than the control in making people choose the trusted vendor. Hypothesis 10 is supported for this behavioural measure, but not for the remaining three. Further analysis shows that a male anthropomorphic character is also more effective than the female anthropomorphic character for this measure.

Table 6: Results of hypotheses testing on female and male anthropomorphic character treatments for each behavioural measure

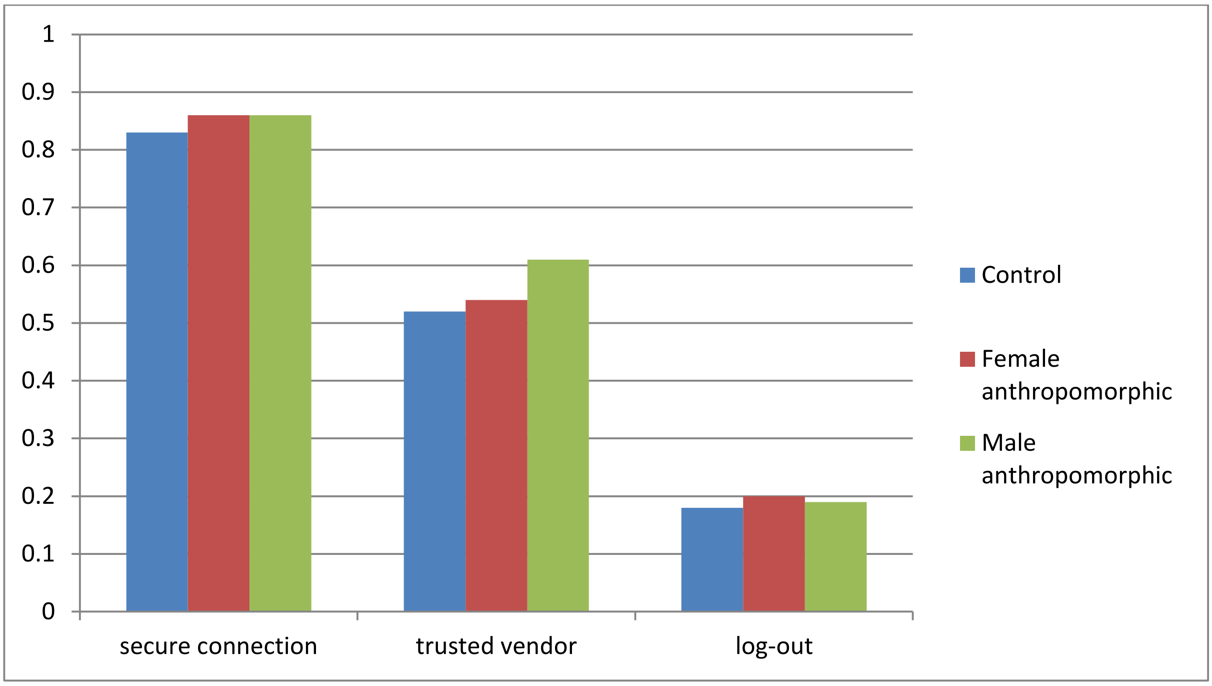
Conditions	**n**	**Mean**	**SD**	**Min - Max**	**t-test**[#] Treatment vs control	**t-test**[#] Treatment vs treatment
Secure connection:						
Control	507	0.83	0.38	0 – 1	NA	-
Female	503	0.86	0.34	0 - 1	0.1304	-
Male	505	0.86	0.34	0 - 1	0.1238	-
Female vs male	-	-	-	-	-	0.9800
Trusted vendor:						
Control	507	0.52	0.50	0 – 1	NA	-
Female	503	0.54	0.50	0 – 1	0.4452	-
Male	505	0.61	0.49	0 – 1	**0.0034*****	-
Female vs male	-	-	-	-	-	**0.0310****

Password strength:						
Control	507	3.03	1.06	0 – 6	NA	-
Female	503	2.99	1.11	0 – 6	0.6012	-
Male	505	3.04	1.06	1 – 6	0.9040	-
Female vs male	-	-	-	-	-	0.5232
Log-out:						
Control	507	0.18	0.38	0 – 1	NA	-
Female	503	0.20	0.40	0 – 1	0.3871	-
Male	505	0.19	0.39	0 – 1	0.6057	-
Female vs male	-	-	-	-	-	0.7272

$^{\#}$ *p*-value

*** $p<0.01$, ** $p<0.05$

Figure 12: Mean scores for the three binary behavioural measures, by female or male anthropomorphic condition



4.1.4 Low-risk, high-impact vs high-risk, low-impact

No difference was found between low-risk, high impact conditions and high-risk, low-impact conditions in any of the four behavioural measures. Hypothesis 11 was not supported.

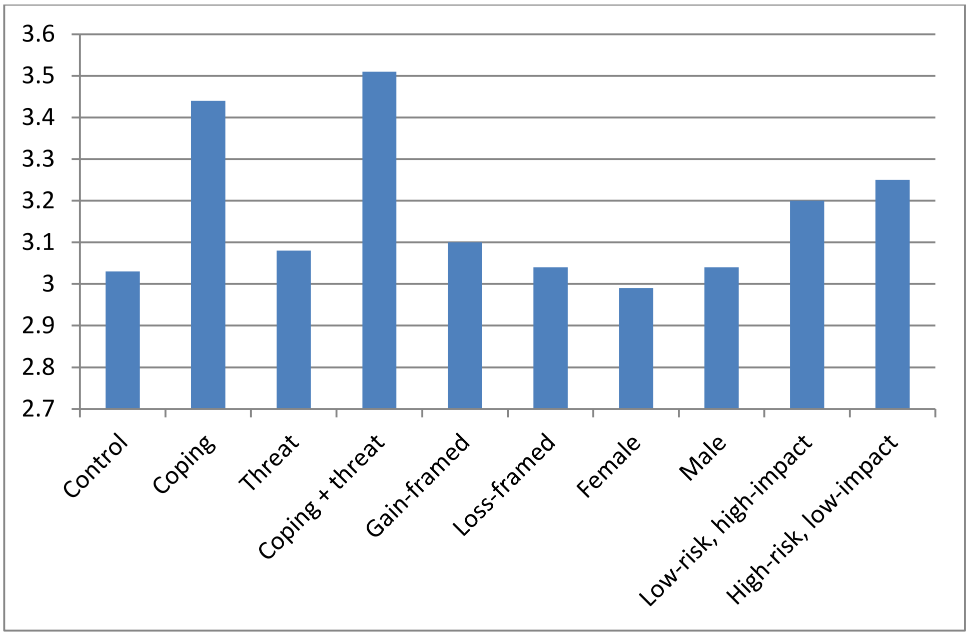
Table 7: Results of hypotheses testing on low-risk, high-impact and high-risk, low-impact treatments for each behavioural measure

Conditions	**n**	**Mean**	**SD**	**Min - Max**	**t-test**[#] Treatment vs. treatment
Secure connection:					
Low-risk, high-impact	511	0.92	0.27	0 - 1	-
High-risk, low-impact	506	0.92	0.28	0 - 1	-
Low-risk, high-impact vs high-risk, low-impact	-	-	-	-	0.9912
Trusted vendor:					
Low-risk, high-impact	511	0.78	0.41	0 – 1	-
High-risk, low-impact	506	0.73	0.44	0 – 1	-
Low-risk, high-impact vs high-risk, low-impact	-	-	-	-	0.0559
Password strength:					
Low-risk, high-impact	511	3.20	1.09	0 – 6	-
High-risk, low-impact	506	3.25	1.09	0 – 6	-
Low-risk, high-impact vs high-risk, low-impact	-	-	-	-	0.4694
Log-out:					
Low-risk, high-impact	511	0.21	0.41	0 – 1	-
High-risk, low-impact	506	0.26	0.44	0 – 1	-
Low-risk, high-impact vs high-risk, low-impact	-	-	-	-	0.0617

[#] *p*-value

The graphs presented above exclude the behavioural measure 'password strength'. The reason for this is that all the other behavioural measures are binary, whereas password strength is measured on a scale from 1 to 6. Figure 13 presents results of participants' score for password strength according to experimental treatment. As discussed earlier, only in *coping appraisal message* and *coping + threat appraisal message* was there a difference with the control group. Quite simply, participants needed to be told what a secure password looked like. The same can easily apply in other realms of cybersecurity behaviour: people cannot be expected to know as much about security behaviour as those who design IT systems.

Figure 13: Mean scores for 'password strength', by female or male anthropomorphic condition



4.1.5 The cybersecurity index: a different perspective

While all of the behavioural measures capture one aspect of secure online behaviour, there is no single measure that summarises how securely a participant behaved throughout the experiment. To overcome this, we proposed a composite behavioural measure. In this 'cybersecurity index', all measures were equally weighted, as there was no evidence *a priori* that any of them should be reinforced. The following formula was applied:

$$Cyber\ security\ index = \frac{secure\ connection + \frac{password\ strength}{6} + trusted\ vendor + log\ out}{4}$$

There are limitations to this index. For one, the assumption that all four behavioural measures should carry the same weight can be challenged. Therefore, we do not propose relying on it in isolation; rather, it should be understood in the context of the study and together with the individual behavioural measures. On the plus side, the index does provide a starting point and invites divergent viewpoints on the matter.

A similar indicator was developed by the authors for a lab experiment in a previous study (Rodriguez-Priego & van Bavel, 2016). However, it cannot be compared to this one, for two reasons. Firstly, it includes one more behavioural measure (secure connection) that was not used as indicator in the lab experiment because of a ceiling effect. Secondly, *password strength* is measured on a scale from 1 to 6 in this experiment, while in the lab experiment it is measured from 1 to 7. In the online experiment, participants did not have to provide an email, so security parameter number 7 of the lab experiment, namely a Boolean search (to see whether the password contained the participant's email), could not be checked.

Results show that, for the PMT treatments, *heightened coping appraisal, heightened threat appraisal* and *combined coping + threat appraisal* all have a positive effect on participant's security behaviour when compared to the control group. Moreover, the *combined coping + threat appraisal* condition is more effective than the *heightened threat appraisal* condition.

For the gain and loss-framed treatments, *gain-framed* led to more secure behaviour than the control condition, but *loss-framed* did not. There was no difference between the effectiveness of *gain-framed* and *loss-framed*.

For the *female* and *male anthropomorphic characters*, only the male character had an effect: it led to a higher score on the cybersecurity index than the control and the female character. The female character had no effect compared to the control condition.

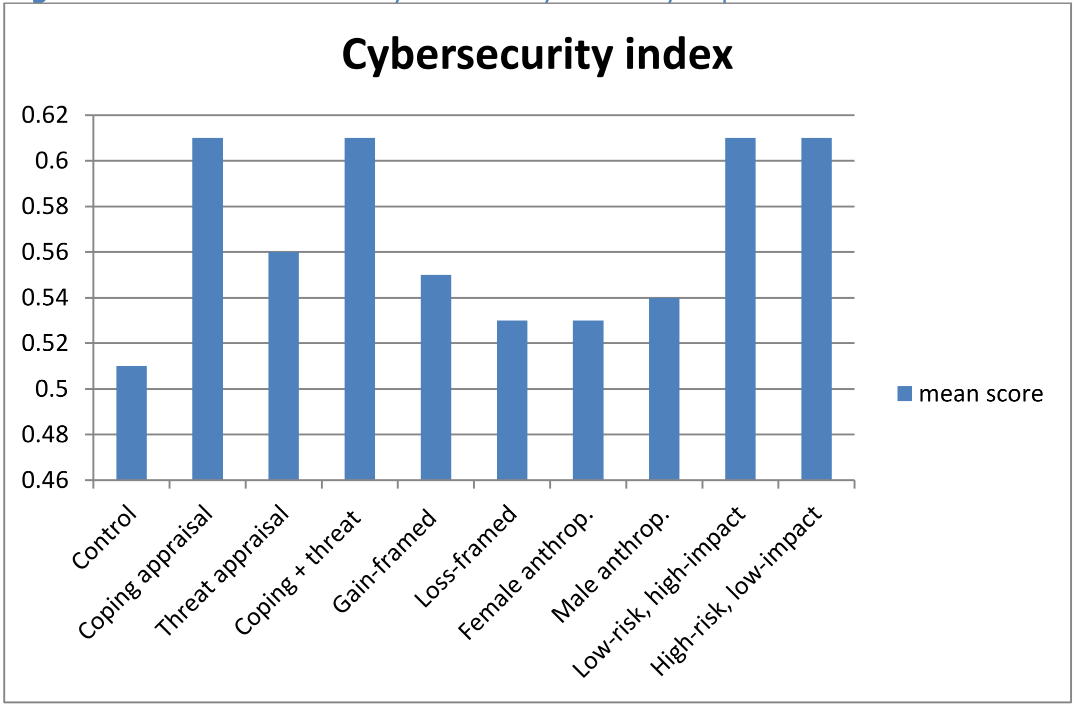
Finally, there was no difference in the scores of the cybersecurity index between the *low-risk, high-impact* condition and the *high-risk, low-impact* condition.

Table 8: Results of all experimental conditions on the cybersecurity index

Conditions	**n**	**Mean**	**SD**	**Min - Max**	**t-test#** Treatment vs. Control	**t-test#** Treatment vs. Treatment
Control	507	0.51	0.22	0.04 – 1	NA	NA
PMT-inspired conditions						
Coping	505	0.61	0.25	0.04 – 1	**0.0000*****	-
Threat	504	0.56	0.20	0.04 – 1	**0.0002*****	-
Coping + threat	508	0.61	0.24	0.04 – 1	**0.0000*****	-
Coping vs coping + threat	-	-	-	-	-	0.9228
Threat vs coping + threat	-	-	-	-	-	**0.0001*****
Gain vs loss-framed conditions						
Gain-framed	507	0.55	0.21	0.04 – 1	**0.0014*****	-
Loss-framed	509	0.53	0.23	0.04 – 1	0.0606	-
Gain vs loss-framed	-	-	-	-	-	0.2126
Female and male anthropomorphic characters						
Female anthropomorphic	503	0.53	0.22	0 – 1	0.1774	-
Male anthropomorphic	505	0.54	0.22	0.04 – 1	**0.0104*****	**-**
Female vs male anthropomorphic	-	-	-	-	**-**	0.2266
Low-risk, high-impact vs high-risk, low-impact						
Low-risk, high-impact	511	0.61	0.20	0.04 – 1	N/A	-
High-risk, low-impact	506	0.61	0.21	0.08 – 1	N/A	-
Low-risk, high-impact vs high-risk, low-impact	-	-	-	-	-	0.9912

p-value, *** *p*<0.01

Figure 14: Mean scores of cybersecurity index by experimental condition



Note: Low-risk, high-impact and high-risk, low-impact conditions are not comparable with the control condition

4.2 Socio-demographic analysis

We conducted demographic analyses to see the effect of country, gender and age on the four behavioural measures and the composite cybersecurity index.

There was no difference by countries on the ‵secure connection' measure. The 'trusted vendor' measure, however, varied considerably: there was a statistical difference between all countries, except between Germany and Spain (see Figure 15). For 'password strength', there was a difference only between Spain and Sweden (3.05 on average vs 3.30, p<0.001), and between Poland and Sweden (3.15 vs 3.30, p<0.001). For 'log-out', only Germany and Spain had significantly different scores (0.27 vs 0.19, p<0.001). Finally, in the cybersecurity index, there was a statistically significant difference between all countries except the following pairings: DE-SE, DE-UK, SP-PL and SE-UK. The pattern of scores for this index resembles that of 'trusted vendor' (Figure 16).

Figure 15: Mean scores for the three binary behavioural measures, by country

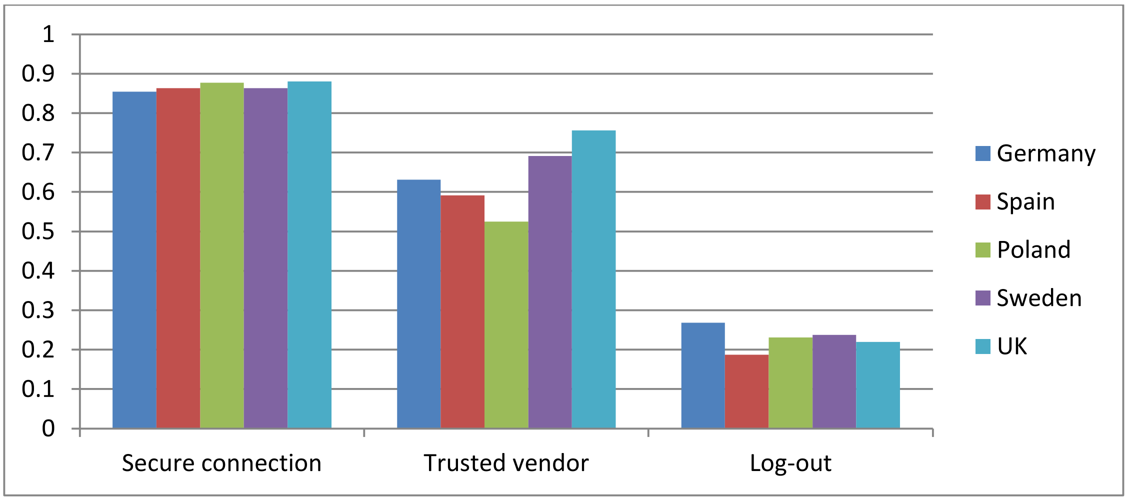
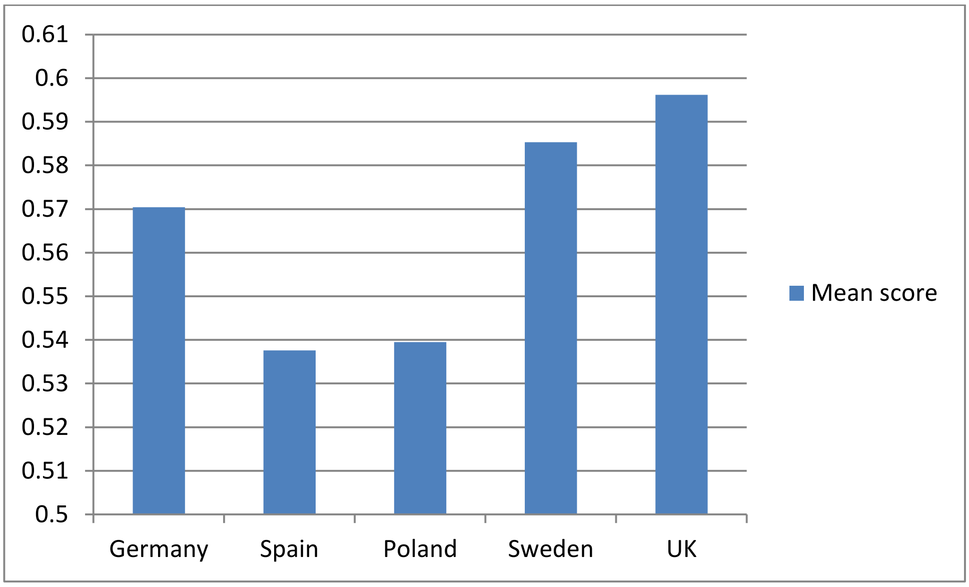
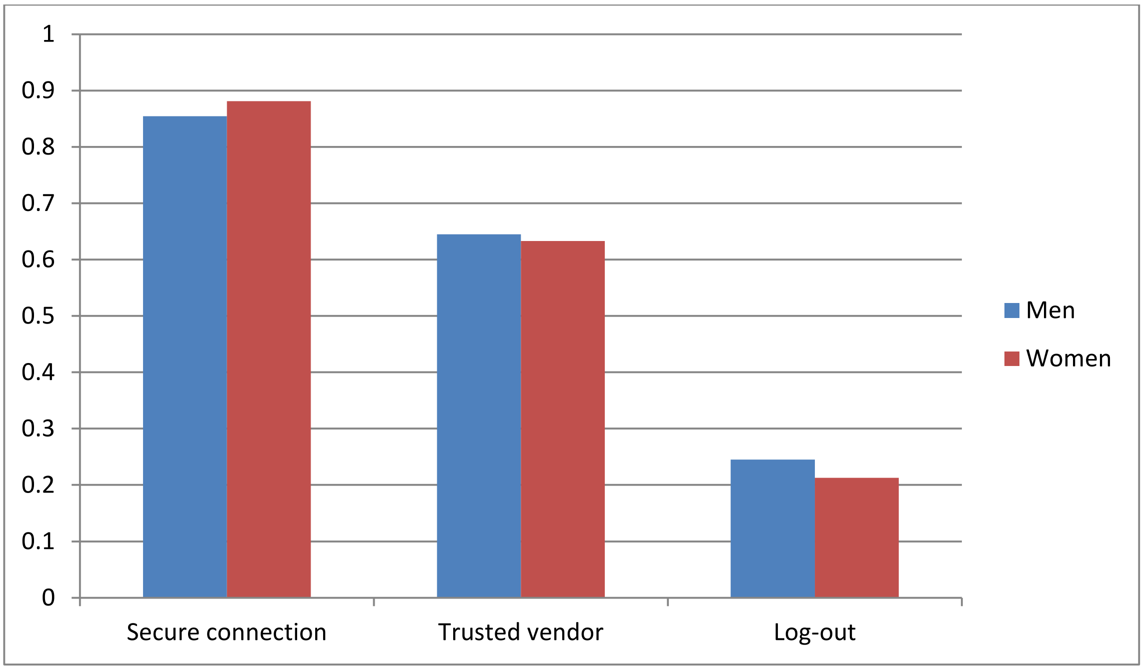


Figure 16: Mean scores for 'cybersecurity index', by country



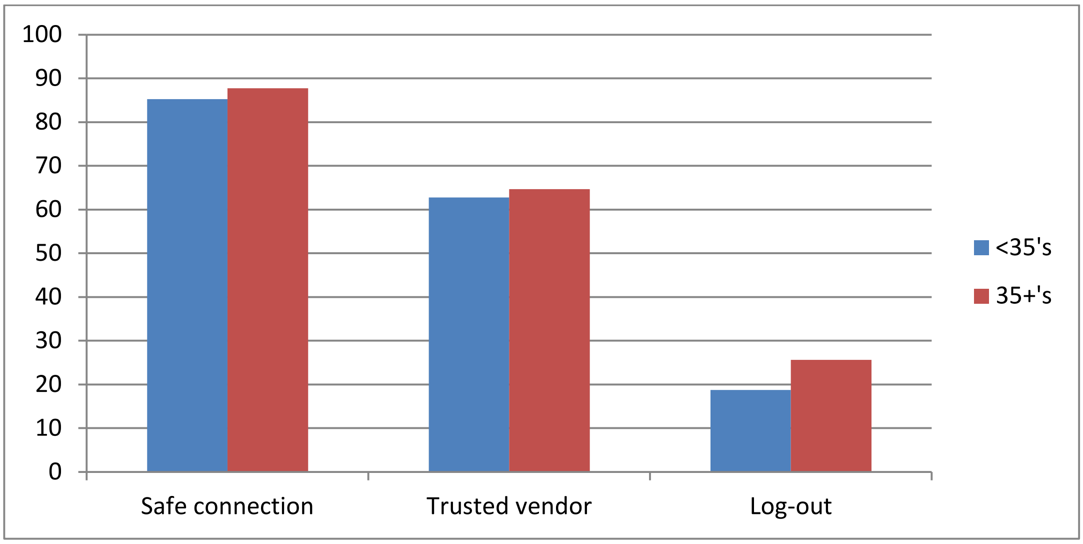
With regard to gender, there was a difference in the mean scores for ‘secure connection'. Men were less likely to choose the secure connection, scoring 0.85 on average vs. 0.89 by women (p<0.001). For 'trusted vendor' there was no difference. For 'log-out' and 'password strength', however, the situation was inverted: men did better than women (log-out: 0.25 vs 0.21, p<0.001; password strength: 3.20. vs 3.14, p<0.05). In the cybersecurity index, these effects cancelled each other out: no statistical difference was found between men and women.

Figure 17: Mean scores for the three binary behavioural measures, by gender



With regard to age, there were differences between participants aged under 35 (<35's) and those aged 35 and over (35+'s) in two behavioural measures. In both cases, older participants were more cautious. In ‘secure connection', <35's scored 0.85 on average vs 0.88 by 35+'s (p<0.001). In 'log-out', <35's scored 0.19 on average vs 0.26 by 35+'s (p<0.001). This difference carried over to the combined 'cybersecurity index', where <35's scored 0.55 on average vs 0.58 by 35+'s.

Figure 18: Percentage of participants who made the 'secure choice' in three binary behavioural measures, by age



4.3 Analysis of questionnaire items

In addition to making a purchase in the mock eCommerce store, participants answered questions on their socio-demographic characteristics, risk aversion, trust online, and knowledge of cybersecurity.

4.3.1 Risk aversion

The risk aversion construct was based on 30 items of the Dospert scale (Blais & Weber, 2006; Weber, Blais & Betz, 2002), and was shown to the participant before the purchasing process began. The items of the construct presented high reliability (Cronbach's alpha: 0.87). Lower values in risk aversion meant that the participant was more risk averse. The purpose of including these items was to find out whether risk aversion was a determining factor in secure online behaviour. Results show that it was.

Risk aversion had a positive and significant effect on: choosing a secure connection, buying from a trusted vendor, and logging-out (see ANOVA and probit results in Table 9). The more risk averse the participant, the more likely they would be to exhibit secure behaviour according to these measures. There are no results for the measure 'password strength' (although the ANOVA provides a p-value under 0.05, the ordered probit regression shows there is no effect).

Table 9: Results from ANOVA and ordered probit regression for the risk aversion construct

	ANOVA	**Probit**
Secure connection	7.4433756 (0.000)	-.230062 **(0.000)**
Trusted vendor	5.510104 (0.000)	-.0967589 **(0.001)**
Password strength	12.297246 0.034	-.0077582 0.751
Log-out	2.8910641 0.003	-.1005444 **(0.002)**

Note: Partial SS values reported for ANOVA (p-value in parenthesis); coefficient values for the ordered probit (p-values in parenthesis).

4.3.2 Trust online

The trust online construct was composed of 6 questions (McKnight, Choudhury & Kacmar, 2002), which were asked directly after the purchasing process was completed. The construct displayed high reliability: with a Cronbach's alpha score of 0.89 and average inter-item covariance of 0.568. The items are presented in the Annex.

The only behavioural measure which correlated with trust online was 'secure connection' (partial SS: 1.312, p-value: 0.022). However, a further probit regression does not confirm this effect (p = 0.546). For all other behavioural measures, there was no significant effect.

4.3.3 Knowledge

A number of items included in the post-purchase questionnaire were related to knowledge (see Annex for a full list). We were interested in whether the experimental conditions had an impact on knowledge in parallel to their effect on behaviour. If they did not affect knowledge, it would be an indicator that the nudge worked by affecting System 1, side-stepping System 2 (Kahneman, 2011). If, on the other hand, knowledge *was* affected, we could reasonably assume that the nudge worked by fortifying participants' deliberative capacity.

Results show that one knowledge variable was affected by experimental conditions: *knowledge_logout.* This variable tested whether participant knew that logging out could help them prevent a cyberattack. *Coping threat appraisal* and *combined coping + threat appraisal* had a significant effect on this variable (see Table 10). This was probably due to the warning messages in these conditions, both of which made explicit reference to logging out. On the other hand, they also made reference to choosing secure connections and using secure passwords, which were behavioural measures that showed no effect.

Table 10: Ordered probit regression to test the effect of the treatments on *Knowledge_logout*

Treatments	Coef.	Std.Err	z	P>|z|	[95% Conf.	Interval]
Gain-framed	.0005365	.0702144	0.01	0.994	-.1370812	.1381543
Loss-framed	.0643528	.0705355	0.91	0.362	-.0738942	.2025998
Female anthropomorphic	.0417738	.0707261	0.59	0.555	-.0968468	.1803945
Male anthropomorphic	.045672	.0705185	0.65	0.517	-.0925417	.1838857
Coping appraisal	.1756793	.0713878	2.46	**0.014**	.0357618	.3155969
Threat appraisal	.0457393	.0706855	0.65	0.518	-.0928018	.1842803
Coping + threat appraisal	.2281604	.0716854	3.18	**0.001**	.0876597	.3686612

Note: Control group is the baseline (low-risk, high-impact and high-risk, low-impact conditions not included as they are not comparable to the control group). Number of observations = 4,048; LR chi2(7)= 18.43; prob > chi2 = 0.0102; log likelihood = -4806.4388; pseudo R2 = 0.0019.

5 Discussion

The fact that the *heightened coping appraisal* condition was the most effective (it had an impact on all four behavioural measures) is one of the main results of the study. It suggests that the reason users often fail to behave securely is not necessarily because they do not care, or because they are unaware of the risks, but rather because they simply do not know what secure behaviour entails. Giving them specific instructions, and reminding them that it is easy and within their grasp, is therefore an effective way of generating secure behaviour. This clearly has implications for policies that seek to make online transactions more secure.

The *heightened threat appraisal* condition was relatively less effective. But a note of caution is needed here. The threat of introducing a virus onto participants' computers may simply have been too realistic. The warnings in the other experimental conditions were limited to damage within the confines of the experiment, i.e. the risk that participants could lose their fee. This may have contributed to an excessive drop-out rate in this condition. Perhaps, those participants who were most susceptible to the message dropped out, leaving the more resilient ones to complete the experiment. This potentially biased sample might have understated the strength of the warning message.

The *combined coping + threat appraisal* was also effective, but in fewer behavioural measures than *heightened coping appraisal* (three out four)*.* Hypothesis 4 suggested that, due to the added effect of the *heightened threat appraisal* message, the *combined coping + threat appraisal* should be more effective than the *heightened coping appraisal*. However, there is no evidence of a statistically significant difference in the means of the three behavioural measures affected by these conditions.

Hypothesis 5 applied the same logic to *heightened threat appraisal* and *combined coping + threat appraisal*. Results show that the latter is more effective than the former, thanks to the addition of the heightened coping appraisal message. While the heightened threat appraisal condition had no impact on any behavioural measure, the combined coping + threat appraisal condition had an impact on three. In sum, *heightened coping appraisal* was effective in generating secure behaviour as a nudge on its own. Moreover, adding it to *heightened threat appraisal* made this latter condition effective, whereas, on its own, it was not.

One question remains. Why, if *heightened coping appraisal* had an impact on all four behavioural measures, did *combined coping + threat appraisal* fail to have an effect on 'secure connection'? In no measure did *heightened threat appraisal* have a negative effect on secure behaviour (not even a statistically non-significant one). So why would its combination with *heightened coping threat* appraisal erode the latter's effectiveness? The most plausible explanation would seem to be that it simply made the message too long, which discouraged participants from reading it fully.

Regarding gain and loss-framed conditions, results show that they both had a positive effect on the 'trusted vendor' behavioural measure. They also differed significantly from the control condition in the dropout rate (participants in these conditions were less likely to drop out after seeing the warning message than in the control). But there was no evidence of any difference between a loss-framed message and a gain-framed one. This result implies that reinforcing a warning message with information on what can be gained or lost will contribute to changing behaviour: it will act as an incentive to stay logged on, and will also lead people to be more cautious when selecting a vendor.

Of the male and female anthropomorphic characters, only the male anthropomorphic character had an effect, and only on one measure ('trusted vendor'). Further research needs to be undertaken to better understand the role of these characters on online behaviour. Finally, the comparison of a message that gave a high probability of losing a small amount with one that gave a low probability of losing a large amount yielded no results.

In sum, the experimental conditions based on insights from protection motivation theory (PMT) were the most effective. PMT therefore holds promise as a theoretical underpinning for policies which seek to make people behave more securely.

Of the behavioural measures, one appears to have been more sensitive to the experimental conditions: 'trusted vendor'. It was affected by all PMT-inspired conditions, the gain and loss-framed message conditions and the male anthropomorphic character condition. The reasons behind this are not immediately clear and cannot be explained convincingly by the empirical data provided here. However, they might have a cultural element to them. 'Trusted vendor' was the measure that showed the greatest variation across countries (there were differences between all countries, except Germany and Spain). It could be that, when faced with the option of paying for something vs getting something for free, the risk of unsecured behaviour becomes more apparent. In other words, exposure to cybersecurity threats may be perceived as the price for getting something for free.

6 Conclusion

This study applies experimental methodology to the study of cybersecurity behaviour. It forms part of a wider trend which applies behavioural insights to policy-making (van Bavel, Herrmann, Esposito & Proestakis, 2013; Lunn, 2014; World Bank, 2015; Obama, 2015; Lourenço, Ciriolo, Almeida & Troussard, 2016). This research has effectively explored the role of changes to the *choice architecture* on online decision-making, or *nudging* (Sunstein, 2013; Thaler & Sunstein, 2008).

It has shown that warning messages built on established behavioural insights in the specific literature, such as those inspired by PMT or gain vs loss-framing, can be effective. It has also shown that some behavioural measures (such as 'trusted vendor') are more susceptible to experimental conditions, and therefore more malleable. Risk aversion has been shown to correlate with security behaviour, and that the nudges have only a limited effect on knowledge of cybersecurity. Finally, security behaviour will vary from country-to-country significantly, and show some differences according to age (older participants being more cautious). Regarding gender, results are inconclusive.

Perhaps one of the main contributions of this study is its emphasis on observed, actual behaviour. Much of the evidence on online behaviour is limited to intention or to self-reported behaviour. In both these cases, the gap with actual behaviour can be significant. Using experimental methodology to test behavioural insights on actual behaviour is promising and could be an effective tool for testing proposed policy interventions. It may not provide all the answers, but it is a robust method for building a body of reliable data on which to base policy.

Bibliography

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221), 509-514.

Anderson, C. L., & Agarwal, R. (2010). Practicing safe computing: a multimedia empirical examination of home computer user security behavioral intentions. *Mis Quarterly*, *34*(3), 613-643.

Bente, G., Dratsch, T., Rehbach, S., Reyl, M., & Lushaj, B. (2014). Do you trust my avatar? Effects of photo-realistic seller avatars and reputation scores on trust in online transactions. In *HCI in Business* (pp. 461-470). Springer International Publishing.

Boer, H., & Seydel, E.R. (1996). Protection motivation theory. In M. Connor and P. Norman (Eds.) *Predicting Health Behavior*. Buckingham: Open University Press.

Crossler, R.E., Johnston, A.C., Lowry, P.B., Hu, Q., Warkentin, M. & Baskerville, R. (2013). Future directions for behavioural information society research. *Computer Security*, 32, 90-101.

Crossler, R. E., Long, J. H., Loraas, T. M., & Trinkle, B. S. (2014). Understanding compliance with bring your own device policies utilizing protection motivation theory: Bridging the intention-behavior gap. *Journal of Information Systems*, *28*(1), 209-226.

European Commission (2016). *European Digital Progress Report*. Available at https://ec.europa.eu/digital-single-market/en/news/commission-releases-2016-european-digital-progress-report-unequal-progress-towards-digital.

Executive Order No. 13707, 3 C.F.R. 56365 (2015).

Groom, V., & Calo, R. (2011, September). Reversing the Privacy Paradox: An Experimental Study. TPRC.

Heckman, C. E., & Wobbrock, J. O. (2000, June). Put your best face forward: Anthropomorphic agents, e-commerce consumers, and the law. In *Proceedings of the fourth international conference on Autonomous agents* (pp. 435-442). ACM.

Herath, T., & Rao, H. R. (2009). Protection motivation and deterrence: a framework for security policy compliance in organisations. *European Journal of Information Systems*, *18*(2), 106-125.

Johnston, A. C., & Warkentin, M. (2010). Fear appeals and information security behaviors: an empirical study. *MIS quarterly*, 549-566.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, 263-291.

Lee, D., Larose, R., & Rifon, N. (2008). Keeping our network safe: a model of online protection behaviour. *Behaviour & Information Technology*, *27*(5), 445-454.

Lee, Y. (2011). Understanding anti-plagiarism software adoption: An extended protection motivation theory perspective. *Decision Support Systems*, *50*(2), 361-369.

Liang, H., & Xue, Y. (2010). Understanding Security Behaviors in Personal Computer Usage: A Threat Avoidance Perspective*. *Journal of the Association for Information Systems*, *11*(7), 394.

Lourençο, J.S., Ciriolo, E., Almeida, S.R. & Troussard, X. (2016). Behavioural insights applied to policy: European report 2016. EUR 27726 EN; doi: 10.2760/707591.

Lunn, P. (2014). Regulatory Policy and Behavioural Economics. OECD Publishing.

Maddux, J.E., & Rogers, R. W. (1983). Protection motivation theory and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology,* 19, 469-479.

Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research*, 26(4), 323-339.

Neuwirth, K., Dunwoody, S., & Griffin, R. J. (2000). Protection motivation and risk communication. *Risk Analysis*, *20*(5), 721-734.

Obama, B. (2015). *Executive order – Using behavioural science insights to better serve the American people* (Executive Order 13707). Washington, DC: The White House. Retrieved from: https://www.whitehouse.gov/the-press-office/2015/09/15/executive-order-using-behavioral-science-insights-better-serve-american on 15 November 2016.

Qiu, L., & Benbasat, I. (2009). Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. *Journal of Management Information Systems*, *25*(4), 145-182.

Rodriguez-Priego, N. & van Bavel, R. (2016). The effect of warning messages on secure behaviour online: Results from a lab experiment in e-Commerce. *JRC Technical Reports*.

Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. *The journal of psychology*, *91*(1), 93-114.

Rogers, R.W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In J. Cacioppo & R. Petty (Eds.), *Social Psychophysiology*. New York: Guilford Press.

Shillair, R., Cotten, S. R., Tsai, H. Y. S., Alhabash, S., LaRose, R., & Rifon, N. J. (2015). Online safety begins with you and me: Convincing Internet users to protect themselves. *Computers in Human Behavior*, *48*, 199-207.

Siponen, M., Mahmood, M. A., & Pahnila, S. (2014). Employees' adherence to information security policies: An exploratory field study. *Information & management*, *51*(2), 217-224.

Stevens, M., Chisnell, D., Sasse, A., Krol, K., Theofanos, M. & Wald, H. (2014) *Report: Authentication Daily Study* (National Institute for Standards and Technology), NISTIR 7983. Available at http://dx.doi.org/10.6028/NIST.IR.7983.

Sunstein, C. R. (2013). Behavioral economics, consumption, and environmental protection. *Forthcoming in Handbook on Research in Sustainable Consumption (Lucia Reisch & John Thøgersen eds.)*.

Thaler, R. & Sunstein, C. (2008). *Nudge*: *Improving decisions about health, wealth, and happiness*. London: Penguin.

van Bavel, R., Herrmann, B., Esposito, G., & Proestakis, A. (2013). *Applying Behavioural sciences to EU policy-making, JRC Scientific and Policy Reports EUR 26033 EN .* http://ec.europa.eu/dgs/health_consumer/information_sources/docs/30092013_jrc_scientific_policy_report_en.pdf.

Waldrop, M. M. (2016). How to hack the hackers: The human side of cybercrime. *Nature*, 533, 164.

Woon, I., Tan, G. W., & Low, R. (2005). A protection motivation theory approach to home wireless security. *ICIS 2005 proceedings*, 31.

Workman, M., Bommer, W. H., & Straub, D. (2008). Security lapses and the omission of information security measures: A threat control model and empirical test. *Computers in Human Behavior*, *24*(6), 2799-2816.

World Bank (2015). *World Development Report 2015: Mind, Society, and Behaviour* (http://www.worldbank.org/en/publication/wdr2015).

Youn, S. (2005). Teenagers' perceptions of online privacy and coping behaviors: a risk–benefit appraisal approach. *Journal of Broadcasting & Electronic Media*, *49*(1), 86-110.

Annex: Questionnaire items

Table 11: Risk aversion

Construct	Question	Answer
Risk aversion	For each of the following statements, please indicate the likelihood that you would engage in the described activity or behaviour if you were to find yourself in that situation. Provide a rating from Extremely Unlikely to Extremely Likely, using the following scale: 1. Admitting that your tastes are different from those of a friend. 2. Going camping in the wilderness. 3. Betting a day's income at a casino. 4. Investing 10% of your annual income in a moderate growth mutual fund. 5. Drinking heavily at a social function. 6. Taking some questionable deductions on your income tax return. 7. Disagreeing with an authority figure on a major issue. 8. Betting a day's income at a high-stake poker game. 9. Having an affair with a married man/woman. 10. Passing off somebody else's work as your own. 11. Going down a ski run that is beyond your ability. 12. Investing 5% of your annual income in a very speculative stock. 13. Going white-water rafting at high water in the spring. 14. Betting a day's income on the outcome of a sporting event. 15. Engaging in unprotected sex. 16. Revealing a friend's secret to someone else. 17. Driving a car wearing a seat belt (reversed item). 18. Investing 10% of your annual income in a new business venture. 19. Taking a skydiving class. 20. Riding a motorcycle without a helmet. 21. Choosing a career that you truly enjoy over a more secure one. 22. Speaking your mind about an unpopular issue in a meeting at work. 23. Sunbathing without sunscreen. 24. Bungee jumping off a tall bridge. 25. Piloting a small plane. 26. Walking home alone at night in an unsafe	Scale from [1] Extremely Unlikely to [5] Extremely Likely.

	area of town. 27. Moving to a city far away from your extended family. 28. Starting a new career in your mid-thirties. 29. Leaving your young children alone at home while running an errand. 30. Returning a wallet you found that contains €200 (reversed item).	

Table 12: Trust in the online environment

Construct	**Question**	**Answer**
Trust online	Please, choose in the table below the level of agreement or disagreement with the statements listed: 1. I am comfortable making purchases or other activities on the Internet 2. I feel that most Internet vendors would act in a customers' best interest. 3. I am comfortable relying on Internet vendors to meet their obligations. 4. I feel fine doing business on the Internet since Internet vendors generally fulfil their agreements. 5. In general, most Internet vendors are competent at serving their customers. 6. I feel confident that encryption and other technological advances on the Internet make it safe for me to do business there.	Scale from [1] Strongly agree [5] Strongly disagree.

Table 13: Perceived knowledge

Construct	**Question**	**Answer**
Perceived knowledge	How well informed do you feel about the risks of cybercrime?	1. Not at all informed. 2. Not very well informed. 3. Somewhat informed. 4. Fairly well informed. 5. Very well informed.

Table 14: Knowledge

Construct	**Question**	**Answer**
Knowledge	Which of the following behaviours do you think can help you prevent from being attacked while online?	Provide a rating from [1] It won't reduce my risk at all to [5] It will reduce my risk extremely
Knowledge_safe	Connecting to a trusted connection.	
Knowledge_pswd1	Using a strong password.	
Knowledge_pswd2	Changing your password frequently.	
Knowledge_pswd3	Avoid using the same password for different sites.	
Knowledge_signup	Providing minimum information.	
Knowledge_trust	Connecting to a trusted site.	
Knowledge_logout	Logging out.	
Knowledge_soft1	Using anti-virus software and firewalls.	
Knowledge_soft2	Updating software to the latest version.	
Knowledge_public	Avoiding access to my personal accounts in public places.	

List of figures

List of tables

LF-NA-28197-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2791/2476

ISBN 978-92-79-63487-1