



**HAL**  
open science

# ”Phishing For (quantum-like) Phools” Theory and experimental evidence

Ariane Lambert-Mogiliansky, Adrian Calmettes

► **To cite this version:**

Ariane Lambert-Mogiliansky, Adrian Calmettes. ”Phishing For (quantum-like) Phools” Theory and experimental evidence. 2019. halshs-02146862

**HAL Id: halshs-02146862**

**<https://shs.hal.science/halshs-02146862>**

Preprint submitted on 20 Jun 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



PARIS SCHOOL OF ECONOMICS  
ÉCOLE D'ÉCONOMIE DE PARIS

WORKING PAPER N° 2019 – 25

**"Phishing For (quantum-like) Phools"  
Theory and experimental evidence**

Ariane Lambert-Mogiliansky  
Adrian Calmettes

JEL Codes:  
Keywords :

# ”Phishing For (quantum-like) Phools”<sup>\*</sup>

## Theory and experimental evidence <sup>†</sup>

A. Lambert-Mogiliansky<sup>‡</sup> A. Calmettes<sup>§</sup>

May 28, 2019

### Abstract

Quantum-like decision theory is by now a theoretically well-developed field (see e.g., Danilov et al. 2018A). We here test the predictions of an application of this approach to persuasion. One remarkable result entails that in contrast to Bayesian persuasion, distraction rather than relevant information has a powerful potential to influence decision-making.

We first provide a quantum decision model for a choice between two uncertain alternatives. We derive the impact of persuasion by means of distractive questions and contrast them with the predictions of the Bayesian model. Next, we conduct an experiment where respondents choose between supporting either one of two projects to save endangered species. We test the impact of persuasion in the form of questions related to different aspects of the uncertain value of the two projects. The experiment involves 900 respondents divided into three groups: a control group, a first treatment group and the distraction treatment group. Our main result is that, in accordance with the predictions of quantum persuasion but in violation with the Bayesian model, distraction significantly affects decision-making. Population variables play no role. Some significant variations between subgroups are exhibited and discussed. The results of the experiment provide support for the hypothesis that the manipulability of people’s decision-making can be explained by the quantum indeterminacy of their subjective representations of reality.

## 1 Introduction

Why is the famous P&G’s (Procter and Gambles) ”Thanks mum” advertisement (2010) showing devoted mothers supporting young athletes, among the most successful ads of all time? Why does Samsung feature blind children ”watching” TV in a remote place to advertise for their services? Why does KIA

---

<sup>\*</sup>This expression is borrowed from Akerlof and Shiller’ book titled ”Phishing for Phools - the Economics of manipulation and deception.”2015.

<sup>†</sup>We would like to thank Jerome Busemeyer for a very valuable suggestion on the design of the experiment as well Frederic Koessler and other seminar participants at Paris School of Economics for their valuable comments.

<sup>‡</sup>Paris School of Economics, alambert@pse.ens.fr

<sup>§</sup>Institute of Economics, Scuola Superiore Sant’Anna, adri.calmettes@hotmail.fr

feature a woman comically failing to save a tree to advertise for a hybrid car?<sup>1</sup> The pervasiveness of informationally irrelevant messaging in advertising is stunning. In this paper we test a theory that explains why distraction i.e., addressing informationally irrelevant issues can be a powerful manipulation technic.

The idea that people are being influenced and manipulated by a systematic exploitation of non rational psychological factors rather than by providing information that is rationally processed, was first forcefully put forward in the seminal book of Vance Packard (1957) "The Hidden Persuador". His thesis is that persuadors rely on psychiatric and psychological technics to address their message to our "wild and unruly subconscious". Later Cialdini developed a "science of persuasion" based on a general behavioral principles (e.g., bias for reciprocity) that can be exploited to influence people's choice (see e.g., [4, 21]). Closer to our approach which focuses on information processing, are early works by Festinger and Maccoby (1964). They published the first experiment showing that distraction can induce attitude change: a message has larger persuasion power among respondents subjected to distraction. Their idea is that distraction in the course of information processing makes attempts to provide counter arguments less successful. This in turn makes people more vulnerable to the persuador's message.<sup>2</sup>

More recently Akerlof and Shiller (2015) provide loads of evidence showing that people are systematically "phished" in economic transactions. The authors suggest that this is due to the significance of the story people tell themselves when making decision. Returning to our "Thanks mum" ad example, putting forward mothers' invaluable support to young athletes suggests a story of proud achievements. The authors write that stories are equivalent to people's focus of mind. Thus, the ad induces consumers to change the focus of their mind from an appreciation of house cleaning products to the pride of being a mother - who also keeps the house clean. Irrelevant information induces a change in focus which eventually affects decision-making. P&G experienced a sharp increase in sales.

Consumer sovereignty is a corner stone of neoclassical economics celebrated in the first theorem of welfare. It relies on an assumption that individuals have stable preferences and consistent beliefs. Indeed individuals are defined by their preferences and beliefs. But Akerlof and Shiller show that manipulation of preferences is not a marginal but rather a systematic feature of the functioning of our economies - the subtitle of their book is "The economics of manipulation and deception". What are the consequences for the welfare properties of the market economy? To answer that question one needs an alternative theory of preferences. Danilov et al. ([8]) provide a theory of decision-making for individuals endowed with contextual preferences. It relies on a single departure from the standard model by proposing that preferences are intrinsically indeterminate and therefore endogenous to interaction. As shown in [10], this model when applied to the issue of persuasion delivers Akerlof and Shiller's idea according to which "just change people's focus and you can change the decisions they make" (p.173 [1]). Our aim in this paper is to experimentally test a central prediction of that theory.

---

<sup>1</sup>The two last examples are the most watched ads in 2017.

<sup>2</sup>For a review on psychological literature on the relationship between distraction and persuasion, see [2].

The theory of Bayesian persuasion was initiated by Kamenica and Gentzkow [17] and further developed in a variety of directions. The subject matter of the theory of persuasion is the use of an information structure (or measurement, questions) that generates new information in order to modify a person’s state of beliefs with the intent of making her act in a specific way. The question of interest is how much can a person, call him Sender, influence another one, call her Receiver, by selecting a suitable measurement and revealing its outcome. An example is in lobbying. A pharmaceutical company commissions to a scientific laboratory a specific study of a drug impact, the result of which is delivered to the regulator. The question of interest from a persuasion point of view is what kind of study best serves the company’s interest.

Receiver’s decision depends on her beliefs about the world. A central assumption in [17] and related works, is that uncertainty is formulated in the standard classical framework as a probability distribution over a set of states of the world. As a consequence Receiver follows Bayes’ rule to update her beliefs.

However as amply documented the functioning of the mind is more complex and often people do not follow Bayes rule. Cognitive sciences propose alternatives to Bayesianism. One avenue of research within cognitive sciences appeals to the mathematical formalism of Quantum Mechanics (QM). A main reason is that QM has properties that reminds of the paradoxical phenomena exhibited in human cognition. In addition, Quantum cognition has been successful in explaining a wide variety of behavioral phenomena such as disjunction effect, cognitive dissonance or preference reversal (see [15], [3]). Moreover there exists by now a fully developed decision theory in the context of non-classical (quantum) uncertainty (see [7, 8]). Clearly, the mind is likely to be even more complex than a quantum system but our view is that the quantum cognitive approach already delivers interesting new insights in particular with respect to persuasion.

In quantum cognition, the object of interest is the decision-maker’s mental representation of the world. It is modelled as a quantum-like system represented by its state - a *cognitive state* which is the equivalent of beliefs in the classical context. The decision relevant uncertainty is therefore of non-classical (quantum) nature. As argued in [11] this modelling approach allows capturing widespread cognitive limitations in information processing when constructing and revising the mental representation of a ‘complex’ alternative (cf [6]). The key quantum property that we use is that some characteristics of a complex mental object may be “Bohr complementary” that is they are incompatible in the decision-maker’s mind: they cannot have a definite value simultaneously. The decision-maker processes information sequentially and order matters. This reflects the fact that a measurement (e.g., a question) modifies the cognitive state in a non-Bayesian well-defined manner.

As in the classical context our rational Receiver uses new information to update her beliefs. Choices based on updated preferences are consistent with ex-ante preferences defined for the condition (event) that triggered the updating.<sup>3</sup> In [8], we learned that a dynamically consistent quantum-like decision-

---

<sup>3</sup>A rational quantum-like decision-maker is a decision-maker who has preferences over items (or actions) that she perceives (represents) as quantum systems. Her preferences over items with uncertain properties (or actions with uncertain consequences), satisfy a number of axioms similar to those in the classical case as we show in [8]. These axioms secure that the

maker updates her beliefs according to Lüders' postulate.<sup>4</sup> In two recent papers, important theoretical results were established. First, it is shown in [9] that in the absence of any constraints on measurements, full persuasion applies: Sender can always persuade Receiver to believe anything that he wants. Next, in [10] the same authors investigate a short sequence of measurements but in the frame of a simpler task that they call "targeting". The object of "targeting" is the transition of a belief state into another specified target state. The main result of relevance to our issue is that distraction i.e., a test or question that generates irrelevant but "Bohr complementary" information has significant persuasion power. A Bayesian decision-maker does not update her beliefs when the information is not relevant to her concern and thus cannot be persuaded in this manner to change her decision.

The present paper aims at testing experimentally the result according to which distraction, defined as a question related to a perspective that is incompatible with the perspective relevant for the decision at stake, can significantly affect decision-making.

The situation that we consider is the following. People are invited to choose between two projects aimed at saving endangered species (Elephants and Tigers). The selected project will receive a donation of 50 euros (one randomly selected respondent will determine the choice). We consider two perspectives of relevance for the choice: the urgency of the cause and the honesty of the organization that manages the donations. As a first step and in a separate experiment we establish that the two perspectives are incompatible by exhibiting a significant order effect which is the signature of incompatible measurements (see [5]). In the main experiment 900 respondents are divided into three groups: a control and two treatment groups. They all go through a presentation of the projects and some questions about their preferences. The difference between the groups is that the first treatment group is invited to answer a question about their beliefs of direct relevance to their choice while the second must answer a question that distracts them from what is relevant to their choice.

At the population level, the results are in accordance with the predictions of the quantum model: the distractive question has a significant impact on the respondents choices as compared with the control. The (introspective) question on decision relevant beliefs has no significant impact compared to the control group. None of the population variables has any impact. We also find some significant variation between subgroups with respect to their responsiveness to distraction. People who care about urgency are much less responsive. This is consistent with the quantum model under the reasonable assumption that those people have more extreme ("end of the world") preferences than those who care about honesty. Finally the impact of the order of presentation of the two projects is discussed by noting that when it coincides with preferences, it is related to a larger responsiveness to distraction.

We conclude with a discussion on rationality in information processing and relate our approach to prominent behavioral theories.

---

preferences can be represented by an expected utility function.

<sup>4</sup>Von Neuman's and the more general Lüder's postulate are building stones in Quantum Mechanics. They are the formula for updating the state following the performance of a measurement.

### *Related literature*

This paper contributes to the literature on Bayesian persuasion and more particularly on its recent development which introduces various kind of imperfections in information processing by Receiver. One strand of literature remains within the frame of the standard rational agent paradigm. An example is the recent work by Boedel and Segal "Persuasion with rational inattention" (2018). They show how Sender optimally exploits Receiver's inattention with complex signals when objectives are misaligned. Another example that is clearly behavioral, is Lipnowsky and Mathevet ([20]). They assume that beliefs directly enter Receiver's utility function in the spirit of psychological games ([13]). Their focus is on how Sender responds to Receiver's problem with temptation and self-control by adapting the signal structure. Our approach is also behavioral since we introduce cognitive limitations (incompatibility of perspectives) in information processing with the consequence that updating follows Lüders' rule instead of Bayes' rule thereby creating a special role for distraction.

## **2 Quantum persuasion**

Let us first briefly describe the classical approach developed by Kamenica and Gentzkow ([17]) called Bayesian persuasion. We have two players Sender and Receiver. Receiver chooses an action among a set of alternatives with uncertain consequences. An action yields consequences for both players. Sender may try to influence Receiver so she chooses an action that is most valuable to him. A crucial element of the Bayesian persuasion approach is that Sender does not choose the information Receiver obtains. If he did that would raise issues of strategic concealment and revelation. Instead Sender chooses an "information structure" (IS) or a measurement that is a test, an investigation or a question. Sender is committed to truthfully reveal the outcome of the IS (e.g., he does not control the entity that produces the study). For instance, an IS can be an investigation by an independent lab on the impact of a new soft drink on teenager's cognitive performance. The outcome may be "no impact, negative impact, positive impact". Another example, closer to our application here, is a question to Receiver: do you believe (Yes or No) that politicians' climate inaction will lead to global catastrophe under this century? The outcome of any IS is information. In our examples above, it is information about the effect of the soft drink or about the opinion(beliefs) of Receiver on the responsibility of politicians. This information generally affects Receiver's beliefs which in turn may affect her evaluation of the uncertain choice alternatives and therefore the choices she makes. Sender chooses an IS to move Receiver's (expected) choice closest to his own preferred choice. In the classical context Receiver updates her beliefs using Bayes rule and therefore the power of Sender is constrained by Bayesian plausibility: the fact the expected posteriors must equal the priors.

The quantum persuasion approach has been developed in the same vein as Bayesian persuasion. A central motivation is that persuasion seems much more influential than what comes out of the Bayesian approach. So instead of assuming that agents are classical Bayesian, it has been proposed that they

are quantum-like. A first line of justification is that people do not make decision based on reality but based on a *representation of that reality*, that is based on a mental object. The assumption is that this mental object does not evolve in response to new information in accordance to Bayes rule but follows instead Lüder’s rule.<sup>5</sup> This modelling approach is at the heart of Quantum Cognition (QC). It captures well documented cognitive limitations as argued in e.g., [3] and [8]. Namely the fact that people face difficulties in combining different types of information *into a stable picture*. Instead, the picture (mental object) that emerges often depends on the order in which information is processed.

The present paper aims at contributing to testing a central prediction of quantum persuasion namely that Sender can use distracting measurements to influence Receiver ([10]). A distracting measurement corresponds to a measurement that generates information that is incompatible (or Bohr complementary) with the information used by Receiver to evaluate the choice alternatives. The objective is to switch the focus of Receiver’s mind (distract her) which changes her cognitive state i.e., her beliefs although no information relevant to her true concern is ever provided.

## 2.1 The model

We next present the quantum persuasion model in the two dimensional case that is tested in the experiment. For a general and detailed exposition of the formal framework see [10].

**The representation of a choice alternative: a cognitive state** We have two animal protection projects Tiger Forever (TF) and Elephant Crisis Fund (ECF). The determination of the initial cognitive state is fully conditioned by the information provided at the beginning: the description of the projects.

A cognitive state is a mental object that gives a probability for every possible state of the world. It is very close to the notion of belief. However, in contrast to beliefs, a cognitive state is not an (imperfect) mirror image of the objective world. It is a mental object, a representation of the objective world that evolves in a way that reflects the cognitive constraints that we focus on: namely that an individual cannot consider all perspectives simultaneously. There exist perspectives that are not compatible in her mind. They are incompatible or Bohr complementary. As a consequence, the picture that arises depends on the order in which different pieces of information are processed. This approach has gained both theoretical and empirical support under the last 20 years (see e.g., [8], [3]).

We use Dirac’s notation that has advantages in the two dimensional case where it allows easily connecting with the geometrical representation. The initial cognitive states that captures the representation the projects are denoted  $|T\rangle$  and  $|E\rangle$  with  $T$  for Tiger Forever and  $E$  for Elephants Crisis Fund. .

Each (represented) project is characterized by two properties which are incompatible. We call them Urgency (of the cause) and Honesty (of the NGO managing the project). The urgency property (or perspective) is represented by a two dimensional basis ( $|U\rangle, |U^\perp\rangle$ ) (corresponding to Urgent respectively

---

<sup>5</sup>Updating following Lüder’s rule has been shown necessary to dynamic consistency in such a context (see Danilov et al. 2018).

not-Urgent) and the Honesty perspective is represented by an alternative basis of the same state space ( $|H\rangle, |H^\perp\rangle$ ) corresponding to Honest respectively not-Honest NGO.

**Preferences** Individual preferences are expressed by

1. A utility value attributed by the individual to the projects in each possible pure (subjective certainty) states e.g.,  $|U\rangle$  or  $|U^\perp\rangle$ . The utility of an uncertain state is calculated as a linear combination of those values.

2. A preferred perspective which we understand as the decision relevant (determinant) characteristics of a project for the individual (e.g., the Urgency of the cause). Formally, the preferred perspective is defined as the one in which a project utility can be represented as a diagonal matrix. The utility value associated with all non-commuting perspectives (characteristics) can be derived from those elements. We assume that the preferred perspective is common to both projects.<sup>6</sup>

This is illustrated in figure 1 in a two dimensional diagram. The two projects are represented each by two distinct cognitive states vector  $|T\rangle$  and  $|E\rangle$ :

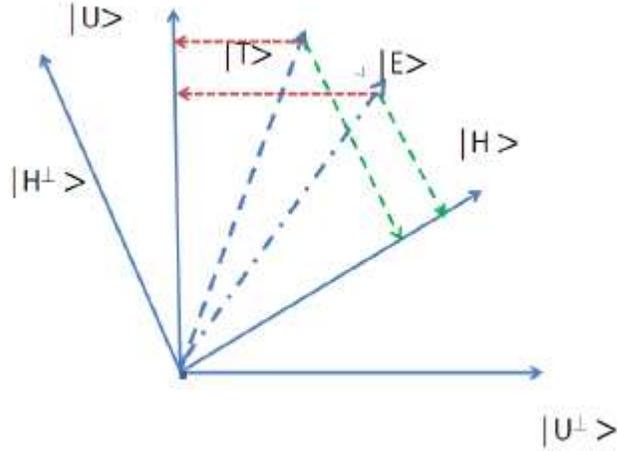


Figure 1

The figure reads as follows. For an individual endowed with preferences that define Urgency (U) as her preferred perspective, the expected utility value of contributing 50 euros to the Elephant Crisis Fund project is denoted  $EU(EFC; E, U)$ . It depends on two things: 1. her beliefs (about the cause's urgency) encapsulated in the cognitive state  $|E\rangle$  and 2. her valuation of contributing to an urgent respectively non-urgent EFC project. We denote these values  $x_U \in \mathbb{R}$  respectively  $x_{U^\perp} \in \mathbb{R}$ . As shown in [8] the utility value of project EFC can be expressed as a Hermitian operator  $E_U = \begin{pmatrix} x_U & 0 \\ 0 & x_{U^\perp} \end{pmatrix}$ . We write:

$$EU(EFC; E, U) = TR(E_U, E) = x_U \cdot |\langle E | U \rangle|^2 + x_{U^\perp} \cdot |\langle E | U^\perp \rangle|^2 \quad (1)$$

<sup>6</sup>An interpretation is that it reflects mega preferences here in the field of projects related to endangered species.

So for instance with  $x_U = 1$  and  $x_{U^\perp} = 0$ , the expected utility value of contributing to ECF when the individual has U-preferences is equal to her subjective probability that the elephant cause is urgent. That probability is, according to Born's rule<sup>7</sup>, the square of the correlation coefficient also called amplitude of probability  $\langle E | U \rangle$ . Graphically, the probability amplitudes are read off in the diagram as the orthogonal projection (green and red dotted lines) of the corresponding cognitive state on the basis vectors.

Similarly, we have for an individual with H-preferences:

$$EU(ECF; E, H) = TR(E_H, E) = x_H \cdot |\langle E | H \rangle|^2 + x_{H^\perp} \cdot |\langle E | H^\perp \rangle|^2 \quad (2)$$

where  $E_H$  the utility matrix is defined in the (preferred)  $(H, H^\perp)$  basis:  $E_H = \begin{pmatrix} x_H & 0 \\ 0 & x_{H^\perp} \end{pmatrix}$ . The corresponding expected utility values of the TF project for individual having U- and H-preferences respectively are:

$$\begin{aligned} EU(TF; T, U) &= TR(T_U, T) = y_U \cdot |\langle T | U \rangle|^2 + y_{U^\perp} \cdot |\langle T | U^\perp \rangle|^2 \\ EU(TF; T, H) &= TR(T_H, T) = y_T \cdot |\langle T | H \rangle|^2 + y_{T^\perp} \cdot |\langle T | H^\perp \rangle|^2 \end{aligned}$$

where  $T_U = \begin{pmatrix} y_U & 0 \\ 0 & y_{U^\perp} \end{pmatrix}$  (as defined in the  $(U, U^\perp)$  basis) and  $T_H = \begin{pmatrix} y_H & 0 \\ 0 & y_{H^\perp} \end{pmatrix}$  (as defined in the  $(H, H^\perp)$  basis) are the operator representing the utility value of choosing TF for a U-individual respectively a H-individual.

**Choice** The individual makes her choice by comparing the expected utility of each project and selects the one that yields the highest expected utility. For the sake of illustration take  $x_U = y_U = 1$  and  $x_{U^\perp} = y_{U^\perp} = 0$ , reading directly from the figure we see that

$$EU(TF; T, U) = |\langle T | U \rangle|^2 > EU(E; U) = |\langle E | U \rangle|^2$$

and

$$EU(ECF; T, H) = |\langle E | T \rangle|^2 > EU(T; H) = |\langle T | H \rangle|^2$$

Which means that in our example a U-individual prefers to contribute to the TF project while a H-individual prefers to contribute to the ECF project.

### 2.1.1 Persuasion

Persuasion is about modifying the cognitive state. This is achieved by means of an informational structure (IS) that is an operation on the cognitive state that triggers the resolution of uncertainty with respect to some aspect. It is important to keep in mind that we are dealing with mental objects (represented world). So in our context an IS can be a question that the individual puts to herself. The outcome is generally some level of conviction (subjective certainty).

<sup>7</sup>Born's rule is the formula for calculating probability in a quantum setting.

In quantum persuasion, an IS can be decomposed into two parts: a measurement device (MD) and an information channel (IC). An IC translates outcomes into signals. In the present context, we confine ourselves to trivial IC, where the signals are the outcomes of the MD. A MD is defined by a set of possible outcomes  $I$ , a collection of probabilities  $p_i$  to reach these outcomes where  $p_i$  depends on the cognitive state,  $p_i = Tr(P_i B P_i)$  where  $B$  is the initial cognitive state (prior). Upon obtaining outcome  $i$ , the cognitive state transits (is updated) into  $B_i = \frac{P_i B P_i}{p_i}$  according to Lüders' rule.<sup>8</sup> In this paper we focus on so called "projective measurements" that is MDs whose outcomes transit the prior into a pure cognitive state, i.e., a state of full conviction (subjective certainty with respect to some aspect).

As in the standard persuasion problem, Sender chooses the MD. In our experiment Sender chooses a question to Receiver. For instance "how urgent do you think it is to protect tigers respectively elephants from the threat of extinction". Such a MD is similar to a procedure that elicits (we prefer to say actualizes<sup>9</sup>) the individual's beliefs about the severity of the threat.<sup>10</sup> A crucial point to emphasize here is that eliciting beliefs does not provide any *informational* justification for modifying those beliefs (cognitive state). Yet, as we next shall see, as Sender asks such a question Receiver's cognitive state changes.

**The impact of measurements** In the development below we focus exclusively on introspective measurements in accordance with the experimental set-up.

**Compatible measurement** The performance of a measurement of the (represented) projects *in the individual's preference perspective* corresponds to actualizing decision-relevant beliefs. In contrast with the classical case (see below) this induces a 'collapse' of the belief states:  $|T\rangle \rightarrow |U_T\rangle$  or  $|T\rangle \rightarrow |U_T^\perp\rangle$  and similarly for  $|E\rangle \rightarrow |U_E\rangle$  or  $|E\rangle \rightarrow |U_E^\perp\rangle$  where the subscript informs about the project and is deleted when it is clear. The 4 possible combinations of cognitive states occur with the corresponding probabilities :  $\begin{pmatrix} |\langle T | U \rangle|^2 |\langle E | U \rangle|^2 & |\langle T | U^\perp \rangle|^2 |\langle E | U \rangle|^2 \\ |\langle T | U \rangle|^2 |\langle E | U^\perp \rangle|^2 & |\langle T | U^\perp \rangle|^2 |\langle E | U^\perp \rangle|^2 \end{pmatrix}$ .

Compared with the case without measurement where our U-individual is selecting TF with probability 1 (see above), we now see that with probability  $|\langle T | U^\perp \rangle|^2 |\langle E | U \rangle|^2$  he selects EFC since  $EU(TF; T, U)_{T=U_T^\perp} = Tr(T_U U_T^\perp) = 0 < EU(ECF; E, U)_{E=U_E} = Tr(E_U U_T) = 1$ . Similarly our H-individual will select TF with positive probability. The change in the cognitive state induces a change in revealed preferences i.e., we already have some persuasion. The impact of mere belief actualization underlines a distinction between the quantum and the classical framework. In the quantum world measurements always changes the state of the measured system (here the beliefs). This is an expression of the fundamental distinction with the classical world where it is assumed that reality preexists any measurement that merely reveals

<sup>8</sup>A behavioral justification for this rule is provided in [10].

<sup>9</sup>The distinction between eliciting and actualizing is that in the latter case we do not assume that the priors is a convex combination. Instead the individual is in a state of superposition which is the expression of indeterminacy. A superposition is a linear combination such that *the square* of the coefficients sum up to 1.

<sup>10</sup>When dealing with projective measurement as we do here the procedure consists in asking the respondents whether they believes the cause is urgent YES/NO. The statistical distribution of answers reflects the beliefs.

it. In the quantum world reality is contextual which means that measurements contribute in determining the state i.e., they do not only reveals a preexisting state, they create that state. Another distinction is the difference in the impact of compatible versus incompatible measurements as we show next.

**Incompatible measurement: distraction** We now turn to distraction i.e., the elicitation of beliefs with respect to features not (directly) relevant to decision-making. Below we depict the case when addressing a U-preference individual. Distraction corresponds to putting a question about NGO's Honesty, an incompatible perspective. Such a question triggers collapse of the cognitive states T and E in the  $(H, H^\perp)$  basis i.e., the state  $T$  ( $E$ ) transits into  $T'$  ( $E'$ ) equal to either  $H$  or  $H^\perp$  and it does so with probability  $|\langle T | H \rangle|^2$  ( $|\langle E | H \rangle|^2$ ) and  $|\langle T | H^\perp \rangle|^2$  ( $|\langle E | H^\perp \rangle|^2$ ). We illustrate this in figure 2 with the green lines for E the cognitive state representing ECF.

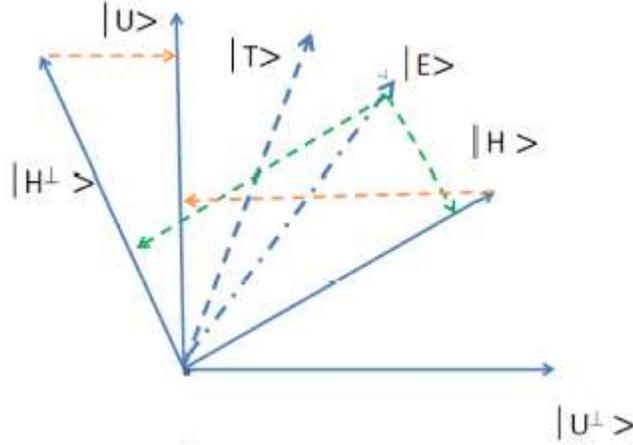


Figure 2: Distraction

The ex-ante expected utility of ECF for a U-individual with cognitive state E is obtained by applying a sequence of two non-commuting operations. First distraction, the state is projected onto the  $(H, H^\perp)$  basis. And then the resulting state ( $H$  or  $H^\perp$ ) is projected onto the preferred basis  $(U, U^\perp)$  in order to evaluate the project

$$EU(ECF; E, U) = \left[ |\langle E | H \rangle|^2 |\langle H | U \rangle|^2 + |\langle E | H^\perp \rangle|^2 |\langle H^\perp | U \rangle|^2 \right] x_U + \left[ |\langle E | H \rangle|^2 |\langle H | U^\perp \rangle|^2 + |\langle E | H^\perp \rangle|^2 |\langle H^\perp | U^\perp \rangle|^2 \right] x_{U^\perp}.$$

Consider the following numerical example for a H-individual where we simplify the matter by assuming  $T = E = D$ , that is the two projects are represented by the same cognitive state meaning that they are subjectively perceived as equally urgent and honest. And let this state in the H-perspective be

$$D = \begin{pmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{pmatrix}$$

Let  $x_H = 8, x_{H^\perp} = 7, y_H = 10, y_{H^\perp} = 4$ . Using the formula in (2) we obtain:

$$EU(ECF; D, H) = 4/5 \cdot 8 + 1/5 \cdot 7 = 39/5$$

And similarly

$$EU(TF; D, H) = 4/5 \cdot 10 + 1/5 \cdot 4 = 44/5$$

Which means that our H-individual chooses to donate to TF. Let us now consider a distraction toward the Urgency perspective that we model for simplicity as a  $45^\circ$  rotation of the H-basis:  $U = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$

$$\text{and } U^\perp = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}.$$

With a  $45^\circ$  rotation the computation simplifies greatly because whether distraction takes the states to  $U$  or  $U^\perp$ , the probability for  $H$  respectively  $H^\perp$  is the same and so is the expected utility

$$\begin{aligned} EU_U(ECF; D, H) &= \left[ |\langle E | U \rangle|^2 + |\langle E | U^\perp \rangle|^2 \right] 1/2 x_H + \left[ |\langle E | U \rangle|^2 + |\langle E | U^\perp \rangle|^2 \right] 1/2 x_{H^\perp} \\ &= 1/2 x_H + 1/2 x_{H^\perp} = 4 + 3.5 = 7.5 \end{aligned}$$

and similarly

$$\begin{aligned} EU_U(TF; D, T) &= \left[ |\langle T | U \rangle|^2 + |\langle T | U^\perp \rangle|^2 \right] 1/2 y_H + \left[ |\langle T | U \rangle|^2 + |\langle T | U^\perp \rangle|^2 \right] 1/2 y_{H^\perp} \\ &= 1/2 y_H + 1/2 y_{H^\perp} = 5 + 2 = 7. \end{aligned}$$

So after the distraction our H-individual chooses to donate to the ECF project. So we note that the impact of distraction can be a total reversal of the choice. This is in contrast with the impact of decision-relevant belief actualization (compatible measurement) which only triggers some partial reversal. For a general theoretical argument on the superior power of distractive IS as compared with compatible IS see [10].

Before moving to the experiment let us briefly remind of the classical subjective uncertainty approach in our example.

### 2.1.2 The classical uncertainty approach

First, we remind that under the classical uncertainty approach Receiver can process any kind of information. All information is compatible and therefore Bayesian updating applies. The individual simultaneously considers Urgency and Honesty and combines them to obtain her expected utility value. Assuming a separable and additive utility function we have

$$\begin{aligned} EU(T) &= \alpha (p_U^0(T) x_U + (1 - p_U^0(T)) x_{U^\perp}) + (1 - \alpha) (p_T^0(A) x_H + (1 - p_T^0(T)) x_{H^\perp}) \\ EU(E) &= \alpha (p_U^0(E) y_U + (1 - p_U^0(E)) y_{U^\perp}) + (1 - \alpha) (p_T^0(E) y_T + (1 - p_T^0(E)) y_{T^\perp}) \end{aligned}$$

where  $p_U^0(T)$  is the subjective probability in cognitive state  $T$  that the TF project is Urgent and similarly for the other probabilities. An individual who announces that Honesty is determinant is an individual with  $\alpha < 1/2$  and similarly for U-individuals ( $\alpha \geq 1/2$ ).

Recall that the measurements that we consider are exclusively introspective i.e., no information appealing to the outside world is called upon. In other words, the questions correspond to eliciting Receiver’s beliefs.

When asked ”do you believe the tiger cause is urgent Yes or No”. With probability  $p_U^0(T)$  Receiver answers YES and with probability  $(1 - p_U^0(T))$  she answers NO. But her beliefs do not change, they remain mixed. Since beliefs are unchanged so is the expected utility for the two projects. As a consequence Receiver has no reason to change her choice after having been exposed to this information structure i.e., after having answered the question. To put it differently, an introspective measurement has no persuasion power whatsoever in the classical context.

This result contrasts starkly with the quantum model where an introspective measurement with respect to both compatible and incompatible perspectives has impact on decision-making. And most importantly incompatible introspective measurements have the strongest potential to affect decision-making. It is precisely this prediction that we aim at testing with the next following experiment.

### 3 Experimental design

Our main experiment uses the property of Bohr complementarity of mental perspectives, i.e., their incompatibility in the mind of Receiver. More precisely it relies on the assumption that two perspectives relevant for evaluating projects aimed at saving an endangered species are incompatible. The two perspectives that we consider are ”the urgency of the cause” and ”the honesty<sup>11</sup> of the organization that manages the funds”. As a first step we provided experimental support for this hypothesis. We know that when two properties are incompatible measuring them in different orders yields different outcomes. Therefore, we started with an experiment to check whether order matters for the response profile obtained. Note that even in Physics there is no theoretical argument for establishing whether two properties are compatible or not. This must be done empirically. Once we experimentally provided support for the hypothesis of incompatibility of the two perspectives of interest, we moved to the main experiment concerned with the impact of distraction. We next describe the procedures and the results in details.

#### 3.1 Testing for perspective incompatibility

We recruited 295 respondents through Amazon’s Mechanical Turk for which data quality has been confirmed by different studies (e.g. Bartneck et al. 2015). The respondents completed the short survey below on the website Typeform. They were paid \$0.1 and spent on average 0:17 minute to complete the

<sup>11</sup>The terms ”honesty” and ”trustworthiness”, or ”trust”, are used interchangeably consistently with the definition given to honesty - see below.

survey.

Participants were first presented a screen with a short description of the situation of refugees in Myanmar including a mention of the main humanitarian NGO present in the field:

”About a million refugees (a majority of women and children ) escaped persecution in Myanmar. Most of them fled to Bangladesh. The Bengali Red Crescent is the primary humanitarian organization that is providing help to the Rohingyas. They are in immediate need of drinkable water, food, shelter and first medical aid.”

They were then asked to evaluate the urgency of the cause and the honesty to the NGO on a scale from 1 (“Not urgent” or “Do not trust”) to 5 (“Extremely urgent” or “Fully trust”). The order of presentation of the two questions was randomized so that half of participants responded to the urgency question before trust (U-T), and the other half conversely (T-U).

The data were processed, cleaned and analyzed with Stata. Responses were clustered into two groups: low level of urgency (resp. trust) (responses $\leq$ 3) and high level of urgency (resp. trust) (responses $>$ 3). Probit regression models were conducted to analyze the effect of the order of the questions on the responses.

As shown in table 1, the results show that the order of the questions significantly impacts the responses given to both urgency (p-value=.050) and trust (p-value=.026).

Table 1: Regression matrix for Order effects

	(1) TRUST	(2) URGENCY
main ORDER	-0.330* (0.148)	-0.323* (0.165)
_cons	0.147 (0.104)	0.918*** (0.122)

Standard errors in parentheses

\*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$

*Notes.* The variable "TRUST" and "URGENCY" take value 0 and 1 given the response to the question; the variable "ORDER" takes value 0 for UT and 1 for TU

The results are consistent with the hypothesis that the two perspectives are incompatible in the mind of people. As well-known there exist other theories for order effects. We observe that the value of the responses to the Trust question tend to be lower when that question comes first (i.e. T-U), whereas the responses to the Urgency question tends to have a higher value when that question comes first (i.e. U-T). Therefore, we can reject both the hypothesis of a recency bias and that of a primacy bias. This strengthens our quantum interpretation. We next proceed to the main experiment using those two perspectives.

### 3.2 Quantum persuasion

More than 900 participants completed the survey on the website Typeform, they were recruited through Amazon’s Mechanical Turk. They were paid either \$1 to 0.75\$ depending on the condition.

The participants were divided into three groups. Two treatment groups and a control group as explained below. All three groups were presented a screen with an introductory message, informing them that the questionnaire is part of a research project on quantum cognition and that they will contribute in deciding which one of two NGOs projects will receive a 50€ donation. The decision will be made by randomly selecting a respondent and implementing his or her decision. Presumably, this created an incentive to respond truthfully. The respondents were next asked to click on a button that randomly assigned them to a specific condition. In all conditions, participants were shown a short text about the situation of elephants respectively tigers and of ongoing actions of two NGOs working for their protection the Elephant Crisis Fund (ECF) and Tiger Forever (TF). The order of presentation of the text was reversed for half of the subjects. This aimed at isolating order effects not relevant to our main point. The screen displayed the following two texts:

“Elephant crisis fund : A virulent wave of poaching is on-going with an elephant killed for its tusks every 15 minutes. The current population is estimated to around 700 000 elephants in the wild. Driving the killing is international ivory trade that thrives on poverty, corruption, and greed. But there is hope. The Elephant Crisis Fund closely linked to World Wildlife Fund (WWF) exists to encourage collaboration, and deliver rapid impact on the ground to stop the killing, the trafficking, and the demand for ivory.”

“Tiger Forever : Tigers are illegally killed for their pelts and body parts used in traditional Asian medicines. They are also seen as threats to human communities. They suffer from large scale habitat loss due to human population growth and expansion. Tiger Forever was founded 2006 with the goal of reversing global tiger decline. It is active in 17 sites with Non-Governmental Organizations (NGOs) and government partners. The sites host about 2260 tigers or 70% of the total world’s tiger population”

It is worth mentioning that the descriptions were formulated so as to slightly suggest that the elephants’ NGO (EFC) could be perceived as more trustworthy (because of its link with of WWF, a well-known NGO). In contrast, the text about tigers suggested a higher level of urgency (the absolute number of remaining tigers is significantly lower than the number of remaining elephants). Thereafter, all respondents were confronted with a choice:

“When considering donating money in support of a specific project to protect endangered species, different aspects may be relevant to your choice. Let us know what counts most to you :

-The urgency of the cause: among the many important issues in today’s world, does the cause you consider belong to those that deserve urgent action? or

-The honesty of the organization to which you donate: do you trust the organization managing the project to be reliable; i.e. do you trust the money will be used as advertised rather than diverted.”

The objective was to elicit an element of their preferences (the preferred perspective, see section 2.1).

The rest of the questionnaire depended on which one of the three groups the participants belonged to.

In the control condition (baseline), they were next asked whether or not they wanted to read the first descriptions again or if they wanted to make their final decision i.e., to make their choice between supporting the Elephant Crisis Fund or Tiger Forever both represented by an image of an adult elephant respectively adult tiger (presented in random order on the same screen).

In the first treatment condition referred to as "QComp", the respondents were redirected to a screen with general information *compatible* with the aspect they indicated as determinant to their choice when making a donation. Importantly, the information did not directly or indirectly favor or disfavor any of the two projects. The information was aimed at triggering a measurement as they were invited to determine themselves with respect to which of the species was most urgent to save respectively which NGO was most trustworthy. We below return to the role and expected impact of the general information screens. Those who responded honesty saw a screen with the following text:

"Did you know that most Elephant and Tiger projects are run by Non-Governmental Organizations (NGOs)? But NGOs are not always honest! NGOs operating in countries with endemic corruption face particular risks. NGOs are created by enthusiastic benevolent citizens who often lack proper competence to manage both internal and external risks. Numerous scandals have shown how even long standing NGOs had been captured by less scrupulous people to serve their own interest.

So a reasonable concern is whether Tiger Forever respectively Elephant Crisis Fund deserves our trust."

Those who responded urgency saw:

"Did you know that global wildlife populations have declined 58% since 1970, primarily due to habitat destruction, over-hunting and pollution. It is urgent to reverse the decline ! "For the first time since the demise of the dinosaurs 65 million years ago, we face a global mass extinction of wildlife. We ignore the decline of other species at our peril – for they are the barometer that reveals our impact on the world that sustains us." — Mike Barrett, director of science and policy at WWF's UK branch. A reasonable concern is how urgent protecting tigers or elephants actually is."

Thereafter the respondents were offered the opportunity to read again the descriptions before making their image choice between ECF and TF.

In the main treatment condition referred to as "QIncomp" (distraction), participants were redirected to a screen with general information on the aspect they did *not* select as determinant to their choice; this is what we call a distraction. So those who selected honesty (resp. urgency) saw the screen on global wildlife decline (resp. NGO's scandals). Thereafter, the respondents were offered the opportunity to read again the initial project description before making their image choice.

Finally, information about their age, gender, education and habits of donation to NGOs was collected before the thank-you message and the end of the experiment.

#### *About the general information screens*

First, we note that the theoretical model does not account for anything like the general information

screen. In the model the only concern is with the implicit questions that follow the general information. Indeed general information plays no role for persuasion since it conveys no new data on the relative urgency or honesty of the specific projects. Therefore it should not affect the choice between the two projects. Before elaborating on that point let us address the question as to what is the role of the screens.

The justification for the screens is to be found in the quantum approach to cognition. Quantum cognition recognizes that people consider project from different perspectives some of which may be incompatible but not mutually exclusive. They are Bohr complementary which implies that the questions related to those perspectives do not commute: order matters. This in turn is an expression of the fact that the cognitive state is modified by responding to a question. Our intuition is that there can be some inertia. Consider a person who declared that Honesty is her priority which we interpret as her being in the Honesty perspective. If you abruptly ask whether the elephant cause is urgent, she might not make the effort to switch perspective in order to respond faithfully. In contrast if you softly accompany her into the switch with an engaging short text, she will find herself capable of responding truthfully without particular effort.

Hence the point with those screens is to accompany the change in perspective. Clearly that is only justified in the distraction treatment but for the sake of symmetry we have a similar screen in the treatment where no change of perspective is required .

But is it justified to expect no impact of general information on beliefs? The main argument is that it contains nothing new about the two projects at stake and therefore no updating is warranted. It could be argued that compatible general information about animal extinction provides some indirect information in favor of Urgency for both projects. In the classical version of the model, the weights given to the two dimensions are not the same for both projects. We could therefore expect some differential impact on the valuation of the two projects. We assume that if there is any effect on the final choice, it should be of non significant magnitude. In contrast the quantum model does not predict any impact.

One may wonder why we do not, in the experiment, simply ask people for their beliefs e.g., do you trust that NGO? The reason is that we wanted to avoid that the response would influence the respondent beyond the impact under investigation. Additional impact can be expected because of perceived dissonance. Assume you care for the urgency of the cause and since there are only 2700 tigers left, you are most likely to choose TF. If you are asked whether you believe that TF is Honest and decide that you don't trust them much, then it becomes psychologically difficult to select TF. On the other hand because we do not ask straightforwardly but use a general text to induce the measurement, people are expected to be less likely to perceive dissonance and choose more spontaneously. These precautions are among the difficult decisions we have to make in quantum cognition when trying to exhibit quantum effect in behavior. Individuals in contrast with particles are thinking systems endowed with among other things a drive toward consistency which can interfere with the intrinsic indeterminacy of preferences (see Discussion in section 5).

### 3.3 Theoretical predictions

Before getting into the results and their interpretation, let us remind of the main theoretical predictions:

- The Bayesian model predicts no effect in both treatment groups.
- The quantum model predicts some milder impact of the question in the first treatment compared with the control group's choice profile. In contrast, the distraction treatment group should exhibit a significantly different allocation of responses compared with the control group. It should be emphasized that since we lack information about the correlation coefficients between the two perspectives and the utility values, we do not have precise quantitative predictions.<sup>12</sup>

## 4 Results

Data were processed, cleaned and analyzed with Stata. 471 participants were removed from the data due to a technical misstep<sup>13</sup> and in order to equally balance the number of participants in each condition. The number of observations that were retained was 900. A number of Probit regressions were conducted to analyze the impact of the variables of interest.

### 4.1 Descriptive Statistics

60,6% of all respondents were male, the average age was 35,4 years and the average education level was undergraduate. Table 2, informs that 72,7% of the participants declared that the Honesty of the NGO rather than the Urgency of the cause is what counts most to their choice. Across the three conditions, 54,6% chose to support with their donation the Elephants Crisis Fund (ECF) and 45,4 the Tiger Forever project (TF). Looking into the different treatment groups, we find that 59% of the respondents in the control condition chose EFC and 56% in the compatible information treatment group. In contrast, in the distraction treatment group only 49% chose ECF. Conditional on revealed preferences, 52% of the respondents who value Urgency most chose to support TF, whereas 57% of those who value Honesty most chose to support EFC. 87,2% made their final decision without reading the description of the projects a second time. They spent on average 1:33 minute to complete the experiment.

---

<sup>12</sup>We know that the less correlated the two perspectives (in the example of section 2 they were fully uncorrelated) the larger the expected impact in terms of switching the choice for given utility values..

<sup>13</sup>Because of an error when performing the HIT on Mechanical Turk, some participants could have responded to the questionnaire twice and so were deleted.

Table 2: Descriptive statistics

Variable	Mean	Std. Dev.
FinalChoice	0.546	0.498
FinalChoice-Control	0.593	0.492
FinalChoice-QComp	0.556	0.498
FinalChoice-QIncomp	0.487	0.501
ChoiceHU	0.727	0.446
DecisionRead	0.872	0.334
Gender	0.606	0.489
Age	35.368	10.522
Education	1.98	0.706
NGO	0.424	0.495

*Notes.* FinalChoice - final choice between Tigers (=0) and Elephants (=1); FinalChoice-Control (resp. QComp, QIncomp) - final choice in the control (resp. QComp, QIncomp) condition; ChoiceHU - choice between Urgency (=0) and Honesty (=1); DecisionRead - decision to read the descriptions again (=0) or not (=1); Gender - females (=0) and males (=1); Education - highest level of formal education between secondary school (=0), high school (=1), undergraduate (=2), graduate and over (=3); NGO - donation of either nothing (=0) or something (=1) in the last 3 years.

## Data Analysis

**General results** The first set of results displayed in Table 3 establishes that distraction - i.e. the question related to the non-determinant perspective - has a statistically significant impact on the final choice. This result stands when the regression includes all variables ( $p=0.011$ ), when it only accounts for compatible and incompatible questions ( $p=0.009$ ) and when in addition it accounts for population variables (age, gender, education, contribution to NGOs) ( $p=0.013$ ). More precisely, it appears that everything else being constant, the predicted probability of choosing ECF is 10.52% lower for an individual in the incompatible condition than for an individual in the control condition (i.e. marginal effects). There is no significant impact of the compatible question in any of those regressions.

Not surprisingly there is a significant impact on the final choice of the declared determinant - i.e. Honesty versus Urgency which captures an element of preferences ( $p=0.024$ ). This influence is robust to the absence of other variables. More precisely, the predicted probability of choosing ECF is 8.45% higher for an individual claiming that Honesty is more determinant than for someone who preferred Urgency.

The compatible question has no impact when preferences are controlled for while distraction has a significant impact ( $p=0.011$ ) when preferences are controlled for. None of the population variables have any impact on the final decision.

In further regressions, we also observed that the decision to reread the descriptions was significantly affected by the extra-information (i.e. being in one of the two treatment conditions) ( $p=0.001$ ), the age ( $p=0.002$ ) and the experience with NGOs ( $p=0.006$ ). More precisely, people that have been exposed to more information, are young and have experience with NGOs tend to reread more.

Interestingly, none of the variables significantly affected the revealed preferences.

Table 3: Regression matrix for Final Choice

	(1)	(2)	(3)	(4)	(5)
	FinalChoice	FinalChoice	FinalChoice	FinalChoice	FinalChoice
FinalChoice					
QComp	-0.0936 (0.103)	-0.100 (0.104)			-0.105 (0.105)
QIncomp	-0.270** (0.103)	-0.259* (0.104)			-0.265* (0.104)
Age		0.00198 (0.00410)		0.00261 (0.00408)	0.00171 (0.00413)
Gender		-0.0768 (0.0877)		-0.0820 (0.0872)	-0.0695 (0.0879)
Education		-0.00852 (0.0606)		-0.0212 (0.0605)	-0.0177 (0.0608)
NGO		0.000512 (0.0866)		0.00357 (0.0862)	-0.00885 (0.0872)
Order		0.00660 (0.0845)		0.00882 (0.0845)	0.00982 (0.0847)
ChoiceHU			0.214* (0.0939)	0.210* (0.0943)	0.213* (0.0945)
DecisionRead					-0.0428 (0.128)
_cons	0.236** (0.0731)	0.225 (0.223)	-0.0408 (0.0799)	-0.0444 (0.221)	0.138 (0.261)

Standard errors in parentheses

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ 

**Advanced results.** As shown by Table 4, it turns out that when looking closer at the impact distraction, it is not homogenous across subgroups. In particular we find that the order of the initial presentation of the projects matters. The impact of distraction is most pronounced for those who were presented the Tiger Forever project first and Elephant Crisis Fund last (TE subgroup see Fig.3). Distraction significantly affected the final choice in that group (corresponding to 50% of the respondents) ( $p=0.009$ ). In addition when combining the TE presentation order with the Honesty subgroup the coefficient gets closer significance at 5% level ( $p=0.003$ ). Similarly, for ET Urgency a switch can be seen (in Fig.3).<sup>14</sup> For the other two subgroups, distraction had no significant impact ( $p=0,915$  and  $p=0,577$ ).

In contrast revealed preferences (i.e., choice between Honesty and Urgency) are independent of the order of presentation - around 70% declare that honesty is determinant irrespective of the condition.

We also notice a clear distinction between the reaction to distraction depending on preferences alone. "Honesty respondents" react consistently with the theoretical predictions while Urgency people are less affected by distraction.

We return to these results in the interpretation below.

<sup>14</sup>We could not establish a significant p-value however. That group is much smaller around 14% of the whole sample. Further research is needed to better understand the pattern that merges - see below interpretation.

Table 4: Regression matrix for Final Choice

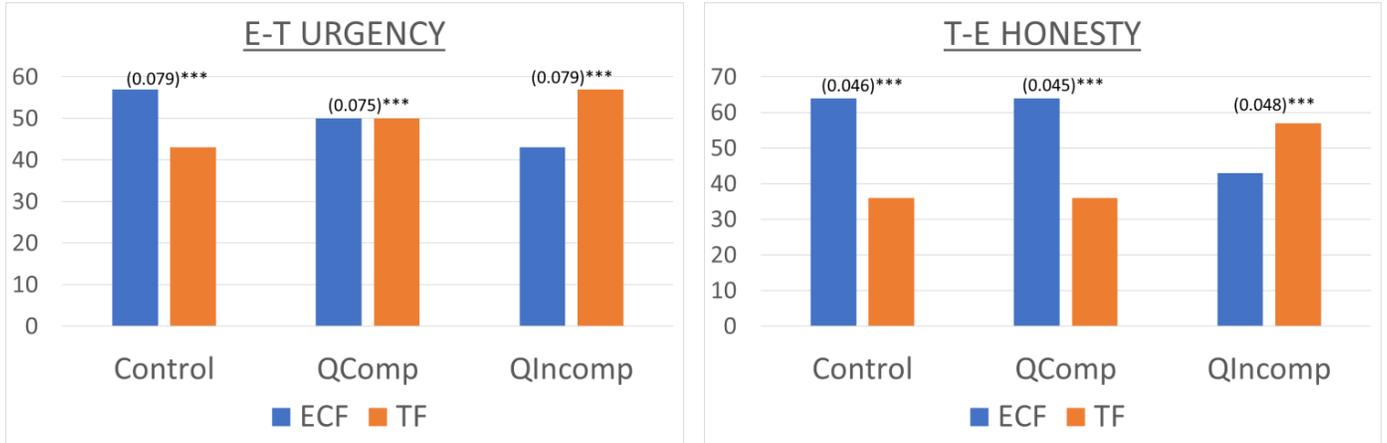
	General	Urgency	Honesty	T-E	E-T	TE-Urg	ET-Urg	TE-Hon	ET-Hon
FinalChoice									
QComp	-0.105 (0.105)	-0.148 (0.204)	-0.0849 (0.124)	-0.0576 (0.150)	-0.166 (0.147)	-0.0706 (0.313)	-0.225 (0.281)	-0.0292 (0.177)	-0.149 (0.175)
QIncomp	-0.265* (0.104)	-0.188 (0.201)	-0.311* (0.123)	-0.392** (0.150)	-0.164 (0.147)	-0.0311 (0.290)	-0.402 (0.288)	-0.535** (0.179)	-0.0962 (0.172)
ChoiceHU	0.213* (0.0945)			0.270* (0.136)	0.159 (0.133)				
DecisionRead	-0.0428 (0.128)	-0.0288 (0.265)	-0.0389 (0.148)	0.00932 (0.183)	-0.116 (0.184)	0.263 (0.415)	-0.220 (0.355)	-0.00222 (0.209)	-0.0967 (0.216)
Age	0.00171 (0.00413)	-0.00789 (0.00860)	0.00524 (0.00476)	-0.00169 (0.00617)	0.00388 (0.00564)	-0.0349* (0.0143)	0.0115 (0.0115)	0.00795 (0.00715)	0.00163 (0.00654)
Gender	-0.0695 (0.0879)	-0.238 (0.170)	0.00507 (0.104)	-0.147 (0.124)	0.0227 (0.126)	-0.449 (0.249)	-0.0617 (0.246)	-0.00972 (0.147)	0.0407 (0.149)
Education	-0.0177 (0.0608)	-0.0230 (0.118)	-0.0188 (0.0717)	0.0253 (0.0866)	-0.0792 (0.0865)	-0.0570 (0.175)	-0.0337 (0.169)	0.0431 (0.103)	-0.0875 (0.102)
NGO	-0.00885 (0.0872)	-0.116 (0.170)	0.0433 (0.103)	-0.220 (0.126)	0.181 (0.122)	-0.169 (0.258)	-0.0501 (0.238)	-0.195 (0.149)	0.259 (0.144)
Order	0.00982 (0.0847)	0.127 (0.163)	-0.0291 (0.0996)						
_cons	0.138 (0.261)	0.547 (0.495)	0.185 (0.299)	0.237 (0.377)	0.135 (0.364)	1.325 (0.780)	0.126 (0.670)	0.0927 (0.435)	0.298 (0.420)

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Notes. This table explains the different regressions on FinalChoice, participants' final decision between Tigers (=0) and Elephants (=1). Info (resp. InfoIncomp) is a dummy variable that takes value 1 when subjects are in the compatible (resp. incompatible) condition, and 0 when they are not (i.e. control condition). Age the age of participants; Gender their gender which was divided into females (=0) and males (=1); Education the highest level of formal education between secondary school (=0), high school (=1), undergraduate studies (=2), graduate and over (=3); NGO the fact of having donated nothing (=0) or something (=1) to an NGO in the past 3 years; Order the order of the projects descriptions as TF-ECF (=0) or ECF-TF (=1); ChoiceHU represents the choice between Urgency (=0) and Honesty (=1); DecisionRead the decision to read the descriptions again (=0) or to make the final decision (=1).

Figure 3: Descriptive histograms



Notes. Standard errors in parenthesis

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Note the reversal tendency between the Control and the QIncomp condition for E-T Urgency and for T-E Honesty

#### 4.1.1 Interpretation

The general results show with no ambiguity that the question triggered by incompatible information (distraction) had a significant impact on the final choice. It induced some extent of switch as compared to both the control group and the compatible information group.

The importance of preferences i.e., the answer to "what is determinant to your choice" for the final choice underlines that the initial texts were well-understood. The description of the Elephant project was designed to suggest more trust to the NGO managing the project and the description of the Tiger project to suggest a higher level of urgency.

As noted there is no significant impact of the compatible question. This is consistent with the Bayesian model. We argue that it is also consistent with the quantum model that predicts some impact *at the individual level*. We remind that an individual who chooses TF on the basis of mixed beliefs will choose ECF with some probability if she is forced to decide for herself whether the cause is urgent YES or NO, prior to decision (see sect. 2.1.1). The same is true for those who choose ECF on the basis of expected utility based on mixed beliefs. At the population level these effects counter balance each other so the overall impact is not significant. Moreover the fact that general compatible information had no impact also supports the view that it is not merely being exposed to an information screen that affects the choice. Instead it is only when information induces a change in perspective that something happens.

We found a significant variation between subgroups. First we could exhibit a clear distinction between the reaction to distraction depending on preferences alone. Honesty people reacted significantly to distraction while Urgency people were less affected by distraction. This can be explained by a systematic distinction in preferences between these two subgroups. It seems reasonable that Urgency people are

more passionate about the situation. In other word they have more extreme preferences in terms of the utility value they attribute to contributing to an urgent project. Such preferences tend to imply a smaller impact of distraction compared with people with more balanced preferences.

In addition we find that the order of the initial presentations has a significant impact. The order of presentation of the projects is an element of the state preparation procedure. The initial information prepares the state of belief. The data suggest that when there is concordance between preferences and the recency order that is between the preferred perspective and the one suggested by the last presentation, distraction has a more powerful impact (see Fig.3). One possible explanation is that "concordant respondents" are more sensitive to distraction because they were made aware of something they had not thought about. This contrast with people who already have been dealing with both perspectives. It also seems reasonable that people would be subject to a recency effect rather than a primacy effect because the first presentation is read without any thought about "what is most important for you" while the last presentation is immediately followed by that question. This line of interpretation goes outside of our stylized theoretical quantum model however.

The time for responding to the whole questionnaire was between 1 and 3 minutes which is rather short. We view this feature as a good thing because the quantum working of the mind is not rational reasoning: no new information of relevance for the choice was provided yet it did affect the choice. The respondent did not take time to reflect, they reacted spontaneously to the distraction. Recall that we do not elicit their preferences for the projects but only for what is determinant in this kind of situation. We found that those determinants are highly correlated with the final choice both in the control and compatible information groups. Yet, the distraction results suggest that respondents were not conscious of that. Therefore, they were not confronted with a (conscious) cognitive dissonance when the distraction changed their focus and eventually affected their decision.

We note that the respondents were given a chance to reassess their understanding of the project presumably in an attempt to be consistent . Only 15% used the opportunity re-read before making their choice.

Finally, an interesting finding is that the results are fully independent of population variables which supports the hypothesis that the quantum-like structure is a general regularity of the human mind.

## 5 Discussion and concluding remarks

We have made a number of simplifying assumptions in the theoretical model. Among them is that the two mental pictures (representing ECF and TF) are assumed not to be "entangled" that is measuring beliefs and obtaining say  $U$  for TF has no direct impact on the outcome of a next following measurement of the beliefs with respect to ECF: the systems are separable. Although this can be questioned, it is important to understand that it does not preclude "cross signalling" effects because of obvious proximity the two representations in the mind of the individual. Cross signalling effects are classical effects. For

instance in the preparation phase, thinking first about elephants may have an influence on the way tigers are perceived e.g., on the determination of the initial cognitive states. That may account the order effect discussed earlier. It is not among the model's predictions because it is not a *quantum* effect in a model that views the two mental pictures as separate systems.

Next, our focus is on introspective questions. This allows establishing a clear distinction between the classical model's predictions and the quantum one. A flipside is that we have little control over the actual resolution of uncertainty. Some respondents may not have bothered to decide for themselves as they were invited to by the information screens. It is not clear whether requesting an explicit answer would have generated more sincere effort. Our approach aimed at minimizing the possible impact of perceived cognitive dissonance. Such a dissonance may arise if a respondent decides that the Tiger cause is urgent while she earlier declared honesty is determinant to her choice. She may feel compelled to support ECF to remain consistent while her current preference favors TF.

In the experiment that we performed, the change of focus brought about by the distractive question affects revealed preferences for uncertain alternatives. But when it comes to information processing related to the outside classical world, we know that the correct updating rule is Bayes' rule which dictates no updating. There is a discrepancy between the properties of the true classical object and the properties of its representation, the mental object. When Receiver processes information about a classical object as if it was a quantum system, she is mistaken. But as amply evidenced by Kahneman's best selling book "Thinking Fast and Slow", information processing is not always disciplined by rational thinking when the brain processes information quickly. However, as argued in [7] the learning process adapted to a quantum-like world may be appropriate when you are interested in actions/decisions produced by other people. This is because they would also base their decisions on a quantum-like representation of the world. It is true that fast quantum-like information processing is inappropriate when dealing with simple decision involving classical objects. But it may be suitable in many situations involving human beings. Interestingly, we learned from Harari in his best selling book "Sapiens - a brief history of humankind", that a defining trait that secured Sapiens' dominance over other species was their capacity to communicate and cooperate with each other rather than to deal with (classical) Nature.

This paper has proposed a quantum approach to persuasion and contrasted it with the classical rational Bayesian approach. There exists other non-Bayesian approaches. Let us return to the one in "Thinking Fast and Slow" because of its relation to the literature on nudge. The basic idea is that the mind functions according to 2 modes called system 1 and system 2. System 1 operates automatically and quickly and is subject to a number of bias and cognitive limitations. System 2 is the rational thinking mode. The two-system approach does open the way for manipulation. The basic idea is that when the individual thinks fast she makes mistakes which can be exploited. We do not believe that the quantum approach is an alternative to most behavioral theories. Instead we believe that it can provide a rigorous foundations to a number of them as argued in e.g., [18]. However the interpretation tends to be very different. Quantum cognition proposes that all forms of thinking are contextual due to the intrinsic

indeterminacy of mental objects including beliefs and preferences. Conscious thinking may however interfere and constrain contextuality. The reason is that individuals have a resistance to changing their mind without “proper reason” due to drive toward consistency. This drive needs be related to rationality however but to an entrenched attachment to a stable identity or ego. The existence of a stable identity has been questioned by numerous experimental results (see e.g. self-perception theory and [18]). Those studies are consistent with a contextual and thus unstable identity ([18]). As in the two-system approach the extent of conscious thinking matters. This is because the drive toward maintaining a coherent ego is more effectual when the individual is conscious about her instability. As argued in [19] cognitive dissonance and its resolution is an expression of that drive in face of instability (arising from intrinsic indeterminacy). We close this short discussion by suggesting that the quantum-like nature of mental objects needs not reflect a cognitive failure but would be the expression of the intrinsic indeterminacy (contextuality) of *human reality*. The question of rationality in such a context deserves further investigation.

Finally, we recognize that quantum cognition experiments do not by far have the precision of physical experiments which prevents making and testing quantitative predictions. To a large part, this is because it is (today) impossible to fully characterize the state of a cognitive system which is incommensurably more complex than that of an atomic particle. Nevertheless, our experimental exercise shows that it may be useful to test some theoretical predictions in contrast with standard classical (Bayesian) ones.

## References

- [1] Akerlof G., and R. Schiller (2015) *Phishing for Phools - the economics of manipulation and deception*, Princeton University Press.
- [2] Baron, R. S., Baron, P. H., & Miller, N. (1973). The relation between distraction and persuasion. *Psychological Bulletin*, 80(4), 310.
- [3] Bruza P., and J. Busemeyer (2012) *Quantum Cognition and Decision-making*, Cambridge University Press.
- [4] Cialdini, R. B. (1984). *Influence: The Psychology of Persuasion* (ISBN 0-688-12816-5)
- [5] *Mécanique Quantique* (2000), Collection Enseignements des Sciences, Herman Editeurs des sciences et arts.
- [6] Chong D. and Druckman James (2007) "Framing Theory" *Annual Review of Political Sciences* 10, 103-126.
- [7] Danilov V. I., and A. Lambert-Mogiliansky (2010). Expected Utility under Non-classical Uncertainty. *Theory and Decision*, 68, 25-47.
- [8] Danilov V. I., A. Lambert-Mogiliansky, and V. Vergopoulos (2018) "Dynamic consistency of expected utility under non-classical (quantum) uncertainty." *Theory and Decision* June 2018, Vol 84/4, 645-670 <http://rdcu.be/JrES> .
- [9] Danilov V. I., and A. Lambert-Mogiliansky (2018A) "Preparing a (quantum) belief system", *Theoretical Computer Sciences* (in press), ArXiv <http://arxiv.org/abs/1708.08250>.
- [10] Danilov V. I., A. Lambert-Mogiliansky (2018B), Targeting in Persuasion Problems. forth-coming in the *Journal of Mathematical Economy*, pre-print <http://arxiv.org/abs/1709.02595>
- [11] Dubois F., and A. Lambert-Mogiliansky (2016). Our (represented) world and quantum-like object, in *Contextuality in Quantum Physics and Psychology*, ed. Dzafarof et al, World Scientific, Advanced Series in Mathematical Psychology Vol. 6, 367-387.
- [12] Festinger, L., & Maccoby, N. (1964). On resistance to persuasive communications. *The Journal of Abnormal and Social Psychology*, 68(4), 359.
- [13] Geanakoplos, John, David Pearce, and Ennio Stacchetti. "Psychological games and sequential rationality." *Games and economic Behavior* 1.1 (1989): 60-79.
- [14] Harari Y. N. " *Sapiens - A brief history of Humankind*" (2014) London Harvill Secker.
- [15] Haven E. and A. Khrenikov editors, (2017), *The Palgrave Handbook of Quantum Models in Social Sciences*, Palgrave MacMillan.

- [16] Kahneman D. (2011) *Thinking Fast and Slow*, Farrar, Strauss and Giroux, New York.
- [17] Kamenica E., and M. Gentzkow (2011). Bayesian Persuasion. *American Economic Review*, 101(6): 2590-2615.
- [18] Lambert-Mogiliansky A. (2012) "Quantum Type Indeterminacy in Dynamic Decision-Making: Self-Control through Identity Management" *Games* 3, 97-118.
- [19] Lambert-Mogiliansky A., S. Zamir and H. Zwirn (2009) "Type-Indeterminacy a Model of the KT(Kahnemann and Tversky)Man", *Journal of Mathematical Psychology* vol 53/5, 349-361.
- [20] Lipnowski, Elliot, and Laurent Mathevet. "Disclosure to a psychological audience." *American Economic Journal: Microeconomics* 10.4 (2018): 67-93.
- [21] Petty, R. E., Cacioppo, J. T. (2012). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer Science & Business Media.