Testing quantum-like models of judgment
for conjunction fallacy

Thomas Boyer-Kassem,* Sébastien Duchêne,† Eric Guerci†

March 13, 2015

1   Introduction

Conjunction fallacy is a bias that was first empirically documented by Kahneman and Tversky (1982, 1983) through a now renowned experiment in which subjects are presented with a description of someone called "Linda":

> "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations."

Then, subjects are shown a list of 8 possible outcomes describing her present employment and activities, and are asked to rank the propositions by representativeness or probability. In this experiment called "within-subject design", two "target" items were specifically tested:

(a) "Linda is a bank teller", proposed in the sixth position among the eight proposals,

(b) "Linda is feminist and a bank teller", proposed in the eighth position.

Empirical results show that people judge (b) more likely than (a), which is a typical fallacy for classical models, since in classical probabilities a conjunction cannot be more probable than one of its components ($P(F \cap B) \leqslant P(B)$).

Because of the surprising nature of the original experimental result, Kahneman and Tversky, but also economists or psychologists more generally, have tried numerous and various experiments to look at the persistence or not of this fallacy (Mellers et al. 2001). New protocols have been proposed, and their effects on experimental results have been observed. Specific methodologies have also been tested in an attempt to reduce the magnitude of the conjunction effect. For example, in the "between subjects" experiment, Kahneman and Tversky (1982) divide participants into two groups. In each group, only seven propositions were written — including just one of the two target items recalled above, (a) or (b). The experimental result was that the fallacy still occurred. Then, Kahneman and Tversky — thinking that this new design might "eliminate the error" — proposed an "increasingly desperate" experiment in which they only ask to select the more probable item between

---
*Archives H. Poincaré (UMR 7117: CNRS, Université de Lorraine), 91 avenue de la Libération, 54000 Nancy, France. E-mail: thomas.boyer@univ-lorraine.fr

†GREDEG (UMR 7321: CNRS, Université de Nice Sophia Antipolis), 250 rue Albert Einstein, 06560 Valbonne, France. E-mail: sebastien.duchene@gredeg.cnrs.fr, eric.guerci@gredeg.cnrs.fr

(a) and (b). The results were still in favor of a persistent conjunction fallacy. In another test, respondents were asked to rank four possible items about the next Wimbledon tournament and the results of Borg's match — or about US politics and Reagan's budget choices. In the same way, as re-described recently by Kahneman (2011) descriptions of dead people caused by floods were tested. Considering these two scenarios, people were asked to "evaluate their probability":

(a) " A massive flood somewhere in North America next year, in which more than 1,000 people drown"

(b) "An earthquake in California sometime next year, causing a flood in which more than 1,000 people drown"

Every time, students confused the most likely scenario with the most consistent, coherent or most "plausible" one. Indeed Kahneman used the word "plausible" in the sense of what can be expected, which seems to be able to be accepted, admitted, and considered as the truth. In the flood example recalled above, people considered the earthquake in California more probable — contrary to logic — because more details were given, making this story more coherent. As proposed by Tversky (1977) or Kahneman and Tversky (1982), "similarity or representativeness can be inscreased by specification of the target" which means that adding details to a target (e.g. a story of flood in California) can render it more likely. Then, in a completely different approach, to avoid terms of probabilities or events, authors used a regular six-sided die with green and red faces and presented combinations of possible outcomes. Even with this new presentation and exercise, students violate the conjunction rule and did not choose the most probable option — which is the most inclusive event.

Under some specific conditions, Kahneman and Tversky were able to reduce the conjunction fallacy, for example by changing percentages or probabilities into frequencies or numbers. In the same direction, in a comprehensive and complete approach, Hertwig and Gigerenzer (1999) offered a "frequency" description of Linda about 200 women and ask respondents :

– "How many of the 200 women are bank tellers?"

– "How many of the 200 women are bank tellers and are active in the feminist movement?"

The results show that in a within-subject's design with such frequencies, conjunction effect disappeared. From theses works, several discussions have taken place about the semantic and the design of Linda's experiments. For example, Gigerenzer (1991) considered that "frequency" representation would eliminate this bias whereas Kahneman and Tversky (1996) showed that in between-subjects experiments — even in a frequency representation —, responses would not respect the conjunction rule. As explained in Mellers et al. (2001), other aspects of Linda's experiment have been discussed, from the ambiguity of the term "and" (Hertwig et al. 2008), that could be interpreted as a "union" operator (Hertwig, 1997) to the exact structure of sentences and the implicit meaning of each word, that are used in the eight proposals. In the same way, some questions were given on the impact of the filler items' presence and contents (i.e the 6 other items that were neither "bank teller" nor "feminist and bank teller", Hertwig and Chase 1998). These many comments and discussions in the literature have enabled to highlight the questions of misunderstanding hypothesis (Moro 2009), of methodology, the weight of the design (Tentori et al. 2004, Wedell and Moro 2007) as well as the major role of semantic and syntax in this fallacy (Tentori and Crupi, 2008). However, the many experimental protocols and presentations

have shown that the conjunction fallacy was robust to several changes (Stolartz-fantino et al. 2003 or Erceg and Galic 2014) — but less in others (Hertwig and Gigerenzer 1999 or Charness et al. 2010) .

The conjunction fallacy has become so famous that behavioral economists have been interested to look at the consequences of this effect for understanding "real life economic behavior". Thus economists have tried to measure the robustness of this bias in an economic context. For example, Charness et al. (2010) tested subjects — in Linda's experiment — with monetary incentives, or allowing consultations and short discussions with others. In both these new protocols, conjunction fallacy is being greatly reduced. Erceg and Galic (2014) or Nilsson and Anderson (2010) studied conjunction fallacy in sport betting.

The topic has generated a large number of papers and debates about it, leading to new questions of rationality, human decision making and reasoning in many fields such as philosophy or economics (Hartmann and Mejs 2012, Tentori and Crupi 2012). A central aim has been to account for the conjunction fallacy, and many mathematical models have been proposed. To give a few recent examples, Von Sydow (2011) suggested a Bayesian logic model, while Franco (2009) or Pothos and Busemeyer (2013) developed quantum-like models.

These last models are called "quantum" because they rely on the mathematics of a major contemporary physical theory, quantum mechanics — note that only some tools of quantum mechanics are exploited, and the models are not justified by an application of quantum physics to the brain. Such quantum-like (QL) models have been flourishing in the last decade, in cognitive sciences and in decision theory, and they intend to account for human judgments that are considered as irrational and paradoxical for a classical framework. The mathematics from quantum mechanics provide some flexibility and help to overcome some of the constraints implied by conventional mathematics like Bayesian models that are usually employed by economists or psychologists to account for human behavior.

Given the growing importance of these QL models in various fields, they deserve to be studied and assessed. In this paper, we focus on these QL models that want to account for the conjunction fallacy, and we adopt a critical viewpoint. Before presenting these models, it is worth reviewing the various applications that have been made with QL models more generally. The field of QL models is vast and is at the crossroads of economics, psychology, physics, philosophy and mathematics.

In a first approach, QL models of judgment have been suggested to explain order effect. This human paradoxical judgement — in which the answers given to two questions depend on the order of presentation of these questions — has been highlighted by Schuman and Presser (1981), Tourangeau, Ribs and Rasinsky (2000), or Moore (2002). Basically, the idea of QL models here is to exploit the non-commutative features of the quantum mathematical tools. Pothos and Busemeyer (2013) or Busemeyer and Bruza (2012) present a QL model that accounts for several types of order effect. Along the sames lines, Wang and Busemyer (2013) and Wang et al. (2014) propose a quantum-like model to explain several polls on political or social issues that consist in asking two questions A and B in two different orders according to groups. Then, Atmanspacher and Römer (2012) suggest that Wang and Busemeyer's model could be extended to mixed states. Also, Conte et al. (2009) offer a quantum-like model that account for order effects in mental states "during visual perception and ambiguous figures".

3

Another famous paradox is the violation of the sure thing principle, one of Savage's axioms for expected utility (1954). This principle states that if someone prefers choosing action A to B under a specific state of the world and also prefers choosing A to B in the complementarity state, then he should choose A over B regardless of the state of the world. The paradoxes of Allais (1953) or Ellsberg (1961) and the disjunction effect highlighted by Shafir and Tversky (1992) show the violation of this axiom. To account for this "irrational behavior" for classical theories, some QL models have been proposed. Busemeyer, Wang and Townsend (2006) compare Markov models with a quantum-like dynamic model and reveal that interference effect can produce new predictions that could explain the disjunction effect. In the same way, Busemeyer and Wang (2007) construct a QL "information processing model" to account for the disjunction effect, and compare it with a markov Model. Khrennikov and Haven (2009) then use the wave function to understand Ellsberg paradox. Pothos and Busemeyer (2009), Busemeyer et al.(2009) or Trueblood and Busemeyer (2011) proposed a dynamic QLmodel — in the Hilbert space, with the Schrödinger's equation — that rely on a time evolution of the mental state to account for some experiments described by Kahneman and Tversky. In another approach, Aerts et al. (2011) introduce the notion of "quantum conceptual thought" to offer a QL model that illustrates the Hawaï problem.

An other empirical feature, such as asymmetry judgments in "similarity", has also been investigated with QL models. Tversky (1977), discussed later by Ashby and Perrin (1988) "provides empirical evidence for asymmetric similarities and argues that similarity should not be treated as a symmetric relation". According to Tversky, asymmetry in judgments of similarity means that "A is like B" is not equivalent to "B is like A" such that "we tend to select the more salient stimulus" for a referent, and the less one for "a subject". To illustrate this concept, Tversky highlights that we used to say "North Korea is like red China", but not "Red China is like North Korea". To account for this fallacy, Pothos and Busemeyer (2011) introduce a quantum probability similarity model that predict violations in similarity judgments.

QL models have been developed not only in decision theory but also in game theory (Wilkens and Lewenstein (1999)). Also, Lambert-Mogiliansky et al. (2009) propose to rethink agents' preferences as being indeterminate. In this way, an agent is described by a state that is a superposition of preferences. This enables to explain several phenomena of non-commutativity in patterns of behavior such as in decision theory or in game theory. More generally, Danilov and Lambert-Mogiliansky (2008) discuss the "cost" — in terms of behavioral assumptions — to go to the Hilbert space model of QL model and connect it with the notion of bounded rationality (Simon 1957).

Finally, QL models have been developed to account for the conjunction fallacy, as it has been indicated above. Arguing that quantum formalism can violate the joint probability rule of classical probability, Franco (2009) proposes a QL model to explain conjunction fallacy's experimental results, relying on the interference effect. Busemeyer *et al.* (2011), Pothos and Busemeyer (2013) and Trueblood *et al.* (2014) offer an account of the conjunction fallacy in Linda's original experiment, supposing that the cognitive process acts in some orderly and sequential way, such that subjects evaluate first "the more probable possible outcome" and then less probable outcomes.

Here, we focus on these QL models for the conjunction fallacy. The central concern of our paper is the empirical adequacy of these models. To what extent have they been empirically tested? The answer from the literature is that these models are able to account for existing data regarding the fallacy (e.g Franco 2009, Yukalov and Sornette 2011 or Busemeyer and Bruza 2012). We propose to go further, and to test *other* empirical

predictions of these models — predictions that are about other situations than the Linda experiment. For these new predictions, we follow the suggestion made by Boyer-Kassem, Duchêne and Guerci (2015), and consider to test the so-called "Grand Reciprocity" equations, which happen in an order-effect situation. We discuss here for which QL models for conjunction fallacy these equations hold. Are these Grand Reciprocity equations actually empirically verified? This is a question that the existing data cannot answer, as noted in 2009 by Franco:

> "There are no experimental data on order effects in conjunction fallacy experiments, when the judgments are performed in different orders. Such an experiment could be helpful to better understand the possible judgment strategies."

We intend to fill this gap here, by running an experiment that collects the needed data — the ones that Franco is talking about are exactly those needed to test the GR equations. Note that the methodology is novel: we are not testing the QL models against data produced by traditional conjunction fallacy experiments that the model were designed to explain, but we are testing them against other data, in a new experimental framework on which the models actually make some predictions. So, the experimental situation we shall consider is different from the usual Linda experiment.

The paper is structured as follows. In section 2, a theoretical and general QL model is introduced, which summarizes several existing QL models in the literature that account for the conjunction fallacy. Then we introduce the Grand Reciprocity equations that the model must satisfy (Section 3), and which are equalities of conditional probabilities. To test these GR equations, we design and run a new Linda-like experiment, that is introduce in section 4. The results (in section 5) show that several quantum-like models for the conjunction fallacy are not empirically adequate. In section 6, we discuss the consequences of our results, and in particular the QL models that are promising and could replace the failed ones. We discuss possible difficulties they might however encounter, and the other models that could be tested in the same way.

2 The conjunction fallacy and its quantum-like models

2.1 Quantum-like models

To explain the conjunction fallacy and to account for the empirical results of Linda's experiment, models that rely on parts of quantum mechanics have been defended, by Franco (2009), Busemeyer *et al.* (2011), Busemeyer and Bruza (2012), Pothos and Busemeyer (2013). For simplicity, we choose here to summarize them with a single model with our own notations, but the correspondence with the various models from the literature can easily be made by the reader. Also, for pedagogical purposes, we shall consider the conjunction fallacy through a Linda experiment, but the generalization to other instances of the conjunction fallacy are straightforward.

According to the quantum-like models, an agent evaluates the propositions in a Linda experiment by answering for herself successively two simple dichotomous yes-no questions: F = "Is Linda a feminist?" and B = "Is Linda a bank teller?" (the way the answers to these questions are used to rank the propositions is detailed in the subsequent section). As a consequence, the quantum-like models are based on these questions and their answers, and not on the propositions to be evaluated themselves.

A vector space is introduced to represent the answers: answers "yes" and "no" to question F are respectively represented by the vectors $\boldsymbol{F_y}$ and $\boldsymbol{F_n}$, and similarly with $\boldsymbol{B_y}$

and $\boldsymbol{B_n}$ for answers to question B. The sets $(\boldsymbol{B_y}, \boldsymbol{B_n})$ and $(\boldsymbol{F_y}, \boldsymbol{F_n})$ respectively represent all possible answers to question B and F, and thus is a basis of the vector space, which is of dimension 2. Note that it is the same vector space that is used to represents answers to questions B and F; we have just indicated two different bases for this space.

The vector space is equipped with a scalar product, thus becoming a Hilbert space: for two vectors $\boldsymbol{W}$ and $\boldsymbol{X}$, the scalar product $\boldsymbol{W} \cdot \boldsymbol{X}$ is a complex number. The order of the vectors within a scalar product here matters: $\boldsymbol{X} \cdot \boldsymbol{W}$ is the complex conjugate of $\boldsymbol{W} \cdot \boldsymbol{X}$. The above bases are supposed to be orthogonal: $\boldsymbol{B_y} \cdot \boldsymbol{B_n} = \boldsymbol{F_y} \cdot \boldsymbol{F_n} = 0$, and of unitary norm: $\boldsymbol{B_y} \cdot \boldsymbol{B_y} = \boldsymbol{B_n} \cdot \boldsymbol{B_n} = \boldsymbol{F_y} \cdot \boldsymbol{F_y} = \boldsymbol{F_n} \cdot \boldsymbol{F_n} = 1$. A representation of the bases can be found on Figure 1 [Left].
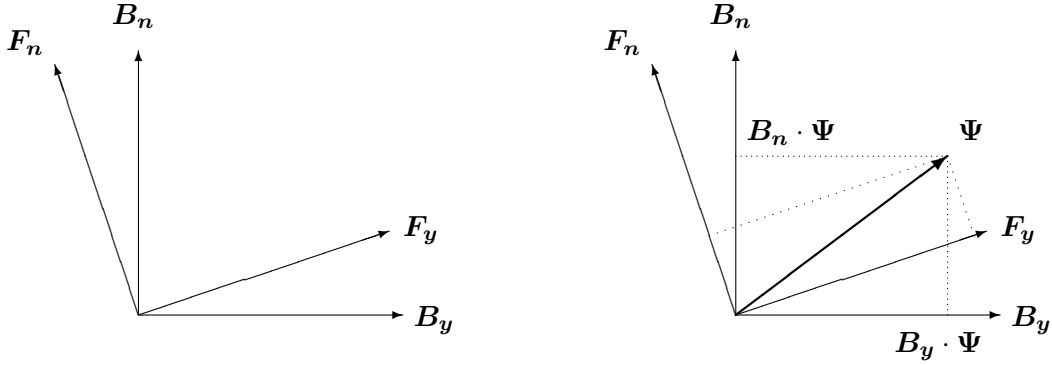

Figure 1: [Left:] The two bases corresponding to the answers "yes" and "no" to questions B and F. [Right:] The state vector $\boldsymbol{\Psi}$ can be decomposed on the two orthonormal bases (the scalar products on $\boldsymbol{B_y}$ and $\boldsymbol{B_n}$ are indicated). These figures assume the special case of a Hilbert space on real numbers.

An agent's state of belief is also represented within this Hilbert space, by a normalized vector $\boldsymbol{\Psi}$. This vector can be decomposed in either of the two above-mentioned bases, as indicated on Figure 1 [Right]:

$$\boldsymbol{\Psi} = (\boldsymbol{B_y} \cdot \boldsymbol{\Psi})\boldsymbol{B_y} + (\boldsymbol{B_n} \cdot \boldsymbol{\Psi})\boldsymbol{B_n} = (\boldsymbol{F_y} \cdot \boldsymbol{\Psi})\boldsymbol{F_y} + (\boldsymbol{F_n} \cdot \boldsymbol{\Psi})\boldsymbol{F_n}. \tag{1}$$

The belief state $\boldsymbol{\Psi}$ gathers all the relevant information needed to predict the behavior of the agent, in the following way. Predictions by the quantum-like models are probabilistic. When a question X ($X = B$ or F) is asked, the probability that the agent answers i ($i = y$ or n) is given by the squared modulus of the scalar product between the belief state and the vector representing the answer:

$$\Pr(\boldsymbol{X_i}) = |\boldsymbol{X_i} \cdot \boldsymbol{\Psi}|^2. \tag{2}$$

This rule is usually called the Born rule, in analogy with the quantum mechanics denomination. Thanks to this rule, one can compute the probability that the agent gives each of the 4 answers, in case questions B or F asked. One easily checks that $\Pr(\boldsymbol{X_y}) + \Pr(\boldsymbol{X_n}) = 1$, because $\boldsymbol{\Psi}$ is normalized. In the case of a real Hilbert space like on Figure 1, a geometric interpretation of the Born rule is the following: to compute the probability to answer, say, "yes" to question B, orthogonally project $\boldsymbol{\Psi}$ on $\boldsymbol{B_y}$ — this gives the length $\boldsymbol{B_y} \cdot \boldsymbol{\Psi}$; the desired probability is just the squared of this length. So, the more $\boldsymbol{\Psi}$ is aligned with a

basis vector $\boldsymbol{X_i}$, the larger the probability is that the agent will answer i if question X is posed (note the "if question X is posed" part: in quantum-like models, the probabilities of answers are defined only in the context when the corresponding question is posed).

The last postulate of the quantum-like model has to do with the way $\boldsymbol{\Psi}$ changes over time. First, when the agent doesn't answer a question, $\boldsymbol{\Psi}$ doesn't change. This conveys the fact that the agent's beliefs are not influenced in some way or another. This hypothesis is supposed to be relevant for cases in which the questions are posed to the agent relatively quickly. Second, when the agent answers a question B or F, her state of belief changes. If her answer to question X is i, then her new state of belief just after giving the answer is:

$$\boldsymbol{\Psi} \quad \longmapsto \quad \frac{\boldsymbol{X_i} \cdot \boldsymbol{\Psi}}{|\boldsymbol{X_i} \cdot \boldsymbol{\Psi}|} \boldsymbol{X_i}. \tag{3}$$

As the fraction is a complex number, the above equation means that the state of belief after an answer X_i is proportional to the vector representing this answer, $\boldsymbol{X_i}$. In the case of a real Hilbert space like on Figure 1, this means that after answering, say, "yes" to question B, $\boldsymbol{\Psi}$ becomes either $\boldsymbol{B_y}$ or $-\boldsymbol{B_y}$, whatever the state of belief was before the question. In other words, after a question X has been posed, the state of belief is bound to be along the basis vectors representing its answers. The Equation 3 can be interpreted as follows: the $(\boldsymbol{X_i} \cdot \boldsymbol{\Psi})\boldsymbol{X_i}$ part represents the fact that $\boldsymbol{\Psi}$ is projected on $\boldsymbol{X_i}$, the basis vector representing the given answer; the $1/|\boldsymbol{X_i} \cdot \boldsymbol{\Psi}|$ part is then just a multiplicative factor that ensures that the new state of belief is normalized. For that reason, the above rule is often called the projection postulate.

Because of the projection postulate, the states before and after an answer will be different, in general. The only case in which they are the same is when the state previous to the answer is proportional to one of the basis vectors representing the possible answers to the question, i. e. when $\boldsymbol{\Psi} = \lambda \boldsymbol{X_i}$, where λ is a complex number such that $|\lambda| = 1$ (in the real case, it amounts to $\boldsymbol{\Psi} = \pm \boldsymbol{X_i}$). In this case, indeed, there is a probability 1 that the agent answers i to question X, and Equation 3 just states that $\boldsymbol{\Psi} \quad \longmapsto \quad \boldsymbol{X_i}$. The fact that, in general, the state of belief changes when a question is answered is a real departure from classical theories. Classically, an answer to a question is supposed to *reveal* a belief, which is pre-existent to the question, and which is precisely the same before and after the question. But here, this cannot be the case: in general, the state of belief is modified by the fact that the question is posed and answered. There is a consolation, however: once a question has been posed, the same answer will be given if the same question is posed again just after.

2.2 Accounting for the fallacy

Now, given this quantum-like model, the explanation of Linda's experiment runs as follows. First, when an agent considers the conjunction "Linda is a bank teller and a feminist", this conjunction is evaluated as a sequence of projections, corresponding to the answers to two successive and dichotomous yes-no questions, namely taken among the set $\{F, B\}$. Secondly, the actual description of Linda "makes it very likely that she is feminist, very unlikely that she is a bank teller". Thirdly, the hypothesis is made that "the more probable possible outcome is evaluated first" — in other words, question F is evaluated before question B.

As illustrated in Figure 1, the probability of considering that Linda is a bank teller corresponds to the squared length of the projection of $\boldsymbol{\Psi}$ onto the bank teller vector $\boldsymbol{B_y}$, and $p(B) = |\alpha|^2$. On the other hand, the probability of considering her to be feminist and

bank teller corresponds to the squared length of the projection of $\boldsymbol{\Psi}$ onto two successive vectors, first the $|F\rangle$, and then $|B\rangle$, and $p(F \cap B) = |\beta|^2$.
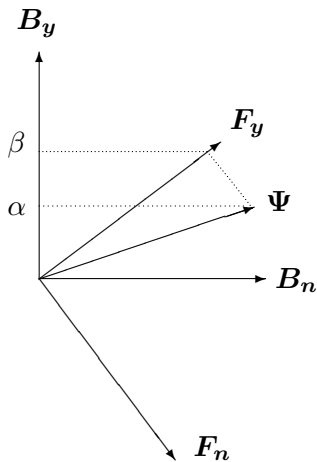

Figure 2: Illustration of QP explanation for Linda's experiment

So, as is highlighted in the literature, there exist some model configurations, like the one plotted in Figure 2, in which the probability to be judged feminist and bank teller is higher than the probability to be judged bank teller, leading to

$$p(F \cap B) > p(B), \tag{4}$$

in contradiction with classical models and in accordance with empirical results. A quantum-like model of the conjunction fallacy has been provided.

3   The Grand Reciprocity equations

From the quantum-like model presented in Section 2.1, some new empirical predictions can be derived. We follow here the work of Boyer-Kassem, Duchêne and Guerci (2015), that have established these predictions in the case of a more general quantum-like model.

As the model is non-degenerate, i.e. the answers are represented by subspaces of dimension 1, it can be shown that a well-known law from quantum mechanics, the law of reciprocity, holds. In our model, this law states that, for $(x, y) \in \{B, F\}^2$, and $(i, j) \in \{y, n\}^2$,

$$p(y_j|x_i) = p(x_i|y_j). \tag{5}$$

This law invites to consider the two questions F and B in one order or in the other, and asserts that conditional probabilities of an answer given another answer are the same whatever the order of the questions B and F. Note that this law is typically quantum: it is not true in general a classical Bayesian model, in which $p(y_j|x_i) = p(x_i|y_j) \times p(y_j)/p(x_i)$, and thus $p(y_j|x_i) \neq p(x_i|y_j)$ as soon as $p(y_j) \neq p(x_i)$.

This law of reciprocity is quite well-known in quantum-like models[1]. Here, it can be instanciated in the following ways:

---

[1]Cf. Franco (2009, 417–418), Wang and Busemeyer (2009, 697–698), Busemeyer *et al.* (2011, 197), Pothos and Busemeyer (2013, 317), Wang *et al.* (2014, 5).

$$\begin{cases} p(B_y|F_y) = p(F_y|B_y), & (6) \\ p(B_n|F_y) = p(F_y|B_n), & (7) \\ p(B_y|F_n) = p(F_n|B_y), & (8) \\ p(B_n|F_n) = p(F_n|B_n). & (9) \end{cases}$$

Some easy computation enable to show that the following equations, called the Grand Reciprocity (GR) equations, hold (cf. Boyer-Kassem *et al.* 2005, Section 3.1):

$$\begin{cases} p(B_y|F_y) = p(F_y|B_y) = p(B_n|F_n) = p(F_n|B_n), & (10) \\ p(B_n|F_y) = p(F_y|B_n) = p(B_y|F_n) = p(F_n|B_y). & (11) \end{cases}$$

These equations 10 and 11 are actually equivalent to one another, and equivalent to the reciprocity law itself. [2] They state that the conditional probabilities that exist in the B-then-F and F-then-B configurations are actually much constrained: among the eight quantities that can be experimentally measured, there is just one free real parameter. In other words, the quantum-like model presented in Section 2.1 actually leaves very little freedom to conditional probabilities.

To the best of our knowledge, the fact that the conditional probabilities are constrained by the GR equations had not been noticed beforehand for quantum-like models for conjunction fallacy. Our bottom line in this paper is to experimentally test these empirical claims the model makes. Note that these empirical predictions are *consequences* of the quantum-like model that is used to explain the conjunction fallacy in the Linda experiment, and that these consequences are observable in experimental situations — B-then-F and F-then-B situations — that are *not* the ones of the original Linda experiment. In other words, the GR equations show that the quantum-like model that is used to explain a Linda experiment can be further tested on another kind of experiment. This is what the next Section considers.

4 Experimental protocol

In order to test the GR equations, which are empirical predictions of the QL models for conjunction fallacy we consider, we have realized the above-mentioned experimental situations in which agents are asked B-then-F questions, or F-then-B ones.

The experiment was conducted October 10th, 2014, in the experimental economics laboratory (LEEN) of the University Nice Sophia Antipolis (France). After entering in the laboratory, subjects received a very brief oral description of the task, namely that they first read a short text, and then answer to some questions. The experiment is a computerized one, with the z-Tree computer program (Fischbacher 1999). The first screen of the computer presented the standard description of Linda (please refer to appendix A for screenshots) and after a few minutes the two distinct questions,

F: "Is Linda feminist?"

B: "Is Linda a bank teller?"

---

[2]For some generalizations of these GR equations, and comments on the link between these equations and other concepts like double-stochasticity, cf. Boyer-Kassem *et al.* (2015), Section 3.

were proposed in sequence. In the first treatment (called FB treatment) questions were asked in the order F-then-B, in the second one (called BF treatment) questions were asked in the inverse order B-then-F. Seven sessions involving $N_{FB} = 187$ students were conducted for FB, whereas six sessions involving $N_{BF} = 167$ students were conducted for BF. In total, 354 undergraduate and graduate students were randomly recruited at the end of classes in economics or in management that were not aware of the original Linda experiment. The experiment was conducted on purpose in only one day: all necessary precautions were taken to organize the sessions in such a way to avoid discussions among students having and not having performed the experiment — these aspects are of course of dramatic importance in an experiment aiming at detecting subtle psychological biases. Because of the "improvised" recruitment (students didn't join the laboratory for a scheduled appointment), of the very short length of the task and of the absence of clear monetary incentive schemes, we have adopted the methodological choice of not paying subjects — a common choice in the literature.

5   Results

Let's define $\mathbf{F} \in \{F_y, F_n\}$ as the binary random variable represented by question F assuming only two possible values F_y for "yes" and F_y for "no". Similarly, $\mathbf{B} \in \{B_y, B_n\}$ is the random variable represented by question B assuming values B_y for "yes" and B_y for "no". Both treatments FB and BF are thus multinomial experiments, that is, statistical experiments described by multinomial distributions. Each treatment has four possible outcomes, for instance $(B_y, F_y), (B_n, F_y), (B_y, F_n), (B_n, F_n)$ for FB. The frequency $N(B_j, F_i)$ is defined as the outcome of the counting process of independent and repeated trials represented by the people responding j to the first question B and then i to the second question F (the order matters in our notation: $N(B_j, F_i)$ refers to the B-then-F experiment).

Table 1 contains the two cross tabulations for both treatements FB and BF. The second row reports the frequencies $N(F_i, B_j)$ and $N(B_j, F_i)$, respectively.

Table 1: Cross tabulations.

FB treatment

Outcomes	(F_y, B_y)	(F_n, B_y)	(F_y, B_n)	(F_n, B_n)	Total
Frequencies	5	8	57	117	187

BF treatment

Outcomes	(B_y, F_y)	(B_n, F_y)	(B_y, F_n)	(B_n, F_n)	Total
Frequencies	7	6	86	68	167

Given such frequencies, one can easily compute relative frequencies $f(F_i, B_j) = \frac{N(F_i, B_j)}{N_{FB}}$ and $f(B_j, F_i) = \frac{N(B_j, F_i)}{N_{BF}}$ and relative conditional frequencies $f(F_i | B_j) = \frac{N(F_i, B_j)}{N(F_i, \cdot)}$ and $f(B_j | F_i) = \frac{N(B_j, F_i)}{N(\cdot, B_j)}$, where the dot notation indicates the marginal distributions.

As we aim at testing the GR equations, we can focus our statistical test on the first set of 6 equations (eq. 10) only, as it is equivalent to the second one (eq. 11). Because the conditional probabilities are not directly observable, conditional relative frequencies are adopted as estimations. The GR condition can be rewritten in terms of conditional

relative frequencies:

$$f(B_y|F_y) = f(F_y|B_y) = f(B_n|F_n) = f(F_n|B_n).$$

which thus consists in 6 two-by-two statistical tests. From a statistical viewpoint, this implies the necessity of performing a correction of the type I error because of the multiple comparisons. We apply the most conservative one, the Bonferroni correction.

In order to compare the conditional relative frequencies y and x, with the null hypothesis that they are equal, a logit transformation of the values is performed:

$$\log\left(\frac{y}{1-y}\right) = \log\left(\frac{x}{1-x}\right),$$

which implies equivalently that

$$\log\left(\frac{x(1-y)}{(1-x)y}\right) = \log(\text{OR}) = 0. \tag{12}$$

The logarithm of the odds ratio X is a standard concept in statistical inference that is characterized by an approximately normal sampling distribution ($X \sim \mathcal{N}(\log(\text{OR}), \sigma^2)$) under the hypothesis of a large sample approximation.

Thus, the first statistical test $f(B_y|F_y) = f(F_y|B_y)$ easily leads to the following equivalent formulation,

$$\log\left(\frac{N(F_y, B_y)N(B_y, F_n)}{N(F_y, B_n)N(B_y, F_y)}\right) = 0, \tag{13}$$

by expressing $x = f(B_y|F_y) = \frac{N(F_y, B_y)}{N(F_y, \cdot)}$ and $(1-x) = f(B_n|F_y) = \frac{N(F_y, B_n)}{N(F_y, \cdot)}$ in eq. (12), and similarly for y. Therefore, the information contained in the two cross tabulations can be used to perform the wanted analysis. Further, the continuity correction is applied, because the normal approximation to the binomial is used, which is effective in particular for small values of $N(F_i, Bb_j)$ or $N(B_j, F_i)$. The above test becomes

$$\log\left(\frac{(N(F_y, B_y) + 0.5)(N(B_y, F_n) + 0.5)}{(N(F_y, B_n) + 0.5)(N(B_y, F_y) + 0.5)}\right) = 0. \tag{14}$$

The standard error of the log odds ratio for the considered test is:

$$\text{SE}_{\text{OR}} = \sqrt{\frac{1}{N(F_y, B_y)} + \frac{1}{N(B_y, F_n)} + \frac{1}{N(F_y, B_n)} + \frac{1}{N(B_y, F_y)}}. \tag{15}$$

It can be easily derived for each of the six tests by considering the square root of the sum of the inverse of the four frequencies which are present in the OR.

Table 2 shows p-values for each of the six tests which have been carried out under the hypothesis of

$$\frac{\log(\text{OR})}{\text{SE}_{\text{OR}}} \sim \mathcal{N}(0, 1), \tag{16}$$

Table 2 shows that 5 out of the 6 statistical tests reject the null of equality between the two conditional frequencies — recall that only one rejection is needed to state that the GR equations are not verified.

The interpretation of these results is quite clear: the non-degenerate QL model we consider is not empirically adequate, as it predicts the GR equations which happen not to be verified in our Linda-like experiment. As a consequence, this model cannot claim to account for the conjunction fallacy.

11

Table 2: *p*-values for each test. The rejection is highlighted (*) at the 1% significance level.

T1	T2	T3	T4	T5	T6
0.00018*	0.00000*	0.00000*	0.00006*	0.25065	0.00000*

6 Discussion: the future of quantum-like models for conjunction fallacy

Beyond the negative rejection of the quantum-like models discussed in the previous section, can we offer a more constructive suggestion? And how should we consider the future of research in the modelling of conjunction fallacy, after these results? Several lines of possible research can be distinguished.

First, it is true that there are other possible lines of research than standard quantum-like models. For instance, Khrennikov has proposed to consider hyperbolic Hilbert space models, instead of complex ones like in this paper. Hyperbolic models enable probabilities that are not constrained in the same way, and can account for some data that are out of reach for complex quantum-like models (see Khrennikov 2010 for a synthesis). Also, one could try to abandon or modify other axioms of standard quantum-like models, like the projection postulate or the Born rule.

But all quantum-like models have not been ruled out here: those to which the GR equations do not apply, i.e. which consider degenerate eigenvalues, have not been tested. So, a natural line of research is to turn to degenerate quantum-like models, where answers are represented by sub-spaces of dimension 2 or larger. Our results do not indicate the end of quantum-like models *in general*. They can be seen as suggesting instead more research on quantum-like models, albeit on degenerate ones.

However, some difficulties with switching to degenerate models have to be acknowledged. First is the question whether the replacing degenerate model is empirically equivalent to, and thus can be reduced to, a non-degenerate model. If this is the case, then the GR equations do apply. So, once a non-degenerate model has been ruled out, a degenerate model is not an escape route unless it is shown not to be reducible to a non-degenerate one. Second, degenerate models are not *a priori* freed from any constraint. Some other equations than the GR equations might hold — some research is urgently needed here. So, we suggest that, when research efforts are made to have degenerate models fit some data sets, similar efforts be also directed towards the testing of these models, in all possible forms, so as to prevent later "bad surprises". Another difficulty with developing degenerate models is that introducing supplementary dimensions of degeneracy should be somehow justified, so as not to be accused of being just *ad hoc*. In quantum physics for instance, degeneration is usually theoretically justified with a symmetry in the Hamiltonian, and the degeneration can be experimentally removed by introducing some factor that was not present (e.g., the spin degeneration is removed by introducing a magnetic field). For quantum-like judgment models as well, one could try to provide a similar theoretical justification and an experimental way of removing the degeneration. In this respect, the idea that there exists fundamental questions and rays could be helpful; the work is then to find out which fundamental dimensions are hidden behind the degenerate eigenspace.

In spite of the possible difficulties that we have just discussed, we remain convinced that quantum-like models are a promising research line. They have brought to discussion many provoking and seminal ideas, such as the hypotheses that preferences might be

underdetermined instead of only unknown, or that non-classical probabilities could be considered. At the same time, our contribution has clarified some important aspects of quantum-like probabilistic models applied to human cognition. In the past, these aspects have been neglected by the quantum literature. We believe that our results pose the appropriate challenge to the scientific community to guide them on where to make future investigations in the field. Indeed, our results suggest that non-degenerate quantum-like models should be considered more as toy models than as empirically adequate models, and that future investigations should focus on degenerate models.

Acknowledgments

References

AERTS, Diederik (2009), "Quantum Structure in Cognition", *Journal of Mathematical Psychology* 53: 314–348.

AERTS, Diederik, Liane GABORA and Sandro SOZZO (2013), "Concepts and Their Dynamics: A Quantum-Theoretic Modeling of Human Thought", *Topics in Cognitive Sciences* 5: 737–772.

ALLAIS, Maurice (1953), "Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école Américaine", *Econometrica* 21:503–546.

ASHBY, F. Gregory and Nancy A. PERRIN (1988), "Towards a unified theory of similarity and recognition", *Psychological Review*, 95:124–50.

BOYER-KASSEM, Thomas, Sébastien DUCHÊNE and Éric GUERCI (2015), "Testing quantum-like models of judgment for question order effects", GREDEG working paper 2015-06, `http://www.gredeg.cnrs.fr/working-papers.html`.

BUSEMEYER, Jerome R. and Peter D. BRUZA (2012), *Quantum models of Cognition and Decision*, Cambridge: Cambridge University Press.

BUSEMEYER, Jerome R., Emmanuel M. POTHOS, Riccardo FRANCO and Jennifer S. TRUEBLOOD (2011), "A Quantum Theoretical Explanation for Probability Judgment Errors", *Psychological Review* 118 (2): 193–218.

BUSEMEYER, Jerome R., Zheng J. Wang, and Ariane LAMBERT-MOGILIANSKY (2009), "Empirical comparison of Markov and quantum models of decision making", *Journal of Mathematical Psychology* 53: 423–433.

CHARNESS, Gary, Edi KARNI, and Dan LEVIN (2010), "On the conjunction fallacy in probability judgment: New experimental evidence regarding Linda", *Games and Economic Behavior* 68: 551–556.

CONTE, Elio, Andrei Yuri KHRENNIKOV, Orlando TODARELLO, Antonio FEDERICI, Leonardo MENDOLICCHIO, and Joseph P. ZBILUT (2009), "Mental States Follow Quantum Mechanics During Perception and Cognition of Ambiguous Figures", *Open Systems & Information Dynamics* 16(1) 1–17.

ELLSBERG, Daniel (1961), "Risk, ambiguity, and the Savage axioms", *Quarterly Journal of Economics* 75(4): 643–669.

ERCEG, Nikola and Zvonimir GALIC (2014), "Overconfidence bias and conjunction fallacy in predicting outcomes of football matches", *Journal of Economic Psychology* 42 :

52–62.

FRANCO, Riccardo (2009), "The conjunction fallacy and interference effects", *Journal of Mathematical Psychology* 53: 415–422.

GIGERENZER, Gerd (1991), "How to make cognitive illusions disappear: Beyond "heuristics and biases."", *European Review of Social Psychology* 2: 83–115.

HAVEN, Emmanuel and Andrei KHRENNIKOV (2013), *Quantum Social Science*, Cambridge: Cambridge University Press.

HARTMANN, Stephan and Wouter MEIJS (2012), "Walter the banker: the conjunction fallacy reconsidered", *Synthese* 184: 73–87.

HERTWIG, Ralph (1997), "Judgment under uncertainty: Beyond probabilities", Berlin, Germany: Max Planck Institute for Human Development. Unpublished manuscript.

HERTWIG, Ralph, and Valerie M. CHASE (1998), "Many Reasons or Just one: How Response Mode Affects Reasoning in the Conjunction Problem" *Thinking and Reasoning* 4(4): 319–352

HERTWIG, Ralph, and Gerd GIGERENZER (1999), "The 'Conjunction Fallacy' Revisited: How Intelligent Inferences Look Like Reasoning Errors" *Journal of behavioral decision making* 12: 275–305.

HERTWIG, Ralph, Björn BENZ and Stefan KRAUSS (2008), "The conjunction fallacy and the many meanings of and", *Cognition* 108: 740–753.

KAHNEMAN, Daniel, and Amos TVERSKY (1996), "On the reality of cognitive illusions: A reply to Gigerenzer's critique" *Psychological Review* 103: 582–591.

KHRENNIKOV, Andrei (2010), *Ubiquitous Quantum Structure. From Psychology to Finance*, Heildelberg: Springer.

KRUMHANSL, Carol L. (1978), "Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density", *Psychological Review* 85: 445–63.

MELLERS, Barbara, Ralph HERTWIG and Daniel KAHNEMAN (2001), "Do frequency representations eliminate conjunction effects? An Exercise in Adversarial Collaboration", *Psychological Science* 12: 269–275.

MOORE, David W. (2002), "Measuring new types of question-order effects", *Public Opinion Quarterly*, 66: 80–91.

MORO, Rodrigo (2009), "On the nature of the Conjunction Fallacy", *Synthese* 171: 1–24.

NILSSON, Hâkan, and Patric ANDERSSON (2010), "Making the seemingly impossible appear possible: Effects of conjunction fallacies in evaluations of bets on football games", *Journal of Economic Psychology* 31: 172–180.

POTHOS, Emmanuel M. and Jerome R. BUSEMEYER (2013), "Can quantum probability provide a new direction for cognitive modeling?", *Behavioral and Brain Sciences* 36: 255–327.

SCHUMAN, Howard and Stanley PRESSER (1981), *Questions and answers in attitude surveys: Experiments on question form, wording, and context*, Academic Press.

SHAFIR, Eldar and Amos TVERSKY (1992), "Thinking through uncertainty: nonconsequential reasoning and choice", *Cognitive Psychology* 24:449–74.

STOLARZ-FANTINO, Stepanie, Edmund FANTINO, Daniel J. ZIZZO, and Julie WEN (2003), "The conjunction effect: New evidence for robustness", *American Journal of Psychology* 116(1):15–34.

TENTORI, Katya, Nicolao BONINI, and Daniel OSHERSON (2004), "The conjunction fallacy: a misunderstanding about conjunction", *Cognitive Science* 28: 467–477.

TENTORI, Katya and Vincenzo CRUPI (2008), "On the conjunction fallacy and the meaning of *and*, yet again: A reply to Hertwig, Benz, and Krauss (2008)", *Cognition* 122: 123–

134.

Tentori, Katya and Vincenzo Crupi (2012), "How the conjunction fallacy is tied to probabilistic confirmation : Some remarks on Schupbach (2009)", *Synthese* 184: 3–12.

Tourangeau, Roger, Lance J. Rips and Kenneth A. Rasinski (2000), *The psychology of survey response*, Cambridge University Press.

Tversky, Amos (1977), "Features of similarity", *Psychological Review* 84(4): 327–52.

Tversky, Amos and Daniel Kahneman (1982), "Judgments of and by representativeness", in: *Judgment under uncertainty: Heuristics and biases*, ed. by D. Kahneman, P. Slovic & A. Tversky, p. 84–98, Cambridge University Press.

— (1983), "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment", *Psychological Review* 90(4): 293–315.

von Sydow, Momme (2011), "The Bayesian logic of frequency-based conjunction fallacies", *Journal of Mathematical Psychology* 55: 119–139.

Wang, Zheng J. and Jerome R. Busemeyer (2013), "A quantum question order model supported by empirical tests of an a priori and precise prediction", *Topics in Cognitive Science* 5(4): 689–710.

Wang, Zheng J., Tyler Solloway, Richard M. Shiffrin and Jerome R. Busemeyer (2014), "Context effects produced by question orders reveal quantum nature of human judgments", *PNAS*, `www.pnas.org/cgi/doi/10.1073/pnas.1407756111`.

Wedell, Douglas H., and Rodrigo Moro (2008), "Testing boundary conditions for the conjunction fallacy: Effects of response mode, conceptual focus, and problem type", *Cognition* 107: 105–136.

A   Screenshots

We report the three screenshots of the experiment ordered according to the F->B treatment. The text is in French.
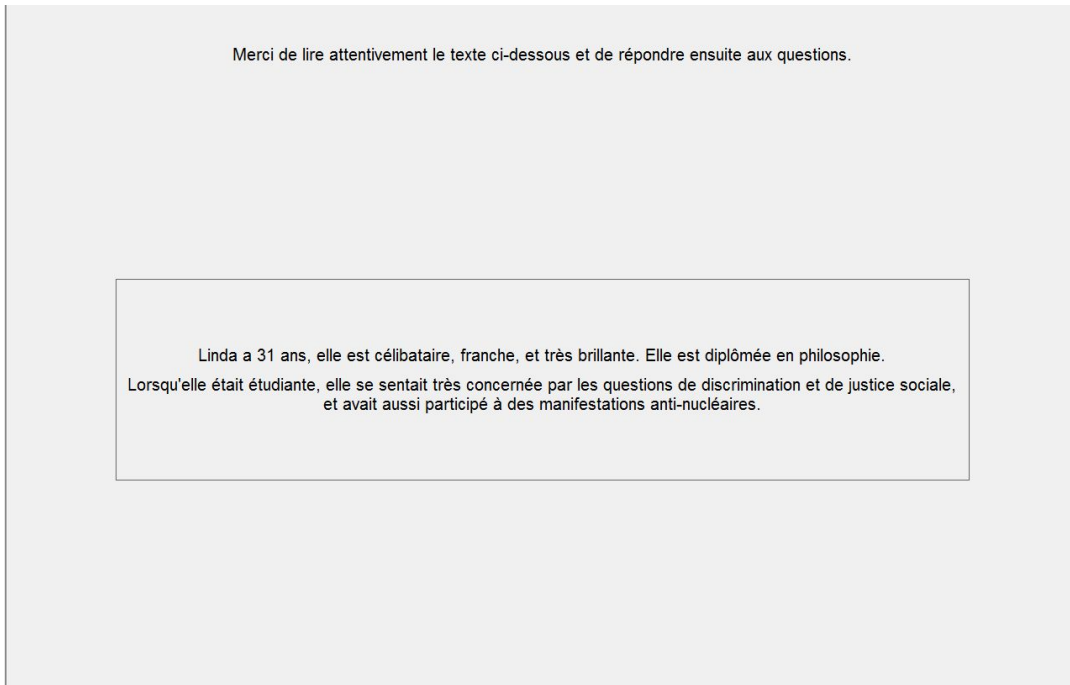
Merci de lire attentivement le texte ci-dessous et de répondre ensuite aux questions.

Linda a 31 ans, elle est célibataire, franche, et très brillante. Elle est diplômée en philosophie.

Lorsqu'elle était étudiante, elle se sentait très concernée par les questions de discrimination et de justice sociale, et avait aussi participé à des manifestations anti-nucléaires.

Figure 3: First screen - Description of Linda

Première question (cliquez sur le bouton que vous souhaitez sélectionner):

Selon vous,
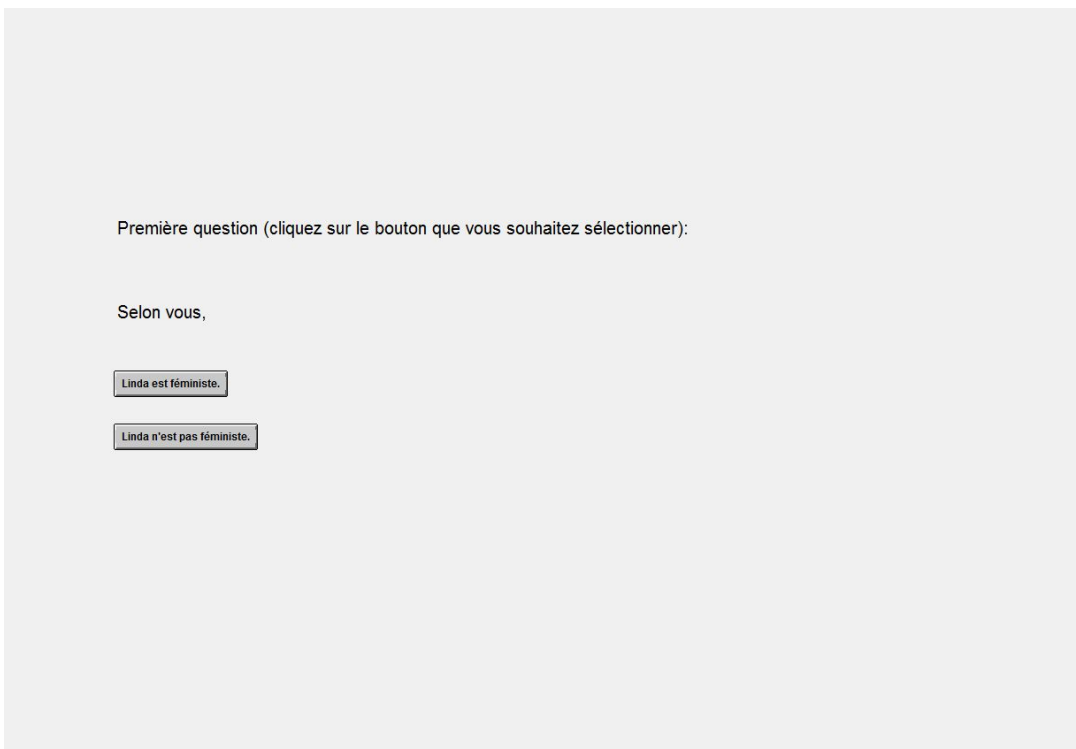
Linda est féministe.

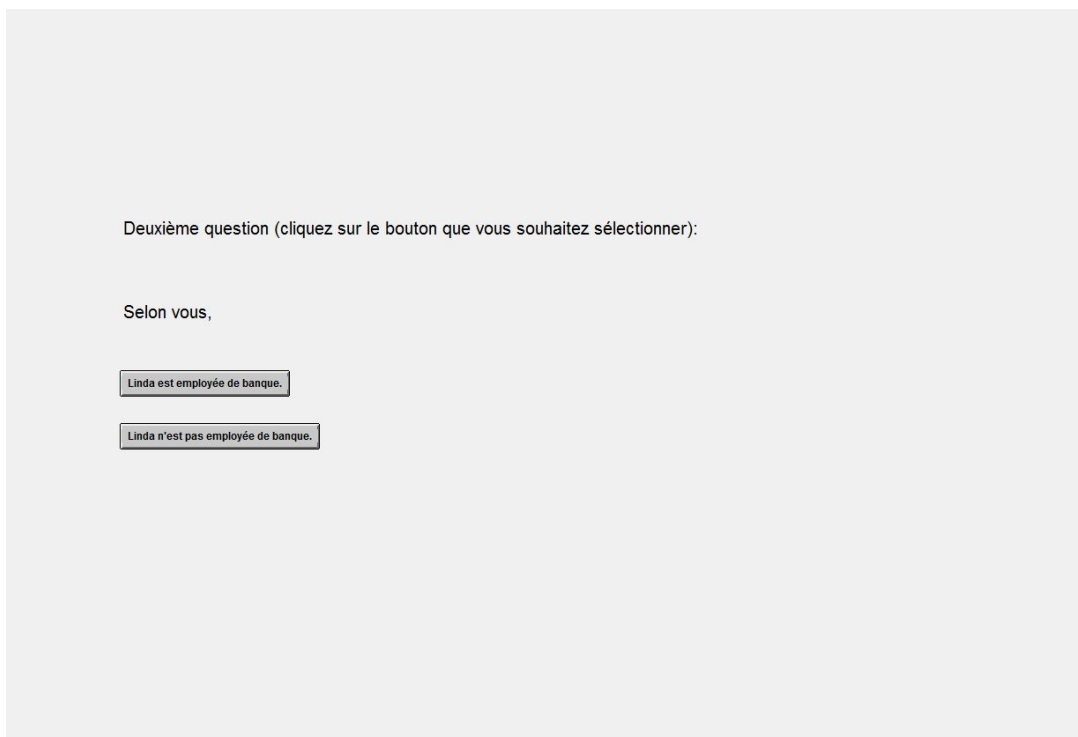Linda n'est pas féministe.

Figure 4: Second screen, on Linda being a feminist.

Figure 5: Third screen, on Linda being a bank teller.