Empirical constraints on computational level models of interference effects in human probabilistic judgements.

James M. Yearsley (james.m.yearsley@vanderbilt.edu)
Department of Psychology, Vanderbilt University, Nashville TN, USA

Emmanuel M. Pothos (emmanuel.pothos.1@city.ac.uk)

Albert Barqué-Duran (albert.barque-duran@city.ac.uk)
Department of Psychology, City University London, UK

Abstract

Decades of research in decision making have established that there are some situations where human judgments cannot be modelled according to classical probability theory. Yet if we abandon classical (Bayesian) probability theory as an overarching framework for constructing cognitive models, what do we replace it with? In this contribution we outline a way to divide the space of possible computational level models of probabilistic judgment into a hierarchy of levels of increasing complexity, with classical Bayesian probability models occupying the lowest level. Each level has a unique experimental signature, and we examine which level is best able to describe human behavior in a particular probabilistic reasoning task.

Keywords: Probabilistic reasoning, disjunction fallacy, quantum theory.

Introduction

2017 marks 35 years since the publication of the volume "Judgment under Uncertainty: Heuristics and Biases" edited by Kahneman, Slovic and Tversky (1982), which has since become a seminal text for those interested in violations of normative reasoning. The years since have seen a great deal of research aimed at better understanding the conditions under which human decision makers do or do not make 'normative' decisions, that is, decisions that can be thought of as being 'correct' by some measure. Note that in this contribution, in common with many other studies, we take the definition of 'normative' to be essentially equivalent to compatibility with classical/Bayesian probability theory. That is, behaving in a normative way is less about providing judgments that are 'accurate' in the sense of reflecting real likelihoods, and rather about self consistency (Oaksford & Chater, 2009).

Despite the large amount of research that has been done to understand when and why these deviations from normative prescription occur, there are some notable gaps in our understanding of the structure of non-normative judgments. In part this has been caused by a tendency amongst researchers to define non-normative judgments by the properties they do not possess, ie a simple relation to an underlying probability distribution representing a belief state. This has led to a peculiar splintering of the study of decision making, most visible in the heuristics programme, where different decision making tasks are assumed to be executed with the aid of different heuristics (eg Gigerenzer et al, 2015). Few have considered the relationships between heuristics, or whether they reflect some deeper structure.

A different approach, advocated by some adherents of the Bayesian cognition program (see Jones & Love, 2011), has been to deal with violations of normative (Bayesian) prescription by including extra variables and relations that, while they may be framed in a Bayesian way, conflict somewhat with the spirit of the Bayesian approach. An example is provided by Bayesian efforts to deal with order effects in probabilistic inference, where, for example, decision makers may judge $p(E|A,B) \neq p(E|B,A)$. One possible solution is to posit that decision makers are sensitive also to the order in which evidence is presented, and that the probabilities for judgments therefore should also be conditioned on the order of presentation. Formally this saves face for the Bayesian approach, but it is hard to see how this does anything other than redescribe the problem. A genuine solution to the problem of order effects would also need to explain why decision makers possess this sensitivity, and make testable predictions.

What we want to do in this contribution is approach the problem from a different angle. Suppose we assume that there is some sort of computational structure underlying non-normative behaviour, presumably that can be understood as a generalisation of the Bayesian approach, and we ask whether we can constrain the structure of this theory in some way. Specifically, can we derive predictions that are sensitive only to some general fact about the structure of the theory and not to the precise details? If we can do this we can perhaps sketch out a space of acceptable computational level theories which can then be investigated in more detail by other means.

This approach is inspired in part by recent work in the field of quantum cognition (Pothos & Buseyemer, 2013), which is the attempt to use the formal probability theory derived from quantum physics to describe human decision making. The advantage of such an approach, so proponents claim, is that quantum probability theory is a formal, all encompassing framework that is nevertheless able to account for deviations from classical, Bayesian, behaviour. However quantum probability theory has a very particular structure (see Busemeyer & Bruza 2012), and it is not obvious that this is either necessary or sufficient to explain human behaviour. The study reported here can therefore be thought of as an attempt to test the sufficiency of the quantum approach for explaining certain types of non-normative behaviours. However it is also much more general, since we will see that quantum theory is just one example of a specific class of models sharing the same structural properties. In order to avoid getting bogged down in arguments about the suitability of quantum theory to describe human decision making, we will avoid making reference to specific computational models beyond classical probability theory.

The rest of this contribution is structured as follows: In Section 2 we explain how we can situate different computational level models of decision making in a hierarchy, and we show that classical and quantum probability theories belong to two different levels, each with a unique experimental signature. In Section 3 we describe an experiment designed to test which level in the hierarchy is necessary and sufficient to explain human behaviour in a particular probabilistic reasoning task. In Section 4 we conclude and outline some future directions.

A Hierarchy of Formal Theories

Given two disjoint events, A and B, a key property of classical probability theory is that the probability of the disjunction, $A \cup B$ is equal to the sum of the probabilities of the individual events. This can be generalised for an arbitrary number of disjoint events as,

$$p(\cup_i A_i) = \sum_i p(A_i) \tag{1}$$

This result expresses the property of classical probability known as linearity (more formally σ-additivity, Kolmogorov, 1933/1950). Classical probability models are therefore examples of linear models mapping events to probabilities.

In order to allow for the possibility of response errors it is useful to generalise this notion somewhat. Suppose we consider a more general class of models where, $p(A) = f(A) + \varepsilon$, where ε is a constant. We will call these models linear if,

$$f(\cup_i A_i) = \sum_i f(A_i) \tag{2}$$

Classical probability theory is obviously a special case of these linear models. These models can be used to capture the idea that judgments are noisy, so for example the probability assigned to the null event is not 0. They can also capture simple types or response biases, for example an aversion to using the endpoints on a response scale.

It is important to note that general linear models can violate some properties of classical probability theory. For example for two disjoint events A, B we have,

$$p(A \cup B) = f(A \cup B) + \varepsilon = f(A) + f(B) + \varepsilon = p(A) + p(B) - \varepsilon \tag{3}$$

Which is a violation of the classical law for disjoint events. However it is easy to see that any law of classical probability that is violated by a general linear model will be violated by some multiple of ε, and so these sorts of effects are easy to spot. In particular, consider the quantity,

$$\mathscr{I}_2(A,B) = p(A \cup B) - p(A) - p(B) = -\varepsilon \tag{4}$$

We call expressions of this form 'interference' terms, because of the analogy with quantum models. In a simple classical probability theory model this quantity is 0. In a general linear model this term is no longer zero, but it is instead equal to some constant which should be the same for all events A, B. The way to test the general class of linear models is therefore to examine different sets of events and test whether $\mathscr{I}_2(A,B)$ depends on the alternatives under consideration. This generalises the notion of the failure of a simple classical model.

The occurrence of widespread violations of classical probability rules in decision making experiments, including conjunction and disjunction fallacies (Tversky & Kahneman, 1983), has lead many to reject linear models as accounts of the way human decision makers assign probabilities to events. However few have explored proposals for concrete models beyond linear ones.

A natural generalisation of linear models is to consider those which contain a *bilinear* term. The appropriate generalisation is to have,

$$p(A) = g(A,A) + f(A) + \varepsilon \tag{5}$$

where $f(\cdot)$ is linear as above, and $g(\cdot,\cdot)$ is linear in both its arguments,

$$g(\cup_i A_i, B) = \sum_i g(A_i, B), \quad g(A, \cup_i B_i) = \sum_i g(A, B_i), \tag{6}$$

If we now consider the probability assigned to the disjunction $A \cup B$ we see,

$$\begin{aligned} p(A \cup B) =& g(A \cup B, A \cup B) + f(A \cup B) + \varepsilon \\ =& g(A,A) + g(B,B) + g(A,B) + g(B,A) \\ &+ f(A) + f(B) + \varepsilon \\ =& p(A) + p(B) + g(A,B) + g(B,A) - \varepsilon \end{aligned} \tag{7}$$

Considering again the quantity $\mathscr{I}_2(A,B)$ we see,

$$\begin{aligned} \mathscr{I}_2(A,B) =& p(A \cup B) - p(A) - p(B) \\ =& g(A,B) + g(B,A) - \varepsilon \end{aligned} \tag{8}$$

which is generally non-zero, but also crucially will depend on A and B, allowing us to distinguish bilinear models from linear ones.

Although we will not prove this here, the class of bilinear models includes quantum theory as a special case (Sorkin, 1994). However clearly this framework is more general than any specific model.

Once we have made the choice to step beyond linear models, a whole world of possibilities is opened up. Why stop at bilinear models? Why not consider a model containing a *trilinear* a function $h(A,B,C)$? The answer is that such models are possible, and in the same way that the quantity Eq.(8) provides us with a way to test between linear and bilinear models, a generalisation of this lets us distinguish between bilinear models and more complicated theories.

This rule comes from exploring what happens when we have three possible disjoint events A, B, C. Because $g(A,B)$ is bilinear it is straightforward, if tedious, to show,

$$\begin{aligned} p(A \cup B \cup C) = & p(A \cup B) + p(B \cup C) + p(A \cup C) \\ & - p(A) - p(B) - p(C) + \varepsilon \end{aligned} \tag{9}$$

So if we define an analogue of $\mathscr{I}_2(A,B)$ for three events, $\mathscr{I}_3(A,B,C)$, then we have,

$$\begin{aligned} \mathscr{I}_3(A,B,C) = & p(A \cup B \cup C) - p(A \cup B) - p(B \cup C) - p(A \cup C) \\ & + p(A) + p(B) + p(C) \\ = & \varepsilon \end{aligned} \tag{10}$$

so this three way interference term is constant for all events A, B, C if the underlying model is bilinear. This does not hold in higher order theories, such as ones based on a trilinear function (Sorkin, 1994). Therefore this provides a test of bilinear models versus more complex theories. Note also that this relation holds trivially for a linear model.

Although we will not show it here, it is straightforward to prove by induction that at every level in this hierarchy of possible model types there is a corresponding interference term $\mathscr{I}_n(A,B,C...)$ which is constant for any model at that level or below (Sorkin, 1994). In other words we can make a very general statement; The experimental signature of a level n model (ie a model based on a 'n-linear' function) is that a) The quantity $\mathscr{I}_n(A,B,C...)$ depends on the events $A, B, ...$, but b) the quantity $\mathscr{I}_{n+1}(A,B,C...)$ is constant.

We know from previous work that any theory that can capture human behavior, in particular disjunction fallacies, must be at least level two. The question is whether a level two theory is also sufficient, that is, whether there are particular effects in human decision making that arise when considering three alternatives. We will provide experimental evidence below that a level two theory is sufficient to explain behavior in a particular decision making task, however we can also motivate sufficiency on general grounds. Looking at the structure of Eq(8) for a bilinear theory, we see that the important terms involve the general function $g(\cdot,\cdot)$ evaluated on two different events, A, B. This framework is a computational level account, but presumably such a term must arise from a process level account wherein the two events A, B are processed in parallel. In contrast a linear theory only ever involves functions with a single argument, and thus would not require a process level account with simultaneous consideration of multiple events. This strongly suggests that there is a specific sense in which a bilinear model requires a more complex underlying process to instantiate it. By analogy, the analagous term in a level n model will involve a function of n arguments, and, presumably, would require a process model in which n events are considered in parallel.

If a higher order computational theory requires a more complex underlying process to produce it, then we can argue on general grounds that it is unlikely that human decision making is described by a theory of very high order. Of course, this argument does not tell us whether to expect a bilinear, trilinear etc model, only that a lower level is likely to be preferred over a very high one. The question of exactly what level is needed is an empirical one, which we shall now examine.

Experiment

We want to test the hypothesis that a bilinear model is necessary and sufficient to capture non-normative effects in human probabilistic reasoning. To do this we need to find sets of at least three disjoint alternatives such that are robust two way disjunction fallacies, in the sense that the interference term $\mathscr{I}_2(A,B)$ depends on the events A, B. Our approach will be to set up three scenarios, each with three disjoint events, and show that the term $\mathscr{I}_2(A,B)$ can be manipulated by introducing joint causes for some of the events. This will prove that a cognitive model capturing these judgments must be at least bilinear. We will then examine the higher order interference term $\mathscr{I}_3(A,B,C)$ in each scenario to check that a bilinear model is sufficient. The joint causes we will introduce will either cause the three events to be grouped into two natural sets, with one element shared between sets, or into one set with one singleton event. There are therefore two different ways of presenting each scenario, so we run each as a between participants condition.

Methods

We recruited 300 participants, equally split into two between participants counterbalancing conditions. Recruitment was through Amazon Mechanical Turk, restricting geographical location to North America. Participants required approximately 20 minutes to complete the task and they were compensated \$1 for their time.

Both conditions consisted of three scenarios and each scenario described a hospital ward in a fictional town, specializing in a particular type of ailment. For example participants were told of a cancer ward, treating only patients of three types, those with lung cancer, stomach cancer, or throat cancer. For each scenario, participants were given some information creating a common cause between ailment pairs. For example, in one case participants were told that throat and lung cancers are caused by smoking, but throat and stomach cancers by alcoholism. Some rationale was provided to justify each association between ailments. All relations were constructed to look semi-plausible (the authors independently assessed this), but we did not aim for medical accuracy. The between participants condition implemented a counterbalancing manipulation, that presented the same scenarios but varied the common causes.

Participants went through each scenario in a blocked format presentation, so that, for example, no information about a subsequent scenario would be presented prior to finishing all questions relevant to the current scenario (scenario order was randomized). The block for each scenario had analogous format. Participants were first presented with the information about the hospital ward, the ailments treated there, and the causal relations. Subsequently participants went through four or five multiple choice questions testing knowledge of the causal relations. The questions were meant to be straightforward and answerable on the basis of simple understanding of the presented information. Participants received corrective feedback, specifically if they responded incorrectly they were told so and asked to try again until they answered correctly (there were more than two alternatives for each question).

Once the training part was over, participants were told that they would be asked to make judgments about the proportion of various categories of patients at the fictional hospital. With each question, the text describing the hospital ward and the causal dependencies was included so that participants did not have to memorize anything, just understand the information provided. Each of the questions was prompted with the statement that each patient was brought to the hospital ward for only a single type of ailment (e.g., a single cancer type or a single fracture, depending on the scenario). Then, participants were asked to indicate on a 0 (None of them), to 100 (All of them) slider the proportion of patients likely to be admitted for ailment A in some questions, A or B in other questions, and A or B or C in another question; note, each combination of possibilities was shown only once. The triple disjunction was implemented as a catch question, since the total number of patients was fixed at 100. An additional three catch questions were included, where participants were just told to select a particular response, as a check that they were paying attention.

After participants finished responding to the questions for the three scenarios they were given a version of the the Cognitive Reflection Test (Frederick, 2005). The CRT is designed to discriminate between participants adopting either a more intuitive or a more deliberative thinking style (Toplak et al., 2011). The CRT has previously been shown to correlate well with measures of non-normativity in probabilistic judgments such as conjunction fallacies (Yearsley et al, 2016).

Aside on the physical analogue of this experiment It may be useful for those familiar with interference in physics to outline the analogy between this phenomena and 'interference' effects in human probabilistic decision making. This might help to motivate some of the experimental design, but this subsection can be safely skipped by any reader who wishes.

In the classic two slit interference experiment a particle can arrive at a given point via one of two paths. In quantum theory because of the wave-like nature of particles two things happen: 1) The particle in some sense (which we don't intend to make precise here) takes 'both' paths, and 2) The phase of the particle's wave-function depends on the details of the paths taken. By choosing the paths in a particular way we can cause the two paths to interfere in either a constructive way (so that the total probability of arriving at a point is greater than the sum of the probabilities to follow either path) or a destructive way (so that the total probability of arriving at a point is less than the sum of the probabilities to follow either path.)

The analogy in decision making is that a disjunction $A \cup B$ of two disjoint events can happen in one of two ways. By manipulating the information we give about the events, for example by introducing a possible common cause, we can, empirically, cause the disjunction to be judged more likely than the sum of the probabilities for the individual events. This is the analogue of constructive interference in the physical set up. This helps us understand why the key experimental manipulation is essentially the stories we tell about the relationship between the different ailments. A pictorial representation is given in Fig 1.

Of course, this analogy is not meant to constitute a formal theory of quantum cognition, but such a theory can be formulated (Busemeyer & Bruza, 2011; Yearsley & Busemeyer, 2016).

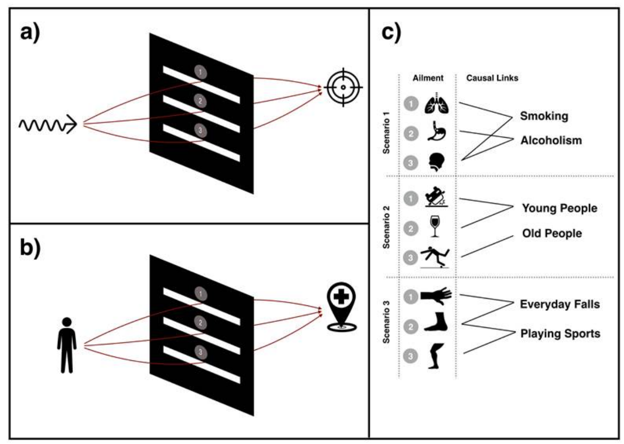

Figure 1: Sketch of the analogy between the physical experiment and the decision making one. a) In a physical interference experiment, a quantum state can take one of three possible paths to a detector, and the different alternatives interfere. b) In our experiment, a patient ends up in a hospital ward due to one of three ailments. c) A pictorial representation of the different ailments for Condition 1. In Scenario 1 the three ailments were Lung, Throat and Stomach cancers. Lung and Throat cancers were linked (smoking), and Throat and Stomach cancers were linked (alcoholism). In Scenario 2 the three ailments were auto accidents, alcohol poisoning, and falls. Auto accidents and alcohol poisoning were linked (young people) and falls (old people) was a singleton. For Scenario 3 the ailments were fractures to wrists, ankles, or lower legs. Wrist and ankle fractures were given a common cause (everyday falls) and angle and lower leg fractures likewise (playing sports).

Results

There are two critical tests to perform; firstly we must check that the two way interference terms $\mathscr{I}_2(A,B)$ vary depending on the events in each scenario and condition. Second, we must examine three way interference $\mathscr{I}_3(A,B,C)$ for each scenario and condition. In this contribution, we only report Bayesian statistical tests that were performed using JASP (JASP team, 2016). In particular we report Bayes factors for the alternative versus the null hypothesis, so that values > 1

indicate evidence for the alternative hypothesis.

Recall that the two way interference term $\mathscr{I}_2(A,B) = p(A\cup B) - p(A) - p(B)$. To check the behaviour of these terms we perform a Bayesian RM ANOVA, with scenario and event pair as the repeated measures, and the counterbalancing condition as a between subjects factor. We omit the CRT from this analysis to save space, but there are no interesting effects of CRT. The analysis of effects is shown in Table 1.

Table 1: Analysis of effects for Bayesian Repeated Measures ANOVA of two way interference terms

Effect	$p(incl)$	$p(incl\|data)$	$BF_{Inclusion}$
Scenario	0.737	1.000	2.44×10^4
Condition	0.737	0.999	2.49×10^2
Pair	0.737	1.000	"∞"
Scenario*Condition	0.316	0.998	1.37×10^3
Scenario*Pair	0.316	1.000	2.11×10^4
Condition*Pair	0.316	0.998	1.35×10^3
Scenario*Condition*Pair	0.053	0.998	1.12×10^4

Recall that if the best description of this situation is via a linear model, ie if non-normative effects are either absent, or due only to response error, then we expect to see no effect of scenario, condition or pair. In contrast JASP actually returns a BF of inclusion for the pair variable of ∞, indicating an extremely large effect of pair. The other variables, and all the interaction terms, all have large Bayes factors. The large Bayes factors for the interaction terms are unsurprising given the experimental design - the difference between the same scenario in a different condition and a different scenario in the same condition is really a matter of convention.

This analysis shows that a model beyond a linear one is needed to explain these data. Now we perform the analogous test for $\mathscr{I}_3(A,B,C)$ to check if a bilinear model is sufficient.

Recall the three way interference term for each scenario is computed as $\mathscr{I}_3(A,B,C) = p(A\cup B\cup C) - p(A\cup B) - p(B\cup C) - p(A\cup C) + p(A) + p(B) + p(C)$. To check the behaviour of these terms we perform a Bayesian RM ANOVA, with scenario as the repeated measure, and the counterbalancing condition and CRT as between subjects factors. The analysis of effects is shown in Table 2.

Table 2: Analysis of effects for Bayesian Repeated Measures ANOVA of three way interference terms

Effect	$p(incl)$	$p(incl\|data)$	$BF_{Inclusion}$
Scenario	0.737	0.313	0.163
Condition	0.737	0.407	0.245
CRT	0.737	0.141	0.059
Scenario*Condition	0.316	0.299	0.923
Scenario*CRT	0.316	2.92×10^{-4}	6.33×10^{-4}
Condition*CRT	0.316	0.003	0.006
Scenario*Condition*CRT	0.053	2.2×10^{-7}	3.97×10^{-6}

The results are striking; none of the Bayes factors for inclusion are greater than 1, indicating that no model containing any combination of these effects is preferred over a null model. The conclusion then is that the terms $\mathscr{I}_3(A,B,C)$ are constant - they do not vary when we manipulate common causes implied for the events in the way that the terms $\mathscr{I}_2(A,B)$ do. This implies that a bilinear model is sufficient to explain these effects.

The lack of a significant effect of the CRT is actually reassuring; recall that the terms $\mathscr{I}_3(A,B,C)$ should be constant regardless of whether a decision maker is using a linear (classical) or bilinear (quantum) model. The CRT has previously been associated with the strength of various measures of non-normativity (Yearsley et al, 2015) and the fact that it is not predictive here suggests that these effects behave very differently from other measures such as the size of conjunction fallacies.

Conclusions and Future Directions

The empirical finding of so-called probabilistic fallacies in decision making has led to an intense debate over how much (if any) of human cognition should be understood in terms of the principles of classical, Bayesian, probability theory (Tversky & Kahneman, 1983). Those who believe these findings cast doubt on the applicability of classical probability theory have tended to respond by abandoning all together the idea of a formal probabilistic framework for decision making. Recent advances in applying quantum probability theory to modelling human decision making (Pothos & Busemeyer, 2013) raise the possibility that all (or most) of human cognition can be understood in formal probabilistic terms, but the appropriate approach is not classical probability theory but quantum probability.

However at least one objection to using quantum probability theory (there are many) is that it is unclear how exactly this expands the space of possible models. Most accounts of the relationship between quantum and classical models tend to focus on the issue of incompatibility, but this is notoriously hard to make precise. In addition, it is far from clear that quantum probability theory is the only way to generalise classical probability to include incompatible events.

What we have tried to do in this paper is to show how we may take a particular approach to divide up the space of possible computational level accounts of interference effects in decision making. The space of models is split into different levels, of increasing complexity in the sense of higher level interference effects. Classical probability theory, and its generalisations in the form of linear models, occupy the lowest level of this hierarchy, whereas bilinear models such as quantum theory sit at the next level up in complexity. Above these bilinear models are an infinite number of different levels, although we can argue on general grounds that we expect human behavior to be characterised by a relatively low level model.

Each level in the hierarchy has a unique experimental signature, and we used this to show that behavior in a particular probabilistic reasoning task is consistent with a bilinear theory. Of course, whether all current examples of non-normative probabilistic reasoning are likewise consistent with a bilinear model is an open question. This level consists of theories where interference between alternatives is computed pairwise. Quantum theory is situated at this level, however our approach is not able to distinguish between different models in a given level. Further work, for example looking at constraints obeyed by quantum theory but not other non-classical probability theories, could address this.

We finally want to outline some future directions for research. One important question is how well our findings generalise when we consider different kinds of relationships between events. In our scenarios different ailments were related, if at all, by common causes, for example smoking can cause both lung and throat cancer. This means that the associated interference terms $\mathscr{I}_2(A,B) = p(A \cup B) - p(A) - p(B)$, tend to be positive. It would be useful to show that we can generate negative interference terms by implying the appropriate relationships between conditions, and check that our conclusions still hold. Another future direction would be to extend this approach to other areas of cognition where quantum models have been proposed, for example perceptual decision making.

Another future direction is to try and extend this analysis to models which fall just outside our framework. One such example is the classical probability plus noise model due to Costello and Watts (2014). They propose a general linear model but with an error term which, rather than be a constant, depends on the type of event being considered, eg a single event, a conjunction or disjunction etc. The idea is that more complex events are associated with larger error terms, and they showed this can lead to conjunction fallacies in participants responses, even though the underlying belief states obey classical probability theory. This theory has a slightly different experimental signature from linear or bilinear models, but it can still be tested against them in a similar way. The results of this analysis will be presented elsewhere.

Acknowledgments

JMY was supported by NSF grant SES-1556415. EMP was supported by Leverhulme Trust grant RPG-2015-311 and IB by H2020-MSCA-IF-2015 grant 696331.

References

Busemeyer, JR & Bruza, P (2011). Quantum models of cognition and decision. CUP: Cambridge, UK.

Costello, F. & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. Psychological Review, 121(3):463?480.

Frederick, S (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives.* 19, 25-42.

Gigerenzer, G, Hertwig, R & Pachur, T (eds.) (2015). Heuristics: The foundations of adaptive behavior. (OUP).

JASP Team. (2016). Jasp. Retrieved from https://jasp-stats.org

Jones, M. & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. Behavioral and Brain Sciences, 34, 169, 231.

Kahneman, D, Slovic, P & Tversky, A. (eds) (1982). Judgments under uncertainty: Heuristics and biases. (CUP, UK)

Kolmogorov, AN (1933/1950). Foundations of the theory of probability. N.Y. Chelsea Publishing Co.

Oaksford, M. & Chater, N. (2009). Prcis of Bayesian rationality: the probabilistic approach to human reasoning. Behavioral and Brain Sciences, 32, 69-120.

Pearl, J (1988). Probabilistic reasoning in intelligent systems: Networks of plausible inference. Morgan Kaufmann.

Pothos, E. M. & Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modelling? Behavioral & Brain Sciences, 36, 255-327.

Sorkin, RD (1994). Quantum mechanics as quantum measure theory. Mod. Phys. Lett. A, 09, 3119.

Toplak, ME, West, RF, Stanovich, KE (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*. 39, 1275-1289.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjuctive fallacy in probability judgment. Psychological Review, 90, 293-315.

Yearsley, JM & Busemeyer, JR. (2016). Quantum cognition and decision theories: a tutorial. Journal of Mathematical Psychology. 74, 99-116.

Yearsley, JM, Trueblood, JS & Pothos, EM (2016). When are representations of causal events quantum versus classical? In Papafragou, A., Grodner, D., Mirman, D., & Trueswell, J.C. (Eds.) (2016). Proceedings of the 38th Annual Conference of the Cognitive Science Society. (pp. 2447-2452). Austin, TX: Cognitive Science Society.