Looking at "Linda": Is the Conjunction Fallacy Really a Fallacy? *Word count: 2691*

Alina Berendsen, Simon Jonas Hadlich, Joost van Amersfoort

I. The Linda problem

"Judgments are all based on data of limited validity, which are processed according to heuristic rules." (Tversky & Kahneman, 1974, p.1124)

Since Amos Tversky and Daniel Kahneman first encounter in the Hebrew University in 1968, they achieved high reputation among their fellows with their research in the field of human reasoning, decision-making and behavior. The above-mentioned quote is an excerpt from their most famous contribution in the field of probability judgment. It mainly states that generally humans tend to argue irrational, biased by personal, societal, or cultural reasoning.

In 1974, Tversky and Kahneman published a paper about judgement and uncertainty, which includes the "Linda problem". Meanwhile, this example reached an ample amount of fame and is cited frequently. The Linda problem is aimed at exposing the so-called conjunction fallacy and is presented as follows to the the test persons:

> "Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with the issue of discrimination and social justice, and also participated in antinuclear demonstrations.
>
> Which of the following is more probable:
>
> 1. Linda is a bank teller (T)
> 2. Linda is a bank teller and is active in the feminist movement (T and B)" (cited in Cartwright, 2008, p.155)

Tversky and Kahneman (1974) surprisingly found that a vast majority of their subjects chose the conjunction T and B to be more likely than one of its components, T. Surprising, because from a logical perspective, it is clear that a conjunction can never be more likely than any of its constituents. If A /\ B is true, than A and B individually are necessarily true as well. For instance, when rolling two dices with the aim of getting a four and six versus rolling just one die with the aim of getting a six, it becomes obvious that it is more probable to get a six with one die than two specific numbers with two dice (Cartwright, 2008, p.155). Having found that their subjects frequently failed to make use of the conjunction rule when evaluating similar questions, Tversky and Kahneman (1974) argued that human are prone to commit the conjunction fallacy because of their use of heuristics.

In opposition to the findings of Tversky and Kahneman, Ralph Hertwig and Gerd Gigerenzer (1999) argue that the outcome of the "Linda problem" trial can be revised by a replacement of the

questions with a more frequent situation. Their findings have shown that a modification can reduce the frequency of fallacy from over 80 percent to approximately ten to 20 percent.

Therefore, this paper will contrast the approach of Tversky and Kahneman with that of Hertwig and Gigerenzer. For this purpose, we will explain both argumentations and elaborate on the findings of the authors. In a final step, we will determine which position seems more convincing from our point of view.

II. Representativeness

Psychologists Amos Tversky and Daniel Kahneman (1974, 1983) have argued that when making decisions under uncertainty, humans use a limited set of simple judgmental heuristics. Among these strategies are representativeness, availability, and anchoring. In general, these heuristics are useful as they reduce the complex tasks of assessing probabilities or predicting values to simpler judgmental operations. However, as the researchers have shown, they also lead to "severe and systematic errors" (1974).

The representativeness heuristic is a means of assessing the probability of an uncertain event or the value of a quantity by comparing it to a mental model (Tversky & Kahneman, 1983). It is used to assess probabilities of categorization or causation, as Tversky and Kahneman (1974) write: "What is the probability that object A belongs to class B? What is the probability that event A originates from process B? What is the probability that process B will generate event A?" These questions are answered by evaluating the degree to which A is representative of B, which is to say, the degree to which A resembles B. The concept of representativeness suggests that an event B is likely to be caused by event A if both resemble each other; or that individual B belongs to category if the individual is highly representative of the category's features.

The representativeness heuristic has been researched through experiments in which participants were asked to assess the probability that a person held a range of jobs - such as farmer, banker, or physician. For example, consider "Steve", who is described by Tversky and Kahneman (1974): "Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail." One finding from these experiments is that "people order the occupations by probability and by similarity in exactly the same way". As the authors argue, this indicates that the probability of Steve being, for instance, a librarian is "assessed by the degree to which he is representative of, or similar to, the stereotype of a librarian".

While judgments according to representativeness may often be valid - there is a good chance that Steve indeed is a librarian - they are frequently flawed due to a range of biases. As Tversky and Kahneman (1974) show, judgments by the representativeness heuristic are insensitive to the base-rate frequency of the outcomes (for instance, even if a person's description is highly representative of that of an astronaut, the chance of her actually holding this occupation is still low, due to the minuscule number of astronauts there are), sample size, and predictability, as well as subject to misconceptions

about regression and chance (one effect of which is the well-known "gambler's fallacy", which wrongly assumes that after a long streak of red in a game of roulette, black becomes more likely).

In particular, Tversky and Kahneman (1983) have argued that "[a] conjunction can be more representative than one of its constituents", which is to say that intuitive reasoning with the representative heuristic is causal for the frequency with which humans fall for the conjunction error. The researchers posit that as representativeness is nonexistensional - i.e. not determined by frequency, and not bound by class inclusion - "the test of the conjunction rule in probability judgments offers the sharpest contrast between the extensional logic of probability theory and the psychological principles of representativeness." It is this hypothesis that is tested with the "Linda problem".

In order to find out whether people use representativeness heuristics rather than existensional probability reasoning, Tversky and Kahneman (1983) presented their subjects with three distinct tests of the "Linda problem" variety - indirect, direct-subtle, and direct-transparent tests, which work on different hypotheses. In the indirect test, one group assesses the probability of the conjunction (T&F), whereas the other group assesses its constituents (T, F), thus testing whether probability judgments conform to the conjunction rule. In the direct-subtle procedure, the conjunction is compared to its less representative constituent (T&F, T), testing whether "people will take advantage of an opportunity to compare the critical events". Finally, in the direct-transparent test, subjects are asked to evaluate to probabilites of the conjunction and its constituents (T&F, T, F), and the relation between them is highlighted, which tests whether "people will obey the conjunction rule" when they are lead to compare the critical events. These three variants were initially tested on three groups, naive, informed (i.e. familiar with basic probability theory), and sophisticated (doctoral students in a decision science program) subjects.

To the researchers' surprise, all three groups invariably tended to commit the conjunction error - i.e, they ranked the probability of Linda being a feminist bank teller higher than that of her being a bank teller. Using the indirect and the direct-subtle test, around 85% of the participants erred in their judgment; and in direct-transparent tests, only the sophisticated group significantly reduced their error rate. Tversky and Kahneman subsequently attempted to find alternative explanations, such as a misunderstanding of the test question; but even as the subjects were presented with two alternative lines of reasoning - one arguing with resemblance, one with probability, most subjects (65%, n = 58) chose the invalid representativeness argument. Indeed, they found that "naive subjects generally endorse the conjunction rule in the abstract", but fail to appreciate it as decisive. As the researchers write, their application of this rule to the Linda problem is blocked by the compelling impression that T&F is more representative of her than T is." This phenomenon is pervasive, as the researchers show in a series of studies on physicians and others with probabilistic expertise.

As a conclusion from their experiments, Tversky and Kahneman (1983) argue that the majority of people generally accept the conjunction rule in its abstract form, but fail to consequently apply it to concrete cases, such as the "Linda problem". In line with their earlier research, they understand the conjunction error as "only a symptom of a more general phenomenon", in which people have a tendency to overestimate the probabilities of representative events, and underestimate those of less representative ones. Consequently, they argue that the incoherence of human judgments "is more

than skin deep", throwing out a "truth plus error" model. As humans make use of heuristics, "normatively inspired theories that assume coherence are descriptively inadequate": their judgments are ridden by fallacies anchored in our evolutionary history.

III. Non-monotonic pragmatic analysis

Considering the conjunction fallacy from a pragmatic point of view, Hertwig and Gigerenzer (1999) make an interesting prediction: “if asked for probability judgments, people will infer a nonmathematical meaning of `probability', and the proportion of conjunction violations will be high as a result.” After the researchers asked subjects to paraphrase the term possibility, a significant part of the group favoured a non-mathematical meaning. This result was to be anticipated, as it had been reported in earlier studies. However, Hertwig and Gigerenzer continued with four more studies. In their second study, they investigated how one can preserve the relevance maxim. This is one of four "conversational maxims" proposed by philosopher Paul Grice (1975, originally as "maxim of relation"), based on the pragmatics of natural language. It claims that when the relevance maxim applies to a problem “the audience expects the communicator's contribution to be relevant” (Hertwig and Gigerenzer, 1999). To investigate this claim, Hertwig and Gigerenzer set up an experiment with four groups. The first two groups had to "think aloud"; one group had to give a typicality judgment and then a probability judgment, the other group only had to make a probability judgment and was considered the control group. The second pair of groups did not have to think aloud and again one group did a typicality judgment and then a probability, while the other control group only did a probability judgment. The interesting result is that violations of the conjunction rule decreased if the group was first asked to give a typicality judgment. Another aspect is that when thinking aloud, the amount of mistakes was also lower. This was especially apparent in the group that thought aloud and gave a typicality judgment first. The researchers concluded from these observations that forcing people to think about a typicality judgment lowers the impact of the relevance maxim and people are more critical towards the communicator.

Hertwig and Gigerenzer took their investigations one step further and asked participants to paraphrase the word frequency and view the Linda problem from a frequency point of view. This experiment was to investigate the validity of a second prediction, which goes as follows: “If asked for frequency judgments, people will infer mathematical meanings, and the proportion of conjunction violations will decrease as a result.” (Hertwig and Gigerenzer, 1999) To test this they asked participants to estimate the frequency “of the two constituent hypotheses and the conjoint hypothesis”. After this task, they were again asked to paraphrase frequency. This time the participants never committed a conjunction fallacy and almost always paraphrased frequency with mathematical terms. Hertwig and Gigerenzer then investigated a third prediction: “If the term ‘probability’ is replaced by ‘believability’, then the proportion of conjunction violations should be about as prevalent as in the probability judgment.” They tested this prediction using external data, where participants where asked to rank a number of statements to their probability, while another group ranked the same statements

according to believability. The ranked list of statements showed large resemblance between the groups, therefore it could be concluded that prediction holds.

In conclusion, Hertwig and Gigerenzer (1999) found that, when presented with a problem of the "Linda" type, most people will violate the conjunction rule. If they are, however, asked to take a mathematical perspective on the question, violations are drastically reduced. This is indeed coherent with Tversky and Kahneman's observation that when asked to bet on an outcome, people will adhere to the conjunction rule more often (1983).

Hertwig and Gigerenzer (1999) posit that violation of the conjunction rule is not irrational. They argue that the Linda problem is an example of "a situation in which adhering to social norms, here conversational maxims, is rational, although it conflicts with classical rationality", which is defined as "adherence to the laws of probability theory and logic". The researchers claim that "[t]he conjunction rule is [not] the only rule of internal consistency used as a benchmark of sound reasoning" - instead, social rationality should be taken into account.

IV. Is the conjunction error a true fallacy?

The conclusions of Tversky and Kahneman (1974,1983) and Hertwig and Gigerenzer (1999) are drawn from similar observations: Humans generally tend to violate the conjunction rule in practical applications, even though they accept it in its abstract form. Both approaches, too, represent a view of human rationality as "bounded", that is subject to certain constraints (Cartwright, 2008). However, they represent two perspectives on the question whether is violation is, in fact, a fallacy. The controversy is nicely described in a quote by Pelletier and Elio (1993) about non-monotonic reasoning:

> "Which arguments should a logical theory admit as valid? In the case of nonmonotonic logic the criteria are not the same as in the case of classical logic [...]. In the case of default reasoning the whole notion of what is correct (the notion we are trying to model) is defined in terms of what sort of reasoning people actually engage in and what sorts of of reasoning an intelligent agent will have to do to get along in a common sense way. So “mistakes” by common sense reasoners are only possible in the sense that they are different from the majority [...]."

Tversky and Kahneman (1983) perceive the judgmental errors caused by use of the representativeness heuristic as fallacious. They find that the Linda problem unearths an "incoherence" in humans' judgments, which are caused by "the contrast between the extensional logic that underlies most formal conceptions of probability and the natural assessments that govern many judgments and beliefs." Gigerenzer et al. (1999), on the other hand, argue that "the function of heuristics is not to be coherent". From the perspective of social rationality, incoherent judgments can be rational if they are - on average - beneficial in the individual's social context.

Is the conjunction error a fallacy? From the perspective of logic, we can only conclude that violating a basic rule such as the conjunction rule is fallacious. However, Gigerenzer et al.'s (1999) perspective of social rationality is well suited to make us aware that human rationality is bounded by physical constraints as well as frequent lack of information. The human mind has evolved to make use of heuristics to break down complex probability judgments into much simpler judgmental processes. Clearly, this sometimes leads our judgment astray - e.g. because humans seem to have more problems dealing with fractions than with absolute numbers. Nevertheless, the question whether the violation of the conjunction rule is a fallacy has to be answered not on the basis of a single case, but in the context of human social rationality. There, we come to the conclusion that the "Linda problem" does not unearth a fallacy, but a reasoning pattern that has evolved as a means of accomodating the physical and environmental boundaries set to the human brain.

V. References

Cartwright, J. (2008). Evolution and Human Behaviour, 2nd ed. Basingstoke: Palgrave Macmillan.

Hertwig, R., & Gigerenzer, G. (1999). The `Conjunction Fallacy' Revisited: How Intelligent Inferences Look Like Reasoning Errors. *Journal of Behavioural Decision Making, 12*, 275-305.

Gigerenzer, G., Todd, P. M., & ABC Research Group (1999). *Simple Heuristics that Make Us Smart.* New York: Oxford University Press.

Grice, H. P. (1975). Logic and Conversation. In Cole & Morgan (eds.), *Speech acts*, pp. 41-58

Myers, D. G. (2010). *Psychology, 9th ed.* New York: Worth.

Pelletier, F., & Elio, R. (1993). Some Truths about Non Monotonic Reasoning. Technical Report TR 93-12, Department of Computing Science, University of Alberta

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science,*

185, 1124-1131.

Tversky, A., & Kahneman, D. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293–315.