

## A Few Comments on the Linda Problem<sup>1</sup>

ADAM OLSZEWSKI

**ABSTRACT:** This paper discusses an experiment in cognitive psychology called the Linda problem. Firstly, some natural conditions for the correctness of the interpretation of psychological experiments (such as the Linda problem) are formulated. The article is essentially a critique of the interpretation of the results of the Linda problem experiment provided by Kahneman and Tversky as well as – indirectly – their concept of heuristics. It is shown that the interpretation provided by Kahneman and Tversky does not meet the aforementioned conditions for correctness. The main argument is justified utilizing such rules of rationality as conditional probability and Grice’s conversational maxims. It is also pointed out that this argument can be reformulated in terms of the intuitive system of reasoning.

**KEYWORDS:** Conjunction fallacy – cognitive psychology – conditional probability – Linda problem.

In this paper, I present a new interpretation of the “Linda problem”. The “Linda problem”<sup>2</sup> is the name of a psychological experiment performed by psychologists Daniel Kahneman and Amos Tversky (‘KT’ for short) in the

---

<sup>1</sup> Received: 7 October 2016 / Accepted: 20 March 2017

✉ Adam Olszewski  
Department of Philosophy of Logic, Faculty of Arts  
The Pontifical University of John Paul II in Cracow  
Ul. Kanonicza 25, 31-002 Cracow, Poland  
e-mail: atolszad@cyf-kr.edu.pl

<sup>2</sup> In this paper, I will also use terms such as the “Linda task,” often without quotation marks.

1980s. It is described in Kahneman & Tversky (1983); and after a subsequent research and analysis, the summary of the experiment has been presented in Kahneman (2013). The experiment consisted of the description of an imaginary woman named Linda. Her story is as follows:

Linda is thirty-one 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. (Kahneman 2013, 211 ff.)

In one version of the study, the researchers then presented eight possible scenarios for Linda's future. They were:

- Linda is a teacher in primary school;
- Linda works in a bookstore and practices yoga;
- Linda supports the feminist movement (abbreviation: (F));
- Linda is a social worker and helps people with mental disorders;
- Linda is a member of the Women's Electoral League;
- Linda is a bank teller (abbreviation: (T));
- Linda is an insurance agent;
- Linda is a bank teller and is active in the feminist movement (abbreviation:  $(T \wedge F)$ ).

The participants of the study were asked which of the following two propositions is more probable. "Linda is a bank teller" (T) or "Linda is a bank teller and is active in the feminist movement" ( $T \wedge F$ ).

The experiment was repeated many times with different groups and in different versions, as well as with groups whose participants had previously completed courses related to statistics or probability theory. To the surprise of the researchers, most the subjects stated that the conjunction ( $T \wedge F$ ) is more probable than the proposition (T). This violates the axioms of probability theory, which entail that the probability of the conjunction of two propositions is less than or equal to the probability of the individual members of the conjunction  $p(T \wedge F) \leq p(T), p(F)$ . In addition, the subjects were unable to explain why they ascribed probabilities to both propositions the way they did.

Relevant here is a well-known quote from Stephen Jay Gould, who, knowing the right answer, wrote that, "[A] little homunculus in my head

continues to jump up and down, shouting at me – ‘But she can’t just be a bank teller; read the description!’” (Kahneman 2013, 215).<sup>3</sup> Kahneman describes the tremendous impression the results made on him (cf. Kahneman 2013, 213).

The experiment was repeated on a group of PhD candidates from Stanford Graduate School of Business, who had previously participated in probability theory, statistics, and decision theory classes. Ninety-five percent of respondents judged against the laws of probability, and thus, according to the researchers, against the rules of rationality. Moreover, the researchers also carried out tests on a group larger than the initial one and on groups where the participants were considered better (or worse) educated, which substantiated the results.<sup>4</sup>

Kahneman mentions only one group in which 64% of the participants gave the correct answer.<sup>5</sup> They were PhD candidates in sociology from the universities of Stanford and Berkeley.

In their research, Tversky and Kahneman considered, for instance, the transformation of the proposition (T) into the following, “Linda is a bank teller, regardless of her activity in the feminist movement.”<sup>6</sup> After conducting an experiment with this new proposition, the number of incorrect answers was reduced to 57%. On the other hand, when conducting the experiments described in Kahneman & Tversky (1983), KT knew that the respondents might treat (T) as  $(T \wedge \neg F)$  – see Kahneman & Tversky (1983, 299).<sup>7</sup> At the same time – based on the description – KT knew that the

<sup>3</sup> I prefer to quote this paper here because it is more comprehensive and decisive.

<sup>4</sup> KT mentioned studying approximately 3,000 subjects; cf. Kahneman & Tversky (1983, 309-310).

<sup>5</sup> It is assumed by KT that an answer is correct when it is consistent with Probability theory.

<sup>6</sup> “Linda is a bank teller, whether she is active in the feminist movement or not” (Kahneman & Tversky 1983, 299). KT therefore took into consideration different formulations of (T). I also mention these versions of the Linda experiment, as well as other interpretations made by KT, to emphasize the versatility of their analysis and the fact that the interpretation I will present had already been taken into consideration in a certain form.

<sup>7</sup> Their reason for rejecting this interpretation is surprising. “To test this interpretation, we asked a new group of subjects (N = 119) to assess the probability of T and of

events ( $T$ ) and  $(T \wedge F)$  are not independent, so there was no reason for them to assume that the subjects would treat  $(T)$  as  $(T \wedge \neg F)$ , since this would imply that  $(T)$  and  $(T \wedge F)$  are independent, which would be at odds with logic. In this context, another peculiar point is that KT, to exclude an understanding of  $(T)$  as  $(T \wedge \neg F)$ , did not examine the relationship between  $(F)$  and  $(T \wedge F)$ .

What is more, KT knew from previous research that, if Linda's description is limited to the judgment that she is a 31-year-old woman, the results are in accordance with the laws of probability (cf. Kahneman & Tversky 1983, 305). Their conclusion is that  $C$  ( $C$  being the description of Linda), or the *paradigm*, as they call it, has affected the outcome of the experiment. It is my contention that we are dealing here with ordinary conditional probability, where the condition is  $C$ .

Let us ask why this "unexpected" outcome of the experiment had such a strong influence on the authors of the experiment as well as on the scientific community. Since there are, in fact, two questions, I assume that the answer to the first question arises from the answer to the second. Indeed, one of the main arguments of KT is the argument concerning dual process theory.<sup>8</sup> This theory, also known as the dual process account of reasoning, states that there are two systems [minds] in the human brain. System<sub>1</sub> is intuitive, and System<sub>2</sub> is analytical.<sup>9</sup> I believe that KT acted, as I will attempt to prove later on, within a kind of "paradigm" of this distinction.<sup>10</sup>

---

$T$  &  $F$  on a point scale ranging from 1 (extremely unlikely) to 9 (extremely likely)" (Kahneman & Tversky 1983, 299). KT thought that it seems reasonable to assume that the respondents did not interpret  $T$  as  $T \wedge \neg F$ , since there is nothing wrong to estimate the probabilities of two events even though one of them is a part of the other.

"The pattern of responses obtained with the new version was the same as before. The mean ratings of probability were 3.5 for  $T$  and 5.6 for  $T \wedge F$ , and 82% of subjects assigned a higher rating to  $T \wedge F$  than they did of  $T$ " (Kahneman & Tversky 1983, 299).

<sup>8</sup> At the time, the view that mind has two systems emerged at least in 1975. In Kahneman & Tversky (1983), this thesis took the form of division into extensional and intuitive reasoning (stated as early as in the title of the article).

<sup>9</sup> A modern summary of the research and the scientific hypothesis concerning this interesting distinction can be found in Hertwig & Gigerenzer (1999).

<sup>10</sup> It cannot, however, be claimed that KT were explicitly formulating the dual process theory.

Kahneman expressed this, for example, by commenting on Gould's statement, stating "[T]he homunculus in Gould's head was, *of course*, (italics mine – A. O.) insistent System<sub>1</sub>" (Kahneman 2013, 213).

Let us now discuss KT's interpretation of the result of the Linda experiment. I use the word "interpretation" with considerable caution, keeping in mind that it is unclear whether psychologists have any unified and standardized theory concerning the interpretation of the results of such experiments as well as the interpretation of *tasks* given to their subjects. In the case of multiple psychological experiments, we are faced with more than one interpretation, two interpretations being dominant.

Experiments similar to the one in question often feature two intersubjective factors (texts) and four subjective factors (interpretations). The first objective factor is a text ( $t_1$ ), created by the researchers, consisting of the description of the task to be solved by the participants. The second factor is usually also a text ( $t_2$ ), the aim of which is to present the notions of the participants. The interpretations concern the aforementioned intersubjective texts, and represent the subjects' ways of understanding the texts. These are very often subjective with respect to both, researchers and participants. By using ( $i_1$ ), we can denote the interpretation of text ( $t_1$ ) offered by the researchers who performed the experiment. Similarly, by using symbol ( $i_4$ ) we can denote the interpretation of the same text ( $t_1$ ) offered by the subjects of the experiment. Both interpretations should be equivalent, as this is usually the aim of the researchers concerning text ( $t_1$ ). The second aspect is the interpretation ( $i_2$ ) of the experiment's result which consists of the interpretation of text ( $t_2$ ) provided by the researchers who created the experiment and the interpretation ( $i_3$ ) of the experiment's result provided by other researchers interested in the experiment. The interpretations ( $i_2$ ) and ( $i_3$ ) should also be equivalent. Therefore, there are two equivalences necessary for the methodological soundness of both the KT experiment as well as any similar experiments. This is, however, insufficient, as no one falsified the KT statements: <sup>11</sup>

$$(i_1) \equiv (i_4)$$

---

<sup>11</sup> I argue that the interpretations of the Linda problem do not fulfill both of these equivalences. The equivalences are taken here to be stronger than material equivalence, rather as 'to have similar meaning'.

$(i_2) \equiv (i_3)$ .

In the case of the KT experiment, neither equivalence holds, as will be explained later in this paper.

Interpretations  $(i_1)$  and  $(i_2)$  presented by KT were immediately criticized by other researchers in multiple ways. One critical remark was made by German psychologist Gerd Gigerenzer. It referred to the understanding of the term ‘probability’ by the subjects. When the experiment was repeated, with the question regarding the probability of the proposition replaced with the *frequency* of occurrence (the frequency interpretation), a great majority of the subjects answered correctly.

Some researchers have suggested (e.g., Morier & Borgida 1984, Hilton 1995) that the proposition (T) could have been understood by subjects as  $(T \wedge \neg F)$ ; this case will be described further. Morier and Borgida studied such possible understandings of the proposition T and developed an experiment in which the subjects were asked about the propositions: (T), (F),  $(T \wedge F)$  and  $(T \wedge \neg F)$ . The results of the experiment confirmed the presence of *the conjunction fallacy*. However, after constructing a problem which was logically equivalent to the Linda problem, in which the meaning of the propositions was explained more clearly, to their surprise, the probability of error decreased dramatically (cf. Miyamoto, Gonzalez & Tu 1995, 337-338).<sup>12</sup> However, Hilton claims that the interpretation of (T) as  $(T \wedge \neg F)$  is based on rational heuristics (they differ, however, from those desired by KT). The reason for disregarding the rules of probability theory is given by the rules of rational pragmatics, and not by any kind of illogicality (cf. Stanovich 2010, 102).

Returning to the criticism regarding the interpretation of the probability used by KT, it is worth mentioning that there are at least five known interpretations of probability. The participants in the KT experiment were asked to assess the probability of the sentences, which shows that they were essentially asked to make use of Carnap’s logical probability. If this were the case, it would provide a basis for criticism of the KT experiment, since if the participants in the study had only basic knowledge of logical probability, this would open the experiment to allegations of circular reasoning.

---

<sup>12</sup> The authors reflect on conditional probability in the context of the Linda problem.

Let us now turn to the interpretation ( $i_4$ ) of the result of the Linda problem experiment which refers to the criticism of KT put forward by David Morier and Eugene Borgida (see Morier & Borgida 1984), and later by Denis Hilton (see Hilton 1995). They are based on the analytical System<sub>2</sub>.<sup>13</sup> Firstly, I will present the interpretation ( $i_2$ ) as the hypothetical reasoning of a possible participant in the experiment.

- I. Grice's Rules<sup>14</sup> (Maxims of Relevance and Quantity);<sup>15</sup>
- II. Description of Linda C is essential for the task; (from I. and the description)
- III. Based on C, the conditional probability<sup>16</sup> that Linda is a feminist  $p_c(F)$  is higher than that she is not, i.e.:  $p_c(\neg F) < p_c(F)$ .
- IV.  $p(T \wedge F) \leq p(K)$ ,  $p(F)$ ; (from the properties of probability theory; abbreviation: CP)
- V. The sentence K, in the context of Linda's description, may be understood<sup>17</sup> by subjects as an abbreviation for  $(T \wedge \neg F)$ , which justifies the following inference [VI. – XI.]:
- VI. (F) is true or (F) is false; (from the bivalence of classical logic)
- VII. (F) did not occur, therefore (F) is not true; (from Grice's maxims of relevance and quantity)
- VIII. (F) is false;
- IX.  $(\neg F)$  is true;
- X. We have (T);
- XI. Therefore:  $(T \wedge \neg F)$ ; (from classical logic)
- XII.  $p_c(T) = p(T)$ ; (from the independence of (T) and C)
- XIII.  $p_c(T \wedge F) = p((T \wedge F) \wedge C) / p(C)$ ; (from the definition of CP)
- XIV.  $p((T \wedge F) \wedge C) / p(C) \equiv p(T \wedge (F \wedge C)) / p(C)$ ; (from the laws of logic)

---

<sup>13</sup> I am using conditional probability here as an important element.

<sup>14</sup> It means that the subject accepts Grice's conversational rules (maxims).

<sup>15</sup> The Maxim of Quantity requires us to say only what is necessary at a given stage of conversation, and the Maxim of Relevance to say only what is integrally related to the topic of conversation.

<sup>16</sup> The condition of the entire C.

<sup>17</sup> Here, we also require properly designed empirical research.

- XV.  $p(T \wedge (F \wedge C)) / p(C) = (p(T) \cdot p(F \wedge C)) / p(C)$ ; (from VI and independence of (T) and (F))
- XVI.  $p(T) \cdot (p(F \wedge C) / p(C)) > p(T) \cdot (p(\neg F \wedge C) / p(C))$ ; (from III)
- XVII.  $(p(T) \cdot p(\neg F \wedge C)) / p(C) = p(T \wedge (\neg F \wedge C)) / p(C)$ ; (from CP)
- XVIII.  $p((T \wedge \neg F) \wedge C) / p(C) = p_C(T \wedge \neg F)$ ; (from CP);
- XIX. Therefore:  $p_C(T \wedge F) > p_C(T \wedge \neg F)$ ; (from VII, XI, XIV and CP).

As we see in the sequence of this reasoning, premises III, V, and VI are critical to the above structure.

For the sake of clarity, I will characterize the steps of this reasoning also in an informal way. I assume that Grice's maxims can be applied, particularly the maxims of relevance and quantity. I think that description C is important. The probability that Linda is a feminist given description C is higher than the probability that she is not a feminist under the same condition, although there are more non-feminists than feminists. The question whether Linda is a feminist is crucial for our discussion. If it is said that Linda is a bank-teller, but feminism is not mentioned, then it should be assumed that Linda is a bank-teller and not a feminist, which is expressed by step V of the reasoning. Next, using the formal properties of the conditional probability we get (XIX), or: under condition C, the probability that Linda is a bank-teller and a feminist is higher than the probability that she is a bank-teller and not a feminist, and we briefly express it in our reasoning by proposition (T).

Concerning the reasoning presented above, we are dealing with two interpretations of the results of the Linda experiment. The first one ( $i_1$ ), derived from KT, supports System<sub>1</sub>, while the second interpretation ( $i_4$ ) is based mainly on System<sub>2</sub>. As mentioned earlier, KT worked within a certain paradigm that is closely related to the dual process theory, which makes these interpretations mutually inconsistent. Therefore, the second interpretation bears directly on the concept of KT. The following questions arise.

- Is anyone wrong here?
- If so, who is wrong?
- If KT are wrong, why?

A positive answer to the first question stems from the inconsistencies present in both interpretations. This inconsistency arises only if one accepts



the paradigm of KT, particularly the dual process theory.<sup>18</sup> In my opinion, while the estimation of probabilities is made by the intuitive part of the mind, which is similar to what KT desire, the reasons for this estimation come from analyzing the mind, which conflicts with KT's interpretation, i.e. ( $i_2$ ). For KT, the mechanism on which the subjects based their answers is heuristic, and in this case, the representativeness heuristic.<sup>19</sup> This heuristic, per KT, is being activated cognitively by Linda's description. It should be noted that the Linda task is one of the most important lessons for experimentally confirming the existence of the representativeness heuristic.<sup>20</sup> Regarding the second question, assuming the point of view presented in the second interpretation, KT's interpretation is incorrect. The answer to the third question requires a slightly longer argument.

There are three main arguments against the interpretation of KT. The first one was mentioned earlier and is related to *circular reasoning*. While the concept of logical probability is not well-known, the subjects had no choice but to make use of the intuitive concept of logical probability. For KT, seeking an experimental confirmation for their concept of the role of heuristics and intuition in cognition, in a way, "forced" the participants to use intuition in advance. To phrase this differently, they checked whether the subjects would use intuition while simultaneously provoking them to do so. This is quite surprising, since the title of Kahneman & Tversky's (1983) refers to both extensional (classical) probability and intuitive probability.

Secondly, as mentioned previously, KT worked within a certain paradigm, and through this, perhaps, made a so-called *systematic error*, where the researchers, "as a matter of principle," tried to interpret the results of any experiment to be in favor of the concept they had initially assumed.

---

<sup>18</sup> If there were only one system of mind, the KT thesis would become somewhat trivial.

<sup>19</sup> This heuristic is an intuitive method of reasoning which allows one to classify a described object in view of its resemblance to a typical representative of the class in question.

<sup>20</sup> This is not the only experiment that supports the concept of KT, as there have been multiple such experiments. Therefore, it cannot be stated, based on my criticism, that the concept of KT has been countered. My comments pertain to the Linda experiment only.

The confirmation of such a suspicion can be found in and reconstructed from what KT offer as reasons for the rejection of the possibility that the subjects had interpreted (T) as  $(T \wedge \neg F)$ . Let us recall what they say in this context:

Since assessment of probability makes sense even if one event includes another, subjects had no reason to interpret (T) as  $(T \wedge \neg F)$ . The response pattern obtained using the new version was the same as before. (Kahneman & Tversky 1983, 299)

This is a peculiar statement, since it seems that KT use the classical concepts of probability (*event, inclusion*) in this statement, while during the experiment, they asked about logical probability. Furthermore, they think that the event in the conjunction  $(T \wedge \neg F)$  mentioned earlier is contained in the event (T), which, according to KT, is sufficient for rejecting that interpretation of proposition T.<sup>21</sup> Consequently, as KT conclude at one point, comparing the probability of the second event to the probability of the event  $(T \wedge F)$  makes more sense than comparing it to the probability of the event  $(T \wedge \neg F)$  itself.<sup>22</sup> KT also performed additional (control) experiments designed to exclude the interpretation of (T) as  $(T \wedge \neg F)$  by the subjects. However, the subjects continued to attribute higher probability to the conjunction than to its conjuncts (cf. Miyamoto, Gonzalez & Tu 1995, 336-339).<sup>23</sup>

The third objection concerns assuming Linda's task to be the whole task, or a *pars pro toto* error. KT asked the participants to answer the question, (PL) "Which sentence is more probable, (T) or  $(T \wedge F)$ ?" This problem is, even syntactically, a subproblem of the entire Linda task<sup>24</sup> and the correct answer to this question is, of course, that the probability of the conjunction is equal to or less than the probabilities of its conjuncts. Such an answer, held by KT to be the correct one to (PL), is, according to my

---

<sup>21</sup> In citations (and comments) I preserve the signs used by KT.

<sup>22</sup> Such an argument is even more peculiar. I admit I did not understand it correctly.

<sup>23</sup> This trend in KT's research is not clear to me.

<sup>24</sup> Let us take the Linda problem OP and treat it strictly syntactically as a set of expressions of some language. The subproblem of the problem OP, in a syntactic sense, shall be called problem P, where  $(P \subset OP)$ .

interpretation, incorrect. It is not the answer to the *entire* Linda task, or the *entire* text of the task, since it is merely the answer to the question (PL).

Let us try to summarize our discussion so far. We raised some objections to KT's interpretation ( $i_1$ ) of the Linda problem experiment interpretation. This interpretation was guided by their assumption that the judgement of the experiment's outcome was understood in terms of activation of the representativeness heuristic that is, using current terminology, the activation of System<sub>1</sub>. Criticism rested on building a different interpretation ( $i_4$ ), according to which System<sub>2</sub> was used. In my interpretation, there is no need to refer to the heuristics – referring to the analytical System<sub>2</sub> is sufficient. Its effect is presented in the form of interpretation ( $i_2$ ). In general, when it comes to the description of mind, the results of KT's researches in the domain of cognitive psychology and the theory of decision-making are among the most crucial ones. In 2002, Kahneman received the Nobel Prize<sup>25</sup> in economics for his psychological works, which undermined the traditional model of human rationality. To prevent the reader from arriving at any incorrect conclusions regarding this paper, it must be stressed that there are multiple experimental results that confirm KT's conception.<sup>26</sup> Thus, in no case did my paper seek to challenge any general arguments set out by KT. Here I was concerned only with a criticism related to their interpretation of the Linda problem.

It seems that it would be very interesting to perform an experiment in which the reasoning method for a certain problem associated with probability would be compared, on the one hand, in accordance with the representativeness heuristic, and, on the other hand, in accordance with the logic used in mathematics and philosophy, not necessarily classical one. Such an experiment could bring a new look at the Linda problem, although not an attempt at KT's experiment. Taking my considerations into account, it seems very likely that the results of such an experiment would be inconsistent. This would be an interesting conclusion, indicating that the human mind is much more complex, and that it is difficult to predict with certainty the probability of human decisions.

---

<sup>25</sup> A. Tversky died in 1996.

<sup>26</sup> Cf. Kahneman (2013), and the two additions to that work, which are reprints of important KT publications.

## References

- EVANS, J. & STANOVICH, K. (2013): Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science* 8, 223-241.
- HERTWIG, R. & GIGERENZER, G. (1999): The 'Conjunction Fallacy' Revisited: How Intelligent Inferences Look Like Reasoning Errors. *Journal of Behavioral Decision Making* 12, 275-305.
- KAHNEMAN, D. (2013): *Pułapki myślenia. O myśleniu szybkim i wolnym (Thinking, Fast and Slow)*. Poznań: Media Rodzina.
- KAHNEMAN, D. & TVERSKY, A. (1983): Extensional Versus Intuitive Reasoning. The Conjunction Fallacy in Probability Judgment. *Psychological Review* 90, 293-315.
- MIYAMOTO, J. M., GONZALEZ, R. & TU, S. (1995): Compositional Anomalies in the Semantics of Evidence. In: Busemeyer, J., Hastie, R. & Medin, D. (eds.): *Decision Making from a Cognitive Perspective*. San Diego: Academic Press, 319-384.
- STANOVICH, K. (2010): *Decision Making and Rationality in the Modern World*. Oxford: Oxford University Press.