# The Double Disjunction Task as a Coordination Problem

Justine Jacot

University of Lund, Department of Philosophy
justine.jacot@fil.lu.se

**Abstract.** In this paper I present the double disjunction task as introduced by Johnson-Laird. This experiment is meant to show how mental model theory explains the discrepancy between logical competence and logical performance of individuals in deductive reasoning. I review the results of the task and identify three problems in the way the task is designed, that all fall under a lack of coordination between the subject and the experimenter, and an insufficient representation of the semantic/pragmatic interface. I then propose a reformulation of the task, that makes explicit the underlying semantic reasoning and emphasizes the difference of interpretation of the DDT between the experimenter and the subjects.

## 1 Introductory Remarks

### 1.1 The Double Disjunction Task

The *double disjunction task* (DDT), introduced by Johnson-Laird et al. (1992), is an inferential task where subjects are asked to say "What, if anything, follows" from a set of two premises, typically of the form:

Alice is in India or Barbara is in Pakistan [or both/but not bot] ($P_1$)

Barbara is in Pakistan or Cheryl is in Afghanistan [or both/but not bot] ($P_2$)

The optional parameters (in square brackets) are always covariant in premises. In the terminology introduced by Johnson-Laird et al. (1992), the authors determine an *inclusive* and an *exclusive* variant, when (resp.) the first and second values are chosen. These variants are called *affirmative*, as opposed to those obtained when ($P_2$) is substituted with:

Barbara is in Bangladesh or Cheryl is in Afghanistan [or both/but not bot] ($P'_2$)

which are called *negative*, because of the incompatibility between ($P_1$) and ($P'_2$), that gives rise to a contradiction. With subscript for inclusive or exclusive, and superscripts for positive and negative, one obtains four variants, $\text{DDT}_{\text{I}}^{+}$, $\text{DDT}_{\text{E}}^{+}$, $\text{DDT}_{\text{I}}^{-}$ and $\text{DDT}_{\text{E}}^{-}$.

The task was aimed at enforcing some predictions by the mental model theory (see below) according to which the more models an agent has to represent the less she is able to make correct inferences from the premises. Inclusive disjunctions should be therefore more difficult to cope with than exclusive ones, as well as disjunctions involving a contradiction in the premises. For matters of clarity, here is the truth table for the inclusive (noted $\vee$) and exclusive (noted $\underline{\vee}$) disjunctions:

| $P$ | $Q$ | $P \vee Q$ | $P \underline{\vee} Q$ |
|---|---|---|---|
| T | T | T | F |
| T | F | T | T |
| F | T | T | T |
| F | F | F | F |

**Fig. 1.** Truth table for inclusive and exclusive disjunctions

### 1.2 The Mental Model Theory

The *mental model theory* (MMT) has been developed in reaction against the idea of a 'mental logic' according to which individuals have logical rules in their mind and draw logical inferences in line with these rules. According to the MMT, people make logical inferences by following neither the rules of natural deduction nor those of truth tables, but by constructing mental models of the premises from which they derive a model of the conclusion. Deductive reasoning does not depend on syntactic rules but on a three-steps semantic process. First, semantic information is mentally construed as a representation of situations given by the premises. Second, an attempt to draw a conclusion from the mental models of the premises is made. Third, the validity of the conclusion is checked by ensuring that no model of the premises renders it false. For instance, a disjunction such as:

There is a club or there is a spade

calls for two models:

♣

♠

When the following premise is added:

There is no club

the model representing the club is discarded and the informational content of the new premise is added to the remaining model:

¬♣ ♠

The conclusion 'There is a spade' can be drawn, since there is no model of the premises that makes this conclusion false.

Relying on the assumption that individuals are fundamentally rational but make errors in practice, Johnson-Laird identifies three principles for deductions made by 'logically untutored individuals':

> [They] eschew conclusions that contain less semantic information than premises. [...] Similarly, they seek conclusions that are more parsimonious than the premises. [...] They try to draw conclusions that make explicit some information only implicit in the premises. In short, to deduce is to maintain semantic information, to simplify, and to reach a new conclusion. (Johnson-Laird et al. 1992, p. 419)

These constraints on deductive reasoning are close to Grice's maxims of conversation (Grice 1989). These maxims ask for agents involved in a communicative situation to be as informative as is required (maxim of quantity), to try to make their contribution true (maxim of quality), to be relevant (maxim of relevance), and to be perspicuous (maxim of manner). Although Grice is mentioned only in passing by Johnson-Laird et al. (1992, p. 419), their analysis of deduction is recovered in these maxims: maintaining semantic information pertains to the maxim of quantity, non redundancy and simplification to the maxim of manner, and reaching a new conclusion to the maxim of relevance, while deduction guarantees the quality of the conclusion, conditional on the quality of the premises.

Grice stresses that conversations where the maxims apply are a special case of cooperative situations, and that these maxims are instantiations of a more general Cooperation Principle (CP). An experiment is, after all, a cooperative situation, and therefore the effect of CP should be evaluated at the beginning of the task, because deductive reasoning of lay people is constrained by semantic principles as well as pragmatic principles.

The remainder of this paper is organized as follows. The next section firstly discusses the results of the DDT, secondly identifies the main issues at stake in the interpretation of the results and the way the task is formulated. Section 3 shows that most of these problems arise because of a lack of coordination between the subject and the experimenter in the interpretation of the task. Section 4 suggests some possible reformulations that could improve the base rate of originally expected answer, by eliciting both the adequate modal representation, and the selection of relevant information, through pragmatic principles.

## 2 Results and Interpretation of the DDT

### 2.1 Results

Johnson-Laird et al. (1992) discuss only the selection made for the 'positive' DDT, consisting in the following two statements:

Alice is in India and Cheryl is in Afghanistan, or Barbara is in Pakistan, or both $(C_i^+)$

Alice is in India and Cheryl is in Afghanistan, or Barbara is in Pakistan, but not both $(C_e^+)$

The valid conclusion for the negative disjunctions were predicted to be more difficult to make because they suppose the ability to first detect the contradiction in the premises. The percentages of valid conclusions made by the subjects for the four kind of disjunctions are displayed in Fig. 2.

| | Affirmative | Negative |
|---|---|---|
| Exclusive | 21% | 8% |
| Inclusive | 6% | 2% |

**Fig. 2.** Percentage of valid conclusions for the DDT

As predicted by the MMT, the percentage of success decreases when the number of models to build increases, coupled with a floor effect: when a deduction needs three or more models, it is almost impossible for the subjects to make a correct inference, or even to decide whether a conclusion follows from the premises.

But while Johnson-Laird is interested in the link between one particular case of error with one particular subset of the models of the premises (Johnson-Laird et al. 1992, p. 433-434), and although the prediction of MMT for relative rates of success are confirmed, the low base rates remain only imperfectly explained. Besides, a non negligible number of subjects shows an ability to perform disjunctive reasoning. The results reported by Toplak and Stanovich (2002) show a higher success rate for the DDT (37%), because the authors collapsed in one 'super-category' all the conclusions that displayed a "disjunctive pattern" of reasoning.[1]

The two categories collapsed into the disjunctive responses were what Toplak and Stanovich (2002) call 'the *partial contingency*' response and the 'Alice-Cheryl response', which were both given by 12% of the subjects. The partial contingency response, for instance, 'If Alice is in India and Cheryl is in Afghanistan, Barbara is not in Pakistan', is given when people draw "the implications of one of the disjunctive path but not the other." (Toplak and Stanovich 2002, p. 203). According to Toplak and Stanovich (2002), it seems in this case that subjects are able to assign hypothetic truth values to each disjunct in the premises but unable to represent situations where all states of affairs hold, comprising those where Alice is not in India and Cheryl is not in Afghanistan.

### 2.2 Identifying the Problems in the Design of the DDT

MMT explains failure in terms of the number of mental models to be construed in order to solve the task. Reasoning being semantic, the semantic content affects the resolution of the task. The more models individuals have to process, the less they succeed in making correct inferences, due to the difficulty to store in the working memory the totality of the models of the premises needed to draw the correct inferences, particularly when different situations have to be processed in parallel.

While MMT assumes that "naive individuals have a modicum of deductive competence" (Johnson-Laird 1999, Johnson-Laird and Byrne 1991) but that they (sometimes) fail at deductive performance, things have to be viewed the other way round. The theory predicts that, faced with a deductive inference, individuals apply some filtering constraints depending on the kind of inference they have to make. I suggest that they should be viewed as applying some relevance constraints because of the form of the reasoning they have to make, and then draw the conclusions that seem relevant.

Changing the perspective preserves some insights from MMT (i.e. why 'trivial' conclusions are not even considered), but sheds light on other aspects left unexplained by the theory. For instance, answers in disjunctive form are more common than in conditional form (see Toplak and Stanovich 2002), which seems to indicate that people tend to preserve the syntactic form of the premises.

[1] What Toplak and Stanovich (2002) call 'disjunctive reasoning' is the construction of mental models for the premises and at least a partial combination of them in the conclusion. The normative answer (fully disjunctive) was given by only 13% of the subjects, which is lower than in Johnson-Laird et al. (1992).

One may then identify three problematic categories in the formulation of the DDT. The first category of problems concerns the logical form of the premises. First, the conjunction of the two premises is implicit, but this conjunction must become explicit in order for the intended (correct) conclusion to be drawn. Second, it is not clear whether the disjunct 'Barbara is in Pakistan' should be at the same place in the two premises, or at the first place in the first premise and at the second place in the second premise. Indeed, it is not sure that not logically trained individuals always interpret the disjunction as commutative in the DDT. But at the same time, it seems that people want to preserve the syntactic form of the premises: answers in disjunctive form are more common than in conditional form. Although Johnson-Laird et al. (1992, p. 433) explained subjects the difference between an inclusive and an exclusive disjunction, guaranteeing a common interpretation of "... or ... or both" and "... or ... but not both" (between participants and experimenter), they give little detail about this explanation, save for the fact that it did not rely on truth-tables or natural deduction rules.

The second category pertains to the linguistic context opened by the DDT. First, it seems that, in designing the task, Johnson-Laird and his colleagues wanted to avoid the problem of an abstract task by giving the subjects a thematic task. But it is unclear whether the difference would negatively affect the results. Indeed, the premises in the DDT are rather abstract since they present names and places out of the blue. No context or situation supports the story. Second, nothing in the task says that 'Barbara' in both premises is the same individual, for the reasons mentioned above.

The third category of problems relates to the way the DDT is framed in general, and particularly the way the question is asked: 'what, *if anything*, follows from these premises?' Either classical logic is assumed and an infinite number of valid conclusions follows from the premises, and then subjects have no time or no space to give the full answer; or the answer is intended to be the conclusion that preserves the most semantic information from the premises and 'if anything' can be dropped from the question.

In what follows I will show that these issues can be addressed if we reconsider the DDT as a coordination problem between subject and experimenter. I will focus on the second and third category of problems, the first one being answered in the new formulation of the task I propose next.

## 3 A Coordination Problem

### 3.1 Narrowing Down the Context

Experimenters take often for granted that only one possible interpretation of the logical vocabulary is correct, and think that departure from this interpretation is an error. But, as shown by Stenning and van Lambalgen (2008, chap. 3), the problem is that experimenters also leave under-determined this interpretation for the subjects, to such an extent that subjects sometimes complete the under-determined semantic content in an unintended way. If MMT is correct and logical reasoning is indeed semantic, then the DDT should be able to provide tests such that either the semantic content is determined enough to yield a single interpretation, or the normative answer allows for different interpretations.

As noted by Shafir (1994), one problem with the DDT is that it demands subjects to "think through" the disjuncts, and to draw a conclusion from both the premises. In order to do so, subjects must process all the premises at the same time, but they appear on paper as two separate sentences, and the conjunction of the premises remains implicit. While the disjunctive form may trigger some ability to reason by cases (even though the performance's rate is low), what in fact is needed is to reason by cases *together with* the ability to keep track of the different possibilities, as e.g. answers to different questions regarding the premises.

Moreover, as put forward by relevance theory (Sperber et al. 1995, Sperber and Wilson 1995, Wilson and Sperber 2004) there is a discrepancy between the laboratory task and the real life situations in which the subjects might have to perform these kind of inferences. The lack of relevance of the task at stake, or the lack of communicative interaction can yield some pragmatic 'disabilities.' On a linguistic level, some linguistic forms produce some kind of relevance 'filters' content- and context- dependent which, in turn, induce a preference ordering over possible interpretations (Girotto et al. 2001, p. 70). On one hand, in the type of task presented to the subjects, logically relevant and logically irrelevant content are mixed, whereas it is asked to the individuals to pay attention only to the logically relevant content. On the other hand, in order to achieve the task, individuals have to interpret it correctly. They must therefore pay attention to the logically relevant *and* irrelevant content. But the logically irrelevant content sometimes prompts individuals to a certain interpretation, which is not always the interpretation the experimenter has in mind.[2]

Further discussion of the relevance theory explanation of the success and failure of logical reasoning is not necessary, since my argument needs no other agreement than retaining its central insight. The DDT is content- and context-dependent and it seems reasonable to think that subjects interpret it as such, in a way that intermingles semantic and pragmatic inferences in reasoning.

### 3.2 Semantics vs. Pragmatics and the Principle of Cooperation

The way the question is asked in the DDT is disputable, for reasons that are admitted by Johnson-Laird himself, but that he does not seem to factor in the formulation of the question. Indeed, Johnson-Laird acknowledges that "infinitely many valid conclusions follow from any premises" (Johnson-Laird 1999, p. 113) but, by saying 'if anything', nevertheless implies that it could be the case that nothing follows from them. This is obviously triggered by the three principles of deductive reasoning already mentioned in section 1.2. Assuming these facts, it is highly improbable that subjects will draw conclusions such as any conjunction of the premises.

If it is true that logically untrained individuals do not draw 'trivial' conclusions from a set of premises, there is no need to try to lure them in thinking that there might be

[2] Sperber et al. (1995, p. 44): "Past discussions of subjects' performance have tended to focus on the task as already interpreted (by the experimenter). [...] But interpreting the task is part and parcel of performing it, and obeys criteria of rationality in its own right. The study of 'content effects' is the study of sound cognitive processes that are by no means out of place in subjects' performance."

some. Besides, the formulation of the question must indicate that there is some relevant information that can be inferred from the premises. Indeed, the Gricean CP would ask for a speaker's intentions to be recognized by the hearer in order for the speaker's utterance to be correctly interpreted. Why not assuming such a process in the DDT? As previously mentioned, a psychological experiment involves a certain level of cooperation between subjects and the experimenter. Certainly CP governs inferences to intentions and goals of others, and reconstructing them in DDT inevitably involves speculation. But it is not impossible to reconstruct a possible route to $C_i^+$ and $C_e^+$ that takes CP more seriously than the original theory.

## 4 Two Versions of the DDT

### 4.1 The Intended Interpretation

The theory predicts that subjects will: (*a*) compute a representation of three (two) mental scenarios, or 'mental models', for each inclusive (resp.: exclusive) premise; (*b*) combine the representations, eliminating those incompatible with the information; and: (*c*) 'read off' some appropriate conclusion and answer the question.

In this general setting, the intended interpretation requires an abstraction step. In step (*a*), subjects are supposed to replace the premises with a variable $X \in \{A, B, C\}$ where $A, B, C$ stand for each of the individual in the disjunct, to which a value is assigned, equivalent to 'I know/I do not know where is $X$'. Step (*b*), for each variant of the DDT, is tantamount to assigning a value to each variable, compatible with the constraints embodied in the models of the premises.

Comparatively, subjects should be expected to have greater difficulties to solve the 'negative' tasks. Johnson-Laird et al. explain this difficulty by the necessity of obtaining an additional (implicit) premise, which follows from subjects background knowledge:

> Of course, if [Barbara] is in [Bangladesh], then [she] is not in [Pakistan]. According to the model theory, affirmative deductions should be easier than such negative deductions because the latter call for the detection of the inconsistency between the contrary constituents. (Johnson-Laird et al. 1992, p. 433)

The detection of inconsistency is an additional (inferential) step, left unanalyzed by Johnson-Laird et al. This step is equivalent to impose additional constraints to compute the composition of the representations associated with ($P_1$) and ($P'_2$), using background knowledge. The additional complexity of tracking combinations of two values for parameter $B$ in both $\text{DDT}_\text{I}^-$ $\text{DDT}_\text{E}^-$, explains the cost of "detection of inconsistency" by the stress imposed on computational abilities. In addition, $\text{DDT}_\text{E}^-$ outputs an additional mental model, as compared to $\text{DDT}_\text{E}^+$, inducing a heavier load on working memory.

So far, this reconstruction of steps (*a*) and (*b*) of DDT agrees with Johnson-Laird et al. However, reconstruction of step (*c*) raises serious issues with both the design of the experiment, and the interpretation of its results. Johnson-Laird et al. discuss only the selection made for the 'positive' DDT. Their justification for this selection is the "support" mental models yield to these conclusions.

Indeed, treating $A$, $B$ and $C$ as propositions, and substituting known (asserted) and unknown (excluded) locations with 1 and 0, respectively, one obtains distributions of

truth-values that satisfy ($C_i^+$) and ($C_e^+$). However, this explanation is at odds with the claim that reasoning based on mental models does not proceed by building and reading equivalent of truth tables, since "truth tables [...] are too bulky to be mentally manipulated" (Johnson-Laird et al. 1992, p. 421).

Considering the representation Johnson-Laird et al. gives for the set of final mental models in $\text{DDT}_\text{I}^+$ and $\text{DDT}_\text{E}^+$, reproduced Fig. 3, ($C_e^+$) is the most natural of the answers considered. If the mental construction is actually carried as hypothesized by Johnson-Laird et al., ($C_e^+$) can be 'read off' neglecting excluded locations, and using 'or... but not both' to express the existence of (two) alternative scenarios. This method does not rely on explicit substitution with truth-values, but has no straightforward equivalent in $\text{DDT}_\text{I}^+$ that would read ($C_i^+$) off.

| Alice | Barbara | Cheryl | Alice | Barbara | Cheryl |
|---|---|---|---|---|---|
| [I] | | [A] | [I] | [P] | [A] |
| | [P] | | [I] | [P] | |
| | | | [I] | | [A] |
| | | | | [P] | [A] |
| | | | | [P] | |

**Fig. 3.** Mental Models for $\text{DDT}_\text{E}^+$ and $\text{DDT}_\text{I}^+$

One could hypothesize that subjects approximate $\text{DDT}_\text{I}^+$ by addressing it *as if* it were $\text{DDT}_\text{E}^+$, and then simply switch the proviso from "but not both" to "or both." However, this rationale for the choice of ($C_i^+$), is not consistent with the data, in particular the relative difference between the rate of success in $\text{DDT}_\text{E}^+$ and $\text{DDT}_\text{I}^+$. Johnson-Laird et al. may have selected ($C_i^+$) by the above reasoning, or simply by judging ($C_e^+$) intuitively natural enough, and, switching the proviso, assuming that ($C_i^+$) was natural, too. However they give no hint has to what their selection is based on, and rest content with the relative frequencies, that validate the predictions of MMT.

In the absence of an analysis of the 'reading off' method that outputs ($C_i^+$), it seems difficult to argue whether the answer corresponds to some natural method. Likewise, the absence of discussion of answers expected to be given to $\text{DDT}_\text{I}^-$ and $\text{DDT}_\text{E}^-$ makes the assessment of the task more complicated than Johnson-Laird et al. seem to assume.

### 4.2 The Abstraction Problem

The naturalness of ($C_e^+$) depends essentially on the representation of the problem that neglects *specific* locations, and in which one checks only whether a location is known, or not. If one assumes with Johnson-Laird et al. that reasoners somehow automatically select representations that are as implicit as possible, the inescapable conclusion is that the underlying automatisms are not very good at selecting the level of abstraction that makes the task solvable.

Yet, abstraction in variants of the DDT may not be as simple as assumed by Johnson-Laird et al. Reading (or hearing) ($P_1$), ($P_2$) and ($P'_2$), in either their inclusive or exclusive variants, may elicit a very different representation, if the range of values is taken to be

the set of locations. Then, upon reading (hearing) ($P_1$), a subject may consider that the set of relevant values is:[3]

$$v_{(P_1)} = \{i, p, \overline{ip}\} \qquad (v_{(P_1)})$$

(where $\overline{ip}$ corresponds to any other possible location than India and Pakistan). However, reading ($P_2$) produces a new set of values for that premise, as well as an update of ($v_{(??)}$), as follows:

$$v'_{(P_1)} = \{a, i, p, \overline{aip}\} \qquad (v'_{(P_1)})$$

$$v_{(P_2)} = \{a, i, p, \overline{aip}\} \qquad (v_{(P_2)})$$

The effect is even more striking with ($P'_2$), where the number of seemingly relevant values (and therefore of their combinations) increases even more, with:

$$v''_{(P_1)} = \{a, b, i, p, \overline{abip}\} \qquad (v''_{(P_1)})$$

$$v_{(P'_2)} = \{a, b, i, p, \overline{abip}\} \qquad (v_{(P'_2)})$$

Given the complexity of the task, one will typically not complete step ($a$) before the time the answer has to be given. Again, when "unable to construct a set of models or [...] to discern what holds over all of them" subject will respond that *no valid conclusion follows*.

## 5 Conclusion

DDT as it is designed in psychology settings, is not a bad test for reasoning skills in the sense that it would mistake the competence to test. The problem comes from the way experimenters interpret the task itself. The difference in interpretation between experimenters and subjects is what leads to a misinterpretation of the results, and shows the role of relevance and computational complexity in the answers of the subjects. Reconsidering the DDT as a coordination situation brings to light the reasoning process of the subjects in drawing inferences from a double disjunction. This process, as shown in the sketch of the formal reformulation of the task above, is a sequential procedure which starts from a step-by-step construction of representations task, and outputs a global conclusion. Our reformulation emphasizes the discrepancy between the intended interpretation of the DDT by the experimenters, and the task as it may be understood by the subjects.

What remains to do in a further research is, from a theoretical perspective, building a complete formal representation of the DDT displaying the intended interpretation by the experimenter and the possible interpretation(s) by the subjects. This reformulation of the task relies on the difference made by Stenning and van Lambalgen (2008) between reasoning *to* an interpretation and reasoning *from* an interpretation: the authors argue that subjects must always first *reason to* the experimenter's (intended) interpretation before they can *reason from* it to a solution.

[3] A symmetric argument can be given for the case where ($P_2$) or ($P'_2$) are presented first.

From an empirical perspective, we need to design an experiment based on the explanatory hypothesis argued for in this paper, especially concerning the level of abstraction required from the subjects in order to solve the task. To this effect, it seems desirable to handle both sides of the problem by testing subjects' performance in a fairly abstract task where the disjuncts are displayed as propositional variables such as $A, B, C$, and compare the results with a DDT where the premises are contextually fleshed out by a story to fulfill the need for a thematic task.

# Bibliography


Girotto, V., Kemmelmeier, M., Sperber, D., and van der Henst, J.-B. (2001). Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task. *Cognition*, 81:69–76.

Grice, H. P. (1989). Logic and conversation. In *Studies in the Way of Words*, pages 22–40. Harvard University Press.

Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual Review of Psychology*, 50:109–135.

Johnson-Laird, P. N. and Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N., Byrne, R. M. J., and Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, 99:418–439.

Shafir, E. (1994). Uncertainty and the difficulty of thinking through disjunctions. *Cognition*, 50:403–430.

Sperber, D., Cara, F., and Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, 57:31–95.

Sperber, D. and Wilson, D. (1995). *Relevance. Communication and Cognition.* Blackwell, 2nd edition.

Stenning, K. and van Lambalgen, M. (2008). *Human Reasoning and Cognitive Science*. The MIT Press.

Toplak, M. E. and Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, 94:197–209.

Wilson, D. and Sperber, D. (2004). Relevance theory. In Horn, L. R. and Ward, G., editors, *Handbook of Pragmatics*, pages 607–632. Blackwell.