The experimental investigation of the updated traditional interpretation of the conditional statement

András Veszelka

Pellea Human Research & Development Ltd., Budapest, Hungary
andras.veszelka@gmail.com

Abstract. It is well known that the interpretation of the conditional statement in "everyday life" deviates from the official logical approach. It is conceivable, however, that the ancient logicians who first demonstrated the official approach erroneously characterised the "if P or R then Q" relationship in place of the "if P then Q" statement. When fixing this error, it turns out that the equivalent interpretation of the conditional statement, which is traditionally seen as one of the most common everyday fallacies, is in fact exactly the correct interpretation. Since classical logic has not been built on mathematical grounds but rather on philosophical argumentations and insights, its findings can be tested with the tools of today's human sciences, among others, with empirical experiments. The main experimental tools support this updated logical approach, and show that everyday thinking can be made compatible with logic. These results are summarized in this study.

Keywords. scholastic logic, psychology of reasoning, human experimentation, Wason selection task

1 Introduction

The conditional statement is a glaring example of how the abstractions of logic and the everyday use of the logical connectives deviate from each other. Many interpret this as the difference between formal and natural languages (e.g. [1]). This differentiation can be traced back to the beginning of the 20th century, where, for example, Frege [2] argued that the difference between the interpretations of the conditional statement as prescribed in logic and as used in "everyday life" reveals linguistic or psychological components. This is where the search for the linguistic or psychological components deemed different from logic began, first in the philosophy of language, then in linguistics and finally in psychology. However, the so-called everyday interpretation of the conditional statement does not merely deviate from such formal languages as propositional logic, which was born in Frege's era, but also from the classical interpretation of the conditional statement, which is basically the same as that of propositional logic, but which has been clearly created not in a mathematical, but in a linguistic, philosophical and psychological environment. Thus, the discrepancy between the everyday interpretation and the classical abstraction is within the same system.

1.1 The abstraction error in classical logic

That said, when looking back to the classical interpretation, it can be seen that it is erroneous. Instead of the "if P then Q" statement, classical logicians have erroneously abstracted the "if P or R then Q" statement. Let's take an example from Jevons [3] (p. 70), a late scholastic logician:

If the snow is mixed with salt, it melts

As is well known, in this if P then Q statement from the snow mixed with salt (P) antecedent, it is correct to infer the snow melts (Q) consequent (modus ponens, MP), and from the snow not melting (not-Q) it is correct to infer that it has not been mixed with salt (not-P) (modus tollens, MT). However, the denial of the antecedent (DA, if the snow is not mixed with salt (not-P), it does not melt (not-Q)) and the affirmation of the consequent (AC, if the snow melts (Q), it was mixed with salt (P)) are incorrect. Jevons has argued that, for example, from the snow melting (Q), it does not follow that it has been mixed with salt (P) because it can melt by other means as well. It is impossible to find any other explanation, even going back several hundred years, as to why these two latter inferences are incorrect. On the contrary, this interpretation can be traced back even to Aristotle, who wrote that:

The refutation, which depends upon the consequent, arises because people suppose that the relation of consequence is convertible. For whenever, suppose A is, B necessarily is, they then suppose also that if B is, A necessarily is. This is also the source of the deceptions that attend opinions based on sense perception. For people often suppose bile to be honey because honey is attended by a yellow colour: also, since after rain the ground is wet in consequence, we suppose that if the ground is wet, it has been raining; whereas that does not necessarily follow ([4], 167b1ff.).

To complete this argument, the inference that it has been raining does not necessarily follow because there are other possible means to make the ground wet. However, if we refer to additional possible causes, that is, to additional possible antecedents during the abstractions, these have to be denoted. In logic, it is fundamental that "we restricted ourselves to explicitly stated premises" ([5], p 6.). With their denotation, however, it can be seen that with the above explanations we characterized the “if P or R then Q” statement. Jevons' example was therefore the "If the snow is mixed with salt or, for example, the sun is shining, the snow melts" statement that he erroneously characterised in terms of P and Q only. These non-abstracted alternative antecedents can be found in every example provided for the interpretation of the conditional statement in the history of logic.

1.2 The correct abstraction of the conditional statement

The question arises therefore, of what the correct abstraction of the conditional statement can then be, that is, the abstraction of the relationship in which there is no wedging "or R" component. I believe the correct inference pattern is the equivalence, in which all the MP, MT, AC, DA inferences are valid. For example, we endorse the equivalent AC and DA inferences for the "if-then" connective, even in the case of the "if P or R then Q" statement. For instance from Q, we endorse the affirmation of the consequent (AC) inference, and we deduce to "P or R". As the traditional interpretation goes, *within this* we do not infer exclusively to P because it can be R as well. This can be demonstrated the same way in the case of all three other classical inferences as well. On the other hand, classical equivalent statements such as, for instance, "if the sun is in the sky then it is day" are equivalent because the context of these statements does not allow one to wedge any alternative antecedents. There can be day only if the sun is in the sky. Even propositional logic refers to the alternative antecedents, as when it differentiates the equivalence (to use another term, the biconditional) with the artificial expression "If P *then and only then* Q" from the "if P then Q" conditional, in the latter case of which, by parity of argument, several antecedents can lead to Q. All of this is illustrated in further detail by Veszelka [6]. When explaining the interpretation of propositional logic within the if-then statement, Geis and Zwicky [7] reinvented and employed the aforementioned scholastic interpretation, and by mentioning alternative antecedents they managed to block one of the most common fallacies that people commit in the case of the conditional statement, the equivalent/biconditional inferences. This approach has subsequently been implemented in psychology, and its mechanism has to an extent been experimentally verified. Byrne [8] has demonstrated that if the second antecedent is connected to the initial antecedent with an "and" connective, then in terms of P and Q only, the MP and MT inferences would be invalid and the DA and AC inferences would remain valid. This is the case, for instance, in the example of the "If the snow is mixed with salt and it is not extremely cold, it melts" (If P and R then Q) statement. These are very interesting relationships, however, and as a consequence of the historical reasons illustrated in the beginning of this study, this phenomenon is interpreted in linguistics and in psychology as a linguistic, pragmatic effect, which is contrary to logic. It was nevertheless demonstrated above that this phenomenon is actually the update of the classical logical interpretation of the conditional statement. It is the exact definition of what differentiates between the two well-known inference patterns on the conditional statement, the traditionally accepted conditional inference pattern, which allows only MP and MT, and the equivalence. In antiquity, the rule of thumb used was that since the conditional statement can evoke both conditional and equivalent inferences, one should label only those inferences that are prescribed by both of them as valid, that is, the MP and the MT [9]. Obviously, the new definition is more accurate. However, many important psychological experiments are in conflict with this approach.

2 The experimental investigation of the conditional statement

2.1 The demonstration of the biconditional inferences

The most important task of this type, the "single most investigated problem in the literature on deductive reasoning" ([10], p. 224) is Wason's abstract selection task. Consequently, Byrne, who introduced the study of the alternative antecedents into psychology, has rejected the basic biconditional interpretation of the conditional statement [8]. In this task, participants are shown four schematic cards having a letter on one side and a number on the other. Participants are then asked what card or cards they would turn over in order to decide whether, for example, the "If there is a letter E on one side there is a number 4 on the other side" conditional statement is true. On the cards, the "E" (P), "K" (not-P), "4" (Q) and "7" (not-Q) can be seen. In this task, abstract letters and numbers are used in order to assure that the context and the content have no influence on the results and so they accurately display how people interpret the if-then statement itself. The traditional conditional interpretation would be selecting the cards P and not-Q, since these could have falsifying instances on their other side, while the biconditional interpretation would be the selection of all four cards. The customary response is, however, merely the P and Q value. In the psychological field on logical reasoning, the logical negation is expressed in three different ways. It can be implicit (e.g. "A", and its negation = "K"), explicit ("A" and its negation "not-A") and dichotom, which is the same as the implicit, but in which the task instruction states that only two possible values can be found (e.g. "A" and "K"). The result is P and Q with all three negatives [11]. This result constitutes an important basis for many theories in the field. There are three additional main tasks:

- Truth-table evaluation task, in which the given co-occurrences of the truth table of propositional logic, for instance the co-occurrence of P and Q, must be evaluated in terms of whether it verifies or falsifies the conditional statement, or is irrelevant to it.
- Inference task, in which on the basis of the provided conditional statement, people must decide if the given conclusions follow from the minor premises or not, for example whether or not from not-P, not-Q follows.
- Inference production task, in which participants themselves write down what follows from the minor premises.

The available results from the combination of the four tasks and the three types of negatives are shown in Table 1.

Table 1. Results of the main task types with the tree types of negatives

	Implicit neg.	Explicit neg.	Dichotom neg.
Selection task	P&Q	P&Q	P&Q
Truth table task	Def. table [12][1]	Def. table [12][1]	≡ (83%) [13][2]
Inference task	≡ (48%) [14][2]	?	≡ (60%) [14][2]
Inf product.task	?	?	≡ (92%) [13][2]

[1] Defective truth table
[2] Biconditional

As can be seen in Table 1, although there are biconditional solutions, the results are generally inconsistent and there are missing data. For this reason, I have retested all of the tasks [15] except for the abstract selection task, which has robust results for all three types of negatives. Consequently, for the selection task, I tested two thematic tasks that have an evidently biconditional context in order to check if the results of these tasks deviate from the results of the abstract selection task, or if they also evoke the preference of the P and Q values, as was already observed by some researchers. My results are shown in Table 2.

Table 2. My results of the main task types with the tree types of negatives [15]

	Implicit neg.	Explicit neg.	Dichotom neg.
Selection task	P&Q	P&Q	P&Q
Truth table task	Def. truth table	Def. truth table	Bicon (50%)
Inference task	Def. truth table	Def. truth table	Bicon (42%)
Inf product.task	Bicon (73.3%)	Bicon (52%)	Bicon (67%)

The reasoning contained in the defective truth tables[1] require further analysis, although there are several explanations for this phenomenon that are compatible with the updated scholastic approach. It is still apparent that the predominant response is the biconditional. With regard to the selection task, instead of the biconditional responses, both tested biconditional problems evoked the P and Q preference, the characteristic response of the abstract selection tasks. According to my hypothesis [15], which has been also formulated and partially tested by Wagner-Egger [14] one month prior to my study, people avoid the selection of all the cards in the selection task. They believe that selecting all four cards would be contrary to the task instruction, which in

[1] In the defective truth table the co-occurrence of "P and Q" verifies the conditional statement, the co-occurrence of "P and not-Q" falsifies it, and the "not-P and Q" and the "not-P and not-Q" co-occurrences are irrelevant to it.

fact requires them to select *from among* the cards. This is fairly apparent in the case of the following task, which was one of the tasks involving biconditional context that I have tested:

On one side of each card, there is the name of a city and on the other side there is a mode of transportation. Let us suppose that when someone goes to Budapest, he always goes by car, and when he goes to Szeged, he always goes by train. Likewise, when he travels by car, he always goes to Budapest, and when he travels by train, he always goes to Szeged. Mark the card or cards that must be turned over in order to decide whether this is true.

The following statements were printed on the cards: “going to Budapest”, “going to Szeged”, “going by train” and “going by car” ([15], Experiment 3).

In this task, which was tested on 2x20 participants, everyone produced the biconditional answer in the inference production task, but only 10% did so in the selection task. However, as it can be seen, the task was in fact a pseudo-problem, because it contained a clear description of what follows from what, or what value has to figure on the other side of the cards. I obtained the same result on another clearly biconditional problem, the so-called “ball-light” problem [16]. This problem is commonly accepted in the literature as a biconditional task which, being tested on 2x30 participants, has produced biconditional answers in 96% of inference production tasks, but only in 23% of selection tasks [15]. Since participants do not find a better solution than the avoided biconditional response, they finally select those instances that are named in the conditional statement, the P and Q values. Thus, the main experimental tasks altogether support the biconditional approach.

2.2 The demonstration of the response traditionally deemed correct

One half of the updated classical interpretation of the conditional statement, the basic equivalent interpretation, can be therefore experimentally demonstrated. Another empirical obstacle to this approach is to trigger the P and not-Q answer, the traditionally expected response in logic. The elicitation of the “correct” answers has so far been studied almost exclusively with selection tasks, and in so-called thematic selection tasks researchers obtained the allegedly correct P and not-Q response several decades ago. The most cited task of this type is the drinking-age task [17], in which participants have to imagine that they are on-duty police officers who must check if everyone observes the rule that “If someone is drinking beer, he must be older than 18 years”. “Drinking beer”, "Drinking soft drink”, “21 years old,” and “17 years old” appear on the cards. A large proportion of participants select “Drinking beer” and the “17 years old” cards in this task—that is, the P and not-Q cards. This result is interpreted to arise from various effects, such as from pragmatic reasoning schemas [18], from relevance [19], from deontic context [20], from precautions [21], from cheater detection [22], from altruist context [23], from perspective switching [24], or from benefits or costs [5]. However, these are not normatively valid explanations, because

in classical logic or propositional logic, where the abstraction itself has been defined, such components were clearly not present. This can be easily seen in the examples mentioned at the beginning of this study as well. It can be observed, however, that there is a wedging of information in the easy-to-resolve selection tasks as well, which mainly correspond with the effect of the alternative antecedents in the updated interpretation of classical logic: In the above task everyone knows that people above 18 years can drink both alcohol and other beverages, although this is not explicitly communicated in the task instruction. One of the experiments of Hoch and Tschirgi [25] can be seen as a means to test this additional information, in which they used in an abstract task, with the appropriate substitutions, the statement that "Cards with a P on the front may only have Q on the back, but cards with not-P on the front may have either Q or not-Q on the back" ([25], p 203). Although this cue facilitation produced 56% correct results in the experiment of Hoch and Tschirgi [25], in the replication of the experimental condition [26] the rate was only 36% in the usual experimental population, and participants with knowledge of logic were not filtered out; this could evidently improve the result. With the usual experimental population, only a modest improvement was received with this type of facilitation [27]. This task has so far been tested only in selection tasks. In an unpublished experiment (with 21 participants), I also received correct answers only in 14% of selection tasks, but the rate was 76% when the very same task was presented in the inference-production task Pearson Chi-Square (1, 42) = 16.243, $p < .0001$, Cramer's $V = .622$. Perhaps in this case once again, people in the selection task would test the complete relationship, and test, for instance, that both P and not-P can figure on the card with a Q on its other side. This would again require turning over all four cards, and as such the distorting effect mentioned above could reappear. Similarly, it can be observed that, contrary to this facilitation attempt, in the above easy-to-resolve drinking-age task the relationship that people above 18 years can also drink soft drinks is from outside of the task, it is not included in the investigated conditional statement. As a result, it must not be part of the examination. To test this assumption in an abstract selection task, in an unpublished experiment I used the following task:

Imagine that four cards are lying in front of you on the table. On one side of the card there is either the number 4 or the number 6; on the other side, there is either „divisible by two" or „divisible by three". Your task is to check whether the four cards on the table each conforms with the reality, namely, with the rule that

If the number is 4, then it is divisible by only two

Which card or cards would you turn over to check this?

In the control task I replaced 6 with 3 in the instruction and on the second card, and in order to assure better text comprehension, I removed the word "only" from the if-then statement. According to my interpretation, therefore, the two tasks evoke two different relationships as shown in Figure 1.

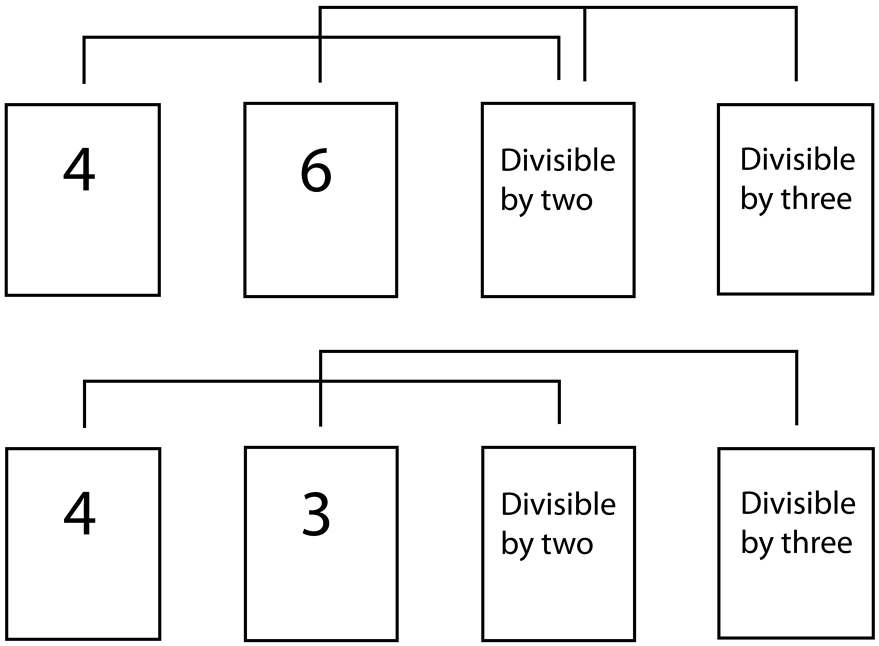

Fig. 1. Cards and evoked relations in the experiment on the "easy to resolve" abstract selection task

With number 6, the task produces the conditional inference pattern, and with number 3 the biconditional pattern. Indeed, with 21x22 participants I obtained P and not-Q responses on the conditional task in 41% of cases. The rate of the "P and Q" and "all" responses being in conformity with the biconditional was only 13.5%, whereas in the control task the rate of the "P and not-Q" answers was merely 4%, and the "P and Q" and "all" comprised 86% of the results. The difference is obvious with Pearson's Chi-Square (3, 43) = 20.157, $p < .0001$, with Cramer's V = .685, and because of the fact that the minimal differences between the tasks explain themselves. However, at a different university, where participants were given twice as much time to resolve the task, I failed to reproduce these results. It was then raised by a colleague that IQ scores have a similar difference between the two universities, and that IQ could possibly also play a role in the way the tasks are resolved. For this reason, with Anikó Kecse Nagy, we tested the task in the summer camp of Mensa HungarIQa. This organization collects Hungarians older than 17 years of age and who obtained a score on the Raven Advanced Matrices IQ test higher than 98% of the general Hungarian population. In this experiment, performed with 20x16 participants, people with and without knowledge of logic participated equally. The results altogether were 69% vs. 30% "P and not-Q" answers, Pearson Chi-Square (1, 36) = 5.355, $p < .021$, Cramer's V = .0386 for the conditional task, versus the biconditional task. The difference among the "all" biconditional answers was also significant in the opposite direction

(25% vs. 0%), Pearson Chi-Square (1, 36) = 4.654, $p < .031$, Cramer's V = .359. Deconstructing the results further, 55% of the participants with no knowledge of logic (9 subjects) gave "P and not-Q" answers to the conditional task, while 25% of them gave the same answers to the biconditional task. Even participants with knowledge of logic produced significantly more P and not-Q answers to the conditional task (86% vs. 33%), Pearson Chi-Square (1, 19) = 4.866, $p < .027$, Cramer's V = .506, and more "all" responses to the biconditional task (0 vs. 33%), Pearson Chi-Square (1, 19) = 2.956, $p < .086$, Cramer's V = .394). Although the 55% rate of correct responses of the Mensa members not familiar with logic is still below the 70-75% rate of easy-to-resolve thematic tasks, Fiddick and Erlich ([28], Exp 1) have received P and not-Q selections in only 54% of cases even when the participants were explicitly instructed to search for the falsifying co-occurrence of P and not-Q in the abstract selection task. It is therefore conceivable that this is the maximum one could obtain from this task.

3 Conclusion

In general, the functioning of the updated classical logical interpretation of the conditional statement can be demonstrated by the main experimental tasks used in experimental psychology. People basically interpret the conditional statement as an equivalent relation, and with the effect of the alternative antecedents this modifies into the relationship known as the conditional. This approach can be defended from the point of view of the history of logic [6], and is normatively valid. Many researchers assume that the description of human inferences necessitates the introduction of non-monotonic logics, or that the everyday interpretation of the conditional statement is not truth-functional [30]. Still, the results presented here could be well described with a merely slightly updated classical logic. In addition, this approach can also describe the everyday interpretation of syllogisms [29]. Non-monotonic logics (e.g. default logic [31], defeasible logic [32]) are introduced with reference to the effect of a certain type of context, without, however, denoting this context. To reiterate, this seems to be a mistake, as in logic "we restricted ourselves to explicitly stated premises" ([5], p 6). Leaving the context undenoted, or for example the traditional interpretation of logical necessity and logical truth is probably the heritage of a classical logic that, in consequence of the erroneous interpretation of the conditional statement, was rigid and unable to develop, and did not allow to describe the effect of the context. When fixing this error, however, the basic effect of the context can be seen even when the equivalent relationship transforms into a conditional relationship—and this can be quite precisely described. The purpose of non-monotonic logics is also to describe such belief revisions. A similar example of the basic context is the otherwise mathematical content, which can be observed at the end of this study in the easy-to-resolve abstract selection task. The conditional statement itself is basically the same in the two experimental conditions "If the number is 4, then it is divisible by (only) two", the underlining relationships (3 or 6) are, however, different. Still, these underlying relationships can be precisely described, they do not require to introduce a specific apparatus just because the conditional statement in one of the cases evokes equivalent,

and in the other case a conditional relationship, and with the addition of further contextual components, could behave again in quite a different way. I believe that more complex contexts and even the concepts themselves behave in accordance with the same principle. Naturally, in a more complex case, we cannot predict the exact context or conceptual network in someone's mind, but without precise information we cannot predict which numbers someone is adding in his mind either. We could make only vague or probabilistic predictions, just as happens in the case of the vague or probabilistic approaches of the conditional statement. However, addition and subtraction written down on paper are still very useful tools.

In another logical approach of the field, Stenning and van Lambalgen [33], [34] worked precisely on defining the components behind the differences of such individual inferences. According to them, participants in the abstract selection task have first to define the parameters, and the differently chosen parameters produce the many different answers, all of which are correct within the given parameters. The authors themselves note that the parameters discussed by them are difficult to demonstrate experimentally, and they assume that further parameters could be discovered. In this respect, this study also defines such parameters, with markedly significant results, such as, for example, the basic equivalent inference, the avoidance of the selection of all cards and the effect of the alternative antecedent. And, of course, the whole literature investigating the relationship between logic and everyday inferences can be interpreted as the search for and testing of such parameters—components that influence how people resolve the tasks. According to this study, however, the many different answers appearing in the abstract selection task are merely artefacts resulting from the avoidance of the equivalent inferences. The same many different answers making altogether the preference of P and Q cards also appear in the evidently biconditional ball-light selection task already mentioned in this study [15]. It is, however, evident that the equivalent inference is the only correct solution in this thematic task. So, in the selection task, the search for the parameters that follows the rejection of the correct equivalent responses does not necessary reveal much about the basics of the inferential processes. Still, they can provide important information on how people try to resolve a situation that was made logically ambiguous. It is true that in the verbal reports presented by Stenning and van Lambalgen participants do not speak about avoiding the equivalent response. However, if logic has been unsure about the interpretation of the conditional statement for 2,400 years, layman participants cannot be expected to formulate a clear picture about this in the 5-10 minutes that they are given to resolve the tasks. They particularly cannot be expected to be so sure about their interpretation that, on the basis of this, they question the hidden instruction in the task, going against the equivalent responses. As a matter of fact, even the good performance on the easy-to-resolve drinking-age thematic task already mentioned in this study drops back to half (75% to 35%) by presenting only two P and two not-Q cards to the subjects, hence requiring the turning over of each of them [35].

In this study, instead of analysing the individual responses I intended to define the overall correct responses and to demonstrate empirically that people generally adhere to them. According to this approach, the greater the extent to which a task can be resolved in the same way, the more it appears easy and evident to the experimental

participants. As the rate of characteristic response drops from 100% to just 20-30%, so the task becomes more and more obscure to the participants. The more the task become obscure, the more contextual effects activate in their mind in a great variation—giving a wider variety of parameters. The most characteristic solution for a task is a sort of vote on what people believe is the correct solution in that task. This study demonstrates that this voting/belief can be equated with some logical rules, which are very simple and hence can probably also be easily programmed into a machine.

Acknowledgements. I am grateful to Professor László Bernáth for all his support during this research, for Péter Bencsik for creating Figure 1 and for the comments of three anonymous reviewers.

References

1. Copi, I. M. (1986). Introduction to logic. (7th ed.) NY: Macmillan.
2. Frege, G. (1923). Compound thoughts. In E. D. Klemke (ed.) Essays on Frege. (pp. 537-558) Urbana, Ill: University of Illinois Press.
3. Jevons, W. S. (1906). Logic. London: Macmillan Philip.
4. Aristotle. (1928). De sophisticis elenchis: The works of Aristotle, edited and translated under the supervision of W.D. Ross. Oxford: Oxford University Press.
5. Evans, J. St. B. T., & Over D. E. (2004). If. Oxford University Press.
6. Veszelka, A. (2008). A feltételes állítás ekvivalens értelmezése: hibás a feltételes állítás tradícionális absztrakciója? (The equivalent interpretation of the conditional statement: Error in the traditional abstraction?), Világosság, 48(1), 15-22. English version available.
7. Geis, M., & Zwicky, A. M. (1971). On invited inferences. Linguistic Inquiry, 2, 561-566.
8. Byrne, R. M. J. (1989). Suppressing valid inferences with conditionals. Cognition, 31, 61-83.
9. Maróth, M. (1983). Arisztotelésztől Avicennáig. [From Aristotle to Avicenna] Budapest, Hungary: Akadémiai.
10. Evans, J. St. B. T. (1996). Deciding before you think: relevance and reasoning in the selection task. British Journal of Psychology, 87(2), 223-240.
11. Johnson-Laird, P. N., & Wason, P. C. (1977). A theoretical analysis of insight into a reasoning task In P. N. Johnson-Laird & P. C. Wason (Eds.), Thinking: Readings in cognitive science. Cambridge, UK: Cambridge University Press
12. Evans, J. St. B. T., Clibbens, J., & Rood, B. (1995). Bias in conditional Inference: Implications for mental models and mental logic. Quarterly Journal of Experimental Psychology, 48A, 644-670.
13. George, C. (1992). Rules of inference in the interpretation of the conditional connective. Cahiers de Psychologie Cognitive/ Current Psychology of Cognition. 12, 115-139.
14. Wagner-Egger, P. (2007). Conditional reasoning and the Wason selection task: Biconditional interpretation instead of a reasoning bias. Thinking & Reasoning, 13(4), 484-505.
15. Veszelka, A. (2007). A feltételes állítás kísérleti vizsgálata: mégis van benne logika? (The experimental investigation of the conditional statement: Does it have logic?), Magyar Pszichológiai Szemle, 62(4), 475-488. English version available.
16. Legrenzi, P. (1970). Relations between language and reasoning about deductive rules. In Giovanni B. Flores D'Arcais & Willem J. M. Levelt (Eds.), Advances in Psycholinguistics. Amsterdam: North-Holland.

17. Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. British Journal of Psychology, 73, 407-420.
18. Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. Cognitive Psychology, 17, 391-416.
19. Sperber, D., Cara F., & Girotto, V. (1995). Relevance theory explains the selection task. Cognition 57, 31-95.
20. Oaksford, M., & Chater, N. (1998). Rationality in an uncertain world: Essay on the cognitive science of human reasoning. Hove, UK: Psychology Press.
21. Manktelow, K. J., & Over, D. E. (1990). Deontic thought and the selection task. In K. J. Gilhooly, M., Keane, R. Logie, and G. Erdos (Eds). In Lines of thought: Reflections on the referpsychology of thinking. Chichester: Wiley
22. Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with Wason selection task. Cognition, 31, 187-316.
23. Brown, W., & Moore, C. (2000). Is prospective altruist-detection an evolved solution to the adaptive problem of subtle cheating in cooperative ventures? Supportive evidence using the Wason selection task. Evolution and Human Behavior, 21, 25-37.
24. Staller, A. Sloman, S., & Ben-Zeev, T. (2000). Perspective effects in nondeontic versions of the Wason selection task. Memory and Cognition, 28, 396-405.
25. Hoch, S. J., & Tschirgi, J. E. (1983). Cue redundancy and extra logical inferences in a deductive reasoning task. Memory & Cognition, 11, 200-209.
26. Hoch, S. J., and Tschirgi, J. E. (1985). Logical knowledge and cue redundancy in deductive reasoning. Memory and Cognition,13, 435-462.
27. Platt, R. D., & Griggs, R. A. (1993). Facilitation in the abstract selection task: The effect of attentional and instructional factors. Quarterly Journal of Experimental Psychology, 46A, 591-613.
28. Fiddick, L., & Erlich, N. (2010). Giving it all away: altruism and answers to the Wason selection task. Evolution and Human Behavior, 31, 131-140
29. Veszelka, A. (in press). Köznapi következtetések és kategórikus szillogizmusok. (Everyday inferences and categorical syllogisms). Világosság. English version available.
30. Edgington, D. (2001). Conditionals. In Lou Goble (ed.). Philosophical Logic. Blackwell.
31. Reiter, R. (1980). A logic for Default reasoning. Artificial Intelligence 13, 81-132.
32. Nute, D. (2003). Defeasible logic. Lecture Notes in Computer Science, vol. 2543/2003, 151-169.
33. Stenning, K., van Lambalgen, M. (2000). Semantics as a foundation for psychology: a case study of Wason's selection task. Journal of Logic, Language and Information, 10, 273-317.
34. Stenning, K., van Lambalgen, M. (2008). Human reasoning and cognitive science. MIT Press.
35. Veszelka, A. (1999). Feltételesen plasztikus (Conditionally pliable). Master's thesis, ELTE University, Hungary