Response Patterns to a Syllogistic Categorical Reasoning Task with Abstract Groups

Joosep Olop[1] & Eve Kikas[1]

[1] School of Natural Sciences and Health, Tallinn University, Tallinn, Estonia

Correspondence: Joosep Olop, School of Natural Sciences and Health, Tallinn University, Tallinn, Estonia. E-mail: joosepol@tlu.ee

Abstract

The current study examined response patterns of young adults (N = 861) to a particular syllogism with abstract categories that contained the fallacy of the undistributed middle. Participants had to evaluate all given conclusions. Results showed that, despite being invalid, conclusions that used the word "some" were more likely to be selected as valid or possible compared to conclusions that used "all" or "none". In addition, we also analyzed participants' solutions to the task at the individual level (i.e., all evaluations to conclusions that contained the end terms). The aim was to detect dominant patterns. Results showed five dominant patterns. The significance of these findings and limitations are discussed.

Keywords: configural cluster analysis, reasoning, reasoning strategies, syllogism

1. Introduction

Reasoning is the process of perceiving stimuli from the environment or from memory and manipulating the stimuli in working memory with the goal of making conclusions about the perceived information. Numerous theories have attempted to explain the underlying mental processes of deductive reasoning. A recent overview describing the variety of theories and their predictive value was carried out by Khemlani and Johnson-Laird (2012b). Most researchers agree that three types of theories can be distinguished: heuristics, formal rules, and reasoning based on diagrams or mental models (Khemlani & Johnson-Laird, 2012b; Roberts et al., 2001). From these, three specific theories currently dominate the field: the theory of mental models (Johnson-Laird, 1984), the heuristic or probabilistic approach (Chater & Oaksford, 1999), and the verbal rules of reasoning (Rips, 1994). More recently, researchers have attempted to unify heuristics and mental models (e.g., Hattori, 2016; Khemlani & Johnson-Laird, 2012a).

On the whole, these theories have been criticized for their assumptions about fundamental underlying mental processes involved in reasoning, their exclusive nature, and their failure to recognize the importance of different strategies used during deductive reasoning. Different strategies are important as they lead to different solutions (Ford, 1995). As note\d by Roberts and Newton (2003, p.24):

"Traditionally, it has been assumed that there exists a monolithic fundamental reasoning mechanism, a device called into play whenever triggered by appropriate material [...] no cognitive processes can be identifiable as fundamental. Instead, people possess a range of strategies that can be applied to various tasks [...]"

Moreover, brain imaging studies have indicated that reasoning is not controlled by a single unitary system, as is commonly assumed by leading theories in the field (Goel, 2007).

In the context of reasoning studies of deduction, it is important, that the participants understand the task as a logic task (cf. Luria, 1976), also the participants interpret the information (i.e., propositions) in logically appropriate ways (Roberts et al., 2001). Manipulations in working memory concern the mental capacity for information processing as well as the reasoning strategies used to reach given conclusions. This capacity is related to language mediation and opens the door to knowledge about phenomena that is said to be "beyond the scope of our senses" (Cherubini et al., 2007, p. 1,496). While this capacity is a fundamental tool for human conduct (Vygotsky & Luria, 1993), it is imperfect. The study of human reasoning in the context of deduction and logical inference has shown that our reasoning is fragile, drawn by biases and directed by heuristics (Calvillo et al., 2020; Wilkins, 1929).

This study aimed to expand current knowledge in the field of psychology of reasoning. Syllogistic reasoning is

typically studied using a production or evaluation task. A conclusion evaluation task is constructed by asking participants to evaluate pairs of premises and all possible conclusions. Data is aggregated and percentages are reported. In contrast with prior studies, we reported dominant individual solution patterns to a particular syllogistic task. These correspond with current theories in the field, but also lend themselves to multiple interpretations and point toward a need to understand individual differences in strategy use.

We also aimed to investigate a highly specific population: Estonian final-year high school and vocational school students who were in the process of transitioning from formal education to working life or higher education. In educational settings, logical thinking is a knowledge domain that needs to be learned but is not directly tied to learning in other domains (Leighton, 2006). Logic is an epistemological strategy that can help detect reasoning quality, including the understanding of whether conclusions are necessarily true, impossible, or possible. In the current study, participants never specifically studied logic in their curriculum but were used to solving abstract problems in mathematics, physics, and biology. Estonian basic school students have shown high academic skills in international comparisons (e.g., OECD, 2019). Formal education can teach students superficial thinking and problem solving (Wertheimer, 2020). With poor reasoning skills, learned material can be used inadequately (Wertheimer, 2020), and the recall of information can be distorted by individual experience (cf. Bartlett, 1932).   In other words, academic skills are getting better, but at the same time some important aspects of education might not be developed. Learning logic might be crucial in education, because while logical reasoning can be a powerful tool, it can also be a source of error. Our aim was to investigate how students use reasoning in a logical task that requires deliberation.

1.1 Syllogisms and Categorical Reasoning

Logical categorical reasoning has been classically presented as syllogisms, which date back to Aristotle and have played a crucial role in the development of western scientific thinking (Politzer, 2004). Syllogisms themselves are formalizations of arguments or types of logical arguments. Traditionally, syllogisms consist of three propositions. Two of the propositions are considered premises, and one is considered a conclusion of the argument. Conclusions which necessarily follow the premises are considered valid. That is, if the premises are true, then it is not possible for the conclusion to be false. In a classical syllogism, there are four possible types of propositions that are distinguished by a quantifier (abbreviations used in this study are designated in parentheses): universal positive "all A is B" (Aab); universal negative "no A is B" (Eab); particular positive "some A is B" (Iab); and particular negative "some A is not B" (Oab). Every proposition consists of two types of terms: 1) the middle term, or the term that appears in both premises but not in the conclusion; and 2) the end terms, or the terms that appear only in the first or second premise and in the conclusions. In a valid syllogism, a middle terms' members should be accounted for by all of its members in both premises. In contrast, if not every member is accounted for (i.e., distributed), the syllogism falls into the category of a syllogistic fallacy. In current study, we used a syllogism that contained the undistributed middle fallacy. The same type of syllogism (premise pairs Aab, Ibc) have been used in several prior studies (Table 1).

Table 1. Response percentages from prior studies to syllgism with premise pairs Aab, Ibc

Study	Aac	Eac	Iac	Oac	NVC[1]
Khemlani & Johnson-Laird (2012b)	2%	1%	70%	1%	16%
Evans et al. (1999)	13%	13%	83%	80%	-

[1] No valid conclusion.

The first known empirical study concerning syllogism as a method of research in human reasoning was carried out by Störring in 1908 (see Politzer, 2004). Störring described different strategies people used to solve syllogisms. These strategies included verbal and spatial behavior and were later rediscovered by Ford (1995) and Bucciarelli and Johnson-Laird (1999). Other studies have tried to go beyond detecting behavioral patterns and further construct theories underlying mental processes (Khemlani & Johnson-Laird, 2012b).

1.2 The Current Study

The current study aimed to detect dominant response patterns to an abstract syllogistic task. We searched for patterns that would occur more often than chance (i.e., dominant response patterns) with the goal of qualitatively explaining typical mistakes. As the tasks involved novel aspects not considered by prior research, we also aimed to compare the results with previous findings.

For the purposes of this study, we constructed a syllogistic task with abstract categories. The task contained the fallacy of the undistributed middle (for detailed description, see Methods). This particular form of syllogism was chosen as it had been used in several prior studies (see Khemlani & Johnson-Laird, 2012b). This provided existing results that could be compared against. Moreover, the chosen form of the syllogism has shown to be difficult to solve. Thus, we hoped this would evoke deliberate reasoning, rather than providing intuitive answers or solutions from memory.

The task in the present study differed from previous studies in certain respects (cf. Khemlani et al., 2012b). First, instead of asking, "What necessarily follows from the given statements?" (necessity instructions) or, "What conclusions are possible concerning the given statements?" (possibility instructions), we instead asked, "Decide if the following conclusions can be made from these statements" (ambiguous instructions). The instructions we used can be interpreted both ways; i.e., as necessity instructions and/or possibility instructions. Second, similar to Evans et al. (1999), the task was constructed in a way that every conclusion given had to be explicitly evaluated. That is, the task did not indicate how many correct conclusions there could be. Third, to reduce random answers, the "Don't know" option was included for every item. Fourth, in addition to conclusions which included end terms (A-C conclusions), two other conclusions were included: a control item, indicating if participants reasoned deliberately, and a conclusion, indicating how participants interpreted the first premise of the syllogism.

Taking into account similarities and differences between tasks used in earlier studies and the specific population that we aimed to study, the following two research questions were asked. First, do task responses differ from prior studies? We expected that there would be a significant trend toward considering the Iac conclusion valid compared with the Aac, Eac, and Oac conclusions (Chapman & Chapman, 1959; Khemlani & Johnson-Laird, 2012b). As methods differ in prior studies, we also expected that, overall, there would be a trend toward accepting conclusions that had a particularity quantifier (Evans et al., 1999). Second, how many participants give responses that are similar to results provided in prior studies, and what kinds of responses are different? These research questions were asked because the instructions were more ambiguous than those used in previous studies. We were also interested in patterns of answers; i.e., whole solutions to the task at the individual level.

2. Method

2.1 Syllogistic Task

Participants were presented with the instructions, two premises (called statements in the task), and six conclusion propositions. All aforementioned text appeared on a single page. The two premises were as follows:

Premise 1: All people in Group A also belong to Group B.

Premise 2: Some people who belong to Group B also belong to Group C.

Participants were then asked to "Decide if the following conclusions can be made from these statements." The conclusions were as follows:

1). All people who belong to Group A also belong to Group C.   (Aac)

2). Nobody who belongs to Group A belongs to Group C.   (Eac)

3). Some people who belong to Group A also belong to Group C.   (Iac)

4). Some people who belong to Group A do not belong to Group C.   (Oac)

5). There is at least one person who belongs to Group A and Group B.   (Iab)

6). All people in Group B also belong to Group A.   (Aba)

For each conclusion, there were three options: Yes, No, and Don't know. Conclusions 1–4 were about end terms (A-C relations). Conclusion 5–6 concerned deductions from Premise 1.

2.2 Participants and Procedure

The syllogistic task was part of a larger test that included other tasks and questionaires. The sample was comprised of participants from a longitudinal study that followed students from kindergarten until Grade 9 (i.e., the end of basic school; for more details, see Kikas et al., 2020). Additionally, new classmates of prior participants were also asked to participate. In total, 54 schools across Estonia participated in the study. In Grade 9, parents and students were invited to take part in a follow-up study. Three years later, a follow-up study was designed with the aim of examining students' academic skills (math and language), motivation, burnout, reasoning skills, and problem solving skills. In order to follow up with the same students, the researchers used an education database to search for

the students' contact addresses and educational institutions. This was done in partnership with the Estonian Ministry of Education and Research. Principals of the schools were contacted, the aims of the study were explained, and principals were asked either to give invitation letters to students or allow researchers to come and carry out the questionnaire in ordinary classroom lessons. Invitation letters explained the aims of the study, stated that participation was voluntary, informed students that they could get feedback about their results if they wanted, and provided a contact email address. Only students who agreed to participate were included in the study.

In total, 907 students participated in the study, and 860 students completed the task (494 female; mean age = 17.1 years; SD = 1.5 years). Of these, 387 students completed the questionnaire independently, and 473 students completed the questionnaire in a classroom under the supervision of a research assistant. In cases when students filled out the questionnaire in school, the classmates of prior participants were also invited to participate. In total, 495 new students filled out the questionnaires. Questionnaires were carried out in 29 high schools and 10 vocational schools.

2.3 Data Analysis Strategy

To answer the first research question, we calculated the percentages of responses at the item level. To answer the second research question, patterns for each individual were formed. A Configural Cluster Analysis (CCA), also called zero-order Configural Frequency Analysis, was conducted to identify solutions that were observed more often than chance (Stemmler, 2020; von Eye, 1990). In CCA, patterns are considered uniform and no main effects or interactions are considered between variables/items. Expected values were calculated by expected = N/T, where T was the number of all possible solutions, and N was number of participants. Expected and observed values were compared with z approximation of χ^2 test. As multiple tests were performed, Bonferroni alpha adjustment was used to determine significant patterns (Stemmler, 2020). Significant patterns were further analyzed qualitatively.

3. Results

3.1 Descriptive Statistics across Separate Items

Percentages of each task response are provided in Table 2. There was a large difference between responses about universal-type conclusions and responses about particular-type conclusions. The majority of participants gave logically valid responses to universal type conclusions (i.e, "All people who belong to Group A also belong to Group C" (Aac; 84%); "No one from Group A belongs to Group C" (Eac; 79%)). In contrast, the majority of respondents gave logically invalid responses to particular-type conclusions (i.e., "Some people who belong to Group A also belong to Group C" (Iac; 80%); "Some people who belong to Group A do not belong to Group C" (Oac; 78%)).

The syllogistic task also contained two conclusions based on the interpretation and manipulation of the first premise (i.e., "All people who belong to Group A also belong to Group B"). The majority of participants (90%) responded correctly that "There is at least one person who belongs to Group A and Group B" (Iab). More than half of participants (65%) provided a logically invalid response: "All people in Group B also belong to Group A."

Table 2. Percentages of different responses to the syllogism task

Responses[1]	Aac	Eac	Iac	Oac	Iab	Aba
No	**84.65%**	**79.42%**	**10.23%**	**13.6%**	5%	**26.98%**
Yes	9.3%	9.88%	80.7%	78.84%	**90.47%**	65.46%
Don't know	6.05%	10.7%	9.07%	7.56%	4.53%	7.56%

Note. Abbreviations are as follows: Aac = all A are C; Eac = no A is C; Oac = Some A is C; Iab = Some A is not C; Iab = Some A is B; Aba = all B is A. 1The task asked if certain conclusions could be made from the given premises, and the choices for each conclusion were "Yes," "No," or "Don't know." Logically valid responses are in bold.

3.2 Individual-Level Dominant Response Patterns

A CCA was conducted to identify dominant response patterns (Table 3). We were interested in response patterns to the conclusions about the relationship between Groups A and C. Participants who gave an invalid response to the control item were excluded, and the item Aba was not included. In total, the analysis was conducted using 818 participants' responses. From 81 theoretically possible patterns, 45 patterns were observed. CCA indicated five response patterns as clusters — i.e., patterns that occurred more often than chance. The significant patterns were

given by 78.2% of the participants.

Table 3. CCA clusters

No	Aac	Eac	Iac	Oac	Observed	Expected	(z)Chi[1]	p.(z)Chi
Pattern 1	no	no	yes	no	35	10.086	7.845	<.0001
Pattern 2	no	no	yes	yes	510	10.086	157.408	<.0001
Pattern 3	no	yes	no	yes	34	10.086	7.530	<.0001
Pattern 4	yes	no	yes	no	35	10.086	7.845	<.0001
Pattern 5	–	–	–	–	26	10.086	5.011	<.0001

Note. Cluster based on: observed > expected and p.(z)Chi < Bonferroni adjusted alfa (<.0006). Abbreviations are as follows: Aac = all A are C; Eac = no A is C; Oac = Some A is C; Iab = Some A is not C; Iab = Some A is B; Aba = all B is A.[1]All local tests df =1. Pattern values: "no" = conclusion cannot be made; "yes" = conclusion can be made; "–" = do not know.

4. Discussion

4.1 Separate Items

When looking at item-level statistics (Table 2), we detected robust similarities and notable differences between our results and those from prior studies (see Khemlani & Johnson-Laird, 2012b). Similar to previous studies, participants tended to believe that universal-type statements (Aac, Eac) could not be made from the given premises. Participants also tended to make the Iac conclusion, which represented another similarity with prior research that was predicted by leading theories in deductive reasoning (i.e., the theory of mental models and the probabilistic approach; e.g., Copeland, 2006; Khemlani & Johnson-Laird, 2012b). Unlike the cited prior research, the majority of participants in this study made the Oac conclusion. This result was also obtained by Evans et al. (1999), wherein participants had to evaluate all given conclusions.

Our study also included inferences about the first premise of the syllogism. When given the premise "All A is B," a small group of participants decided that "At least one A is B" cannot be true. Here, we can suggest two possible explanations. First, these may have been participants who randomly assigned answers to the task without thought or consideration. Second, the task may not have been interpreted as a logic task. In that case, statements may have been evaluated simply based on whether they sounded truthful (i.e., the Gricean interpretation of propositions; cf. Begg & Harris, 1982; Roberts et al., 2001). Therefore, if the first premise is the type "All A is B," then in everyday communication situations, it might seem incorrect to say "At least one A is B."

Concerning the interpretation of the first premise, we included the conclusion type "All B is A." Prior studies have shown that syllogistic reasoning often includes these kinds of interpretations (Chapman & Chapman, 1959; Evans et al., 1999). The majority of participants made this inference. As our study did not show in which order participants made their decisions, we cannot be certain if interpretation of the first premise significantly impacted reasoning about other conclusions. If other conclusions were influenced, the interpretation of the premise changed the task's logical structure. If Groups A and B are identical, then the Iac conclusion is valid. The latter was chosen by the majority of participants.

4.2 Dominant Response Patterns

The following section analyzes clusters from Table 3. We only considered conclusions about the relationship between Groups A and C since these conclusions are typically used in studies regarding syllogisms. Clusters are compared to results from prior studies.

4.2.1 Pattern 1

This cluster included 35 participants who correctly determined that Aac, Eac, and Oac cannot be made from the given premises. A typical mistake in this cluster assumed that Iac can be deduced from the premises. This pattern has been observed in prior studies which asked, "Decide what follows from the premises necessarily" (Khemlani & Johnson-Laird, 2012b; Roberts et al., 2001).

Two specific aspects characterize this pattern. First, respondents decided that only one of the A-C conclusions follows the premises. Second, respondents deduced that the Iac conclusion is correct. This response is well predicted by different theories (e.g., mental models, the probabilistic approach; e.g., Copeland, 2006). It is worth

noting that this type of pattern is usually provided when participants are asked, "What follows from the premises necessarily?" Although this was a significant pattern, only a minority of participants solved the task spontaneously in this manner. Because of methodological differences, this pattern is likely observed in studies where participants are instructed to evaluate every given conclusion (e.g., Evans et al., 1999), but where data is aggregated and particular patterns of individuals are not reported or analyzed.

4.2.2 Pattern 2

This was the dominant cluster that occurred most often (n = 509). These students deduced that universal statements could not be made from the premises, but that particular conclusions could be made. This particular pattern was predicted by its corresponding hypothesis (Wetherick & Gilhooly, 1995). However, this does not necessarily mean that participants solely relied on heuristic strategies without any deliberation. This pattern could also be derived by visualizing the premises in a certain way.

For example, if A is considered a subgroup of B, and if B intersects with C, then the possible answer depends on where A is positioned visually. In this visualization, A has more than one logically correct position. Therefore, no valid conclusion can be drawn and the "Don't know" response could be appropriate. However, these conclusions were made by very few participants. It is possible that some of the students positioned A in such a manner that it intersected with C; as a result, these students may have inferred that Iac and Oac are correct conclusions.

These results can also be tied to the task's instructions. If participants understood the task instructions to be asking what is possible, but not excluded, then participants may have failed to notice that the first two premises could also be made.

4.2.3 Pattern 3 and 4

In Clusters 3 (n = 34) and 4 (n = 35), evaluations of the first premise were important as these evaluations were likely included in the deliberation process. These patterns were similar in that they could be explained by the same mental processes of making inferences from superficial characteristics of the task. It is not uncommon to observe a small group of participants providing similar answers to this type of syllogism (Khemlani & Johnson-Laird, 2012b). The specificity of the task construction in this study shows that, although participants gave superficial solutions, such solutions are not internally illogical because participants fail to notice the initial premises. Thus, different interpretations of the premise may explain these response patterns.

4.2.2 Pattern 5

This pattern only included "Don't know" answers (n = 26). These results may be explained in two ways. First, the task may have been refuted. In other words, it is possible that no reasoning — and particularly, no deliberate logical reasoning took place. On the other hand, since this group also included participants who answered the control question correctly, these participants may have formed different interpretations of inference-making. Adding the "Don't know" option to the task made it possible for participants not to make a conclusion because there was no information in the premises, and thus, participants did not know.

5. Limitations and Conclusions

Some limitations of this study must be discussed. First, participants solved the task in different situations, thus some students may have been motivated to provide a full logical analysis, while others may have settled for the first solution considered. Second, as pattern formulations can be interpreted multiple ways, we cannot deduce what strategies were actually used for solving the task. In addition, questions like, "Did they try out different strategies?" and "How did they choose the one that gave them the solution?" cannot be answered.

The process of making conclusions — or, as in current study, evaluating conclusions against given information — can be a difficult task for young adults. We observed a wide variety of solutions to the task. Clusters included 78.2% of the solutions given by participants, but not all the solutions. In some sense, this reflects the numerous theories in the field. Variability may be a result of different interpretations of the task, premises, and conclusions. Variability could also be related to different strategies used or information processing errors. It is clear that evaluating responses cannot be completed via a single normative framework. In addition, the variety of responses indicates different interpretations of the task and propositions.

Despite this variability, configural cluster analysis demonstrated significant patterns that occurred more often than chance. These patterns show that there is more than one dominant response that young adults tend to give, and some of these differ from those used to support current theories in the field. However, as noted, the descriptions of these patterns lend themselves to multiple interpretations and explanations. It is not clear where heuristic processes begin or where logical thinking originates. Our findings may also be due to low motivation resulting from the

research situation. However, these results mirror the distribution of solutions from situations where students were taking the test — not in a psychology laboratory, but in regular schools or at home.

Our study suggests that reasoning is fragile and influenced by numerous variables: how a task is constructed, how propositions are interpreted, and what information processing issues may occur. Knowledge of these requires further research to systematically understand how strategies of thinking develop (see Lemaire & Fabre, 2005).

References

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology* (pp. xix, 317). Cambridge University Press.

Begg, I., & Harris, G. (1982). On the interpretation of syllogisms. *Journal of Verbal Learning and Verbal Behavior, 21*(5), 595-620. https://doi.org/10.1016/S0022-5371(82)90800-3

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in Syllogistic Reasoning. *Cognitive Science, 23*(3), 247-303. https://doi.org/10.1207/s15516709cog2303_1

Calvillo, D. P., Swan, A. B., & Rutchick, A. M. (2020). Ideological belief bias with political syllogisms. *Thinking & Reasoning, 26*(2), 291-310. https://doi.org/10.1080/13546783.2019.1688188

Chapman, L. J., & Chapman, J. P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology, 58*(3), 220-226. https://doi.org/10.1037/h0041961

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, 38*(2), 191-258. https://doi.org/10.1006/cogp.1998.0696

Cherubini, P., Mazzocco, A., & Minelli, S. (2007). Facilitation and inhibition caused by the orienting of attention in propositional reasoning tasks. *The Quarterly Journal of Experimental Psychology, 60*(11), 1496-1523. https://doi.org/10.1080/17470210601066103

Copeland, D. (2006). Theories of categorical reasoning and extended syllogisms. *Thinking and Reasoning, 12*, 379-412. https://doi.org/10.1080/13546780500384772

Evans, J., Handley, S., Harper, C., & Johnson-Laird, P. (1999). Reasoning about Necessity and Possibility: A Test of the Mental Model Theory of Deduction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1495-1513. https://doi.org/10.1037/0278-7393.25.6.1495

Ford, M. (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition, 54*(1), 1-71. https://doi.org/10.1016/0010-0277(94)00625-U

Goel, V. (2007). Anatomy of deductive reasoning. *Trends in Cognitive Sciences, 11*, 435-441. https://doi.org/10.1016/j.tics.2007.09.003

Hattori, M. (2016). Probabilistic representation in syllogistic reasoning: A theory to integrate mental models and heuristics. Cognition, 157, 296–320. https://doi.org/10.1016/j.cognition.2016.09.009

Khemlani, S., & Johnson-Laird, P. (2012a). The processes of inference. *Argument & Computation, 4*, 1-17. https://doi.org/10.1080/19462166.2012.674060

Khemlani, S., & Johnson-Laird, P. (2012b). Theories of the syllogism: A meta-analysis. *Psychological Bulletin, 138*, 427- 457. https://doi.org/10.1037/a0026841

Kikas, E., Mädamürk, K., & Palu, A. (2020). What role do comprehension-oriented learning strategies have in solving math calculation and word problems at the end of middle school? *British Journal of Educational Psychology, 90, 105-123.* https://doi.org/10.1111/bjep.12308

Leighton, J. P. (2006). Teaching and assessing deductive reasoning skills. *Journal of Experimental Education, 74*(2), 109-136. https://doi.org/10.3200/JEXE.74.2.107-136

Lemaire, P., & Fabre, L. (2005). Strategic aspects of human cognition: Implications for understanding human reasoning. In *Methods of thought: Individual differences in reasoning strategies* (pp. 11-29). Psychology Press.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations* (Trans M. Lopez-Morillas & L. Solotaroff) (pp. xvi, 175). Harvard U Press.

OECD. (2019). *PISA 2018 results (volume I): What students know and can do*. OECD Publishing. Retrieved from https://www.oecd-ilibrary.org/docserver/5f07c754-en.pdf?expires=1591103493&id=id&accname=guest&checksum=9E86E2875A123B34B0E8843B7DB1B8A5

Politzer. (2004). Some precursors of current theories of syllogistic reasoning. In *Psychology of Reasoning* (pp. 223-250). Psychology Press. https://doi.org/10.4324/9780203506936-11

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking* (pp. xiii, 449). The MIT Press.

Roberts, M. J., Newstead, S. E., & Griggs, R. A. (2001). Quantifier interpretation and syllogistic reasoning. *Thinking & Reasoning, 7*(2), 173-204. https://doi.org/10.1080/13546780143000008

Roberts, M. J., & Newton, E. J. (2003). Individual Differences in the Development of Reasoning Strategies. In *Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making* (pp. 23-43). John Wiley & Sons, Ltd. https://doi.org/10.1002/047001332X.ch2

Stemmler, M. (2020). *Person-Centered Methods: Configural Frequency Analysis (CFA) and Other Methods for the Analysis of Contingency Tables* (2nd ed.). Springer International Publishing. https://doi.org/10.1007/978-3-030-49421-6

von Eye, A. (1990). *Introduction to configural frequency analysis: The search for types and antitypes in cross-classifications* (pp. xiv, 270). Cambridge University Press.

Vygotsky, L. S., & Luria, A. R. (1993). *Studies on the history of behavior: Ape, primitive, and child* (pp. xvi, 246). Lawrence Erlbaum Associates, Inc.

Wertheimer, M. (2020). *Max Wertheimer Productive Thinking* (V. Sarris, Ed.). Birkhäuser Basel. https://doi.org/10.1007/978-3-030-36063-4

Wetherick, N. E., & Gilhooly, K. J. (1995). "Atmosphere," matching, and logic in syllogistic reasoning. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, *14*(3), 169-178. https://doi.org/10.1007/BF02686906

Wilkins, M. C. (1929). The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology*, *102*, 83-83.

Copyrights