Kaja Damnjanović[1, 2], Sandra Ilić[1]

[1]University of Belgrade, Laboratory for Experimental Psychology, Department of Psychology, Faculty of Philosophy, Čika Ljubina 18-20, 11000 Belgrade, Serbia

[2]University of Belgrade, Institute of Philosophy, Faculty of Philosophy, Čika Ljubina 18-20, 11000 Belgrade, Serbia

Correspondence:

Kaja Damnjanović

ORCID: 0000-0002-9254-1263

E-mail: kdamnjan@f.bg.ac.rs

Sandra Ilić

ORCID: 0000-0002-5145-0197

E-mail: sandra.ilic@f.bg.ac.rs

Conflict of interest statement

Data availability statement

Ethics approval statement

Funding

Belief inhibition during thinking: not so fast

Abstract

The present study is a conceptual replication of De Neys & Franssens (2009) study about the role of belief inhibition in reasoning, operationalized as the change in reaction times to different categories of words presented after syllogistic reasoning task. As in the original study, we examined the accessibility of cued beliefs after syllogistic reasoning, by presenting participants (N=145) with incongruent (heuristic and normatively correct answer differ) and congruent categorical syllogisms, and lexical decision tasks comprising cued and unrelated words, and imposed methodological restrictions within the original procedure. Mean RT was overall shorter to cued than to unrelated words, and for all combinations of both syllogism congruency and response accuracy on the preceding syllogism, indicating that the full neglect of content is not necessary for correct evaluation of logical status. We registered shorter RTs for words cued by incongruent syllogisms after correct than after incorrect evaluation, which indicates that participants actively process the content of the syllogisim while reasoning, as a form of cognitive control. The successfully conducted Type 2 reasoning enhances lexical access to the cued content, rather than impairing it. In short, findings of the original study were replicated for the priming effects, but not for the inhibition of content.

Keywords: syllogistic reasoning, belief bias, inhibition, lexical access paradigm, dual processes

Introduction

Erroneous answers in syllogistic reasoning tasks in which participants judge the logical validities of categorical syllogisms occur due to the less pronounced logical skills, but also due to the intuitive reasoning that is assumed to take place thus hindering a reasoner from engaging in deliberate logical thinking (De Neys & Franssens, 2009; Bajšanski & Žauhar, 2019; Teovanović, 2019). Drawing or evaluating logical conclusions is not content-independent. When presented with all three elements of a categorical syllogism (two premises and conclusion), and with the instruction to determine whether the conclusion is valid, people should discard the content and beliefs about content and focus only on the logical structure. Concurrently, in problem-solving situations, people usually have some kind of intuitive idea of what the correct answer is, and this idea is supported by the person's knowledge and beliefs about the content of the task. Focusing on content and the believability of the conclusion of syllogism can lead to the erroneous evaluation of logical validity and this phenomenon is dubbed belief bias (Evans, Barston, & Pollard, 1983).

To reach a valid solution, we have to put in some extra cognitive effort, firstly to inhibit our intuitive incorrect answer, and secondly to engage in the slower and cognitively more expensive analytic process, as proposed by the default-interventionist dual-process account (Evans & Stanovich, 2013; Stanovich, 2009). This traditional notion of the duality of cognitive processing is probably most precisely embodied in the experimental practice of using simple tasks to which the great majority of participants' responses fall in one of the two distinctive categories: the typical erroneous ones, and the correct ones. Experimental yielding

of these two types of responses testifies to the dual-process approaches to human reasoning, in which two types of processing algorithms are presumed and conflicted: the encapsulated heuristic (Type 1) which primes intuitive, normatively incorrect, response and the analytical (Type 2) which conveys normatively correct response (Evans, 1989; Stanovich, 2009; Wason & Evans, 1974; Kahneman, 2011; Kahneman & Frederick, 2002; Sloman, 1996; Stanovich & West, 2000; Stanovich, 2018). Computation of the normatively rational responses, in terms of processing, relies on the detection that overriding miserly Type 1 processing is necessary, while the failure of such detection is a processing defect (Stanovich, 2018). The more recent iterations of the DPT, however, propose continuous processing (Melnikoff & Bargh, 2018), and the existence of multiple heuristic processes, which may be both logical and intuitive (Bago & De Neys, 2017; De Neys, 2012; Evans & Stanovich; 2013; Pennycook, Fugelsang & Koehler, 2015; De Neys, Thompson & Pennycook, 2018).

Tasks in which participants judge the logical validities of conclusions of categorical syllogisms convey the described sequence of processing. Such tasks are one of the most commonly employed dual response tasks in the field of higher cognition (Evans, 2003; Evans, 2008). They comprise three categorical propositions: two premises and a conclusion, and participants are asked if the conclusion indeed logically follows from the premises, or not. Categorical syllogisms can be formed as conflict and non-conflict problems (De Neys & Franssens, 2009). In the non-conflict versions, the believability of the conclusion and the logical validity of the conclusion are congruent, while in the conflict versions they are purposefully incongruent, thus simulating a hostile environment in which reasoning can take place. Human reasoning is always immersed and hence interrelated with the surrounding. Two types of surrounding conditions in which our reasoning takes place are benign and hostile environments (Chater, Felin, Funder, Koenderink, Krueger, Noble, Nordli, Oaksford, Schwartz, Stanovich & Todd, 2018). A benign environment is an environment that contains

useful cues, such as beliefs congruent with the logical structure, thus making the task of evaluating the conclusion easy, even for Type 1 processing. A hostile environment for Type 1 processing is one in which none of the available cues are useful, or are even misguiding, causing the substitution of an attribute only weakly correlated with the true target, that is substituting the logical status of the conclusion with its believability (Evans & Stanovich, 2013). The incongruent tasks are designed to simulate a hostile environment, and thus trigger and allow assessment of the miserly cognitive processing.

Sound reasoning in such a hostile environment requires *inhibiting belief* about content, or the content itself, and engaging in analytical processing which enhances the probability of reaching the correct answer. Inhibition, as one of the props of cognitive control, refers to the ability to ignore information or responses that are irrelevant to the task at hand (Gilmore, Göbel, & Inglis, 2018), and as one of the executive functions refers to the propensity to deliberately override dominant, automatic, or prepotent responses (Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000; Myake & Friedman, 2012). However, cognitive control is a costly process. Being the cognitive misers, people are tempted to respond based on the believability of the conclusion, rather than its logical validity (De Neys & Franssens, 2009; Handley, Newstead, & Trippas, 2011; Pennycook, Fugelsang, & Koehler, 2015). In the non-conflict tasks, miserly and normative approaches cue the same response, thus inhibition is not necessary.

One of the first studies in which the role of the belief inhibition in the evaluation of the syllogistic conclusion was investigated was conducted by DeNeys & Franssens (2009). The study rationale was based on findings in the field of memory research that show that when people purposely neglect information, access to that information will subsequently be distorted, and such temporary inaccessibility of information refers to the concept of inhibition

(De Neys & Franssens, 2009). In line with this, authors explored the nature of inhibition failure and the resulting erroneous responses by testing the accessibility of cued beliefs after participants' analytic and miserly reasoning. For this purpose, they used sequences comprising two tasks. In each sequence, the first task was to evaluate the logical status of a conclusion in a conflict (validity and believability incongruent) or a non-conflict (validity and believability congruent) categorical syllogism (ET: evaluation task). The second task in the sequence, which was administered immediately after the ET, was the lexical decision task (LDT). In LDTs participants are presented with letter strings and instructed to indicate whether the presented letter string is a word or not by pressing designated buttons. The 24 stimuli used in the LDT comprised 12 pseudo-words and 12 words. Half of the words were *target* words - core words used in the preceding ET and words closely semantically related to core words, while the other half were words that were unrelated to the six target words. This sequence, the ET+LDT, was repeated 8 times per participant.

The results of this study showed that RTs for lexical decisions on *target* words were longer after evaluating conflict syllogisms (i.e., syllogisms to which normatively correct and belief based but erroneous responses were different), compared to no-conflict syllogisms, while RTs to unrelated words were statistically the same regardless of the type of preceding syllogism. Because only the RTs to target words were affected (longer after conflict syllogism compared to no-conflict), the authors concluded that the distortion of the memory access to words wasn't general, and was due to inhibition of beliefs. In other words, when beliefs cued a response consistent with the logical status of the syllogism, additional cognitive control, that is – inhibition of beliefs, was not required, and RTs for target words in those cases were significantly shorter than in conflict tasks. All participants, regardless of the level of their overall achievement on syllogistic reasoning tasks, displayed the memory distortion after solving *conflict* problems, which suggests that even „the poorest reasoners"

were engaged in fighting the biasing beliefs. Based on that, the authors postulated the occurrence or initiation of the process of belief inhibition in reasoning in all participants, and implied that an erroneous response resulted from a failure to complete the inhibition process, and not from a failure to initiate inhibition (De Neys & Franssens, 2009). Using the same lexical access paradigm, other research confirmed that processing belief–logic conflicts involve effortful belief inhibition (Svedholm-Häkkinen, 2015). Further research on these topics confirmed that distinct heuristic and analytic processing systems underpin reasoning in belief-bias tasks (Stupple, Ball, Evans, & Kamal-Smith, 2011). Findings were generalized from the belief bias paradigm to matching bias tasks (Stupple, Ball, & Ellis, 2013), and it was confirmed that beliefs, indeed, automatically influence reasoning and that ignoring them comes with an attentional cost (Barton, Fugelsang & Smilek, 2009).

Some particular features of the tasks employed in the study conducted by DeNeys and Franssens (2009) could have, to some extent, shaped the findings, and this pertains to both reasoning as a complex cognitive process required for sound judging of the validity of a deducted conclusion, and reasoning as a simpler cognitive process such as deciding if the string of letters is a word of a certain language. First, the authors used different items in the conflict and no-conflict conditions, each with different content. each of the four possible types of syllogistic reasoning tasks (2 conflict, and 2 non-conflict) was represented by two tasks. As those tasks differed by the content, there were altogether 8 tasks, all heterogeneous in terms of the content (themes) that were expected to be discarded in the experimental procedure. Human reasoning is, however, not content independent – i.e., the content, which is to be discarded in syllogisms used to register belief bias, influences task performance, and this notion is supported by voluminous empirical documentation (for a comprehensive review see e.g. Casadio, 2016; Davies, Fetzer & Foster, 1995; Gigerenzer & Hug, 1992, Stanovich, 2018). Second, in the original study, words used in LDT were not controlled for the factors

known to influence RTs within the lexical decision paradigm - word length (Balota, 1994; Hudson & Bergman 1985), word type (Kostić & Katz, 1987; Tyler, Bright, Fletcher & Stamatakis, 2004), and their relative frequency (Gardner, Lapan, & Lafferty, 1987; Brysbaert, Mandera, & Keuleers, 2017). Indeed, the authors established in a pilot study that there were no a priori lexical decision time differences for the two different sets of (only) *target* words presented after *conflict and no conflict* syllogisms. However, it remains possible that the differences in RTs registered at the level of the word type[1] factor (2: target, unrelated), that is – the fact that RTs to target words were shorter than RTs to unrelated words, was at least partly due to accidental systematical differences in length, frequency, and type of words between words within each set, and not solely because of identity or semantic priming (due to exposure to same or semantically related words in the preceding syllogism) that was reduced by belief inhibition in the case of LDT after conflict syllogism.

Our rationale was to test whether DeNeys & Franssens' finding of memory impairment after logical reasoning, on which they (partially) based the notion that the belief inhibition always occurs (but isn't always completed successfully), would still hold after implementing methodological improvements into the original design. For that purpose, we've run a conceptual replication where contents of the syllogisms and characteristics of words were directly matched and controlled. As in the original study, the design of our study is grounded in the relative complexity methodology, in which time is simply used as a measure of how many steps information processing took (Pylyshyn, 1999). More precisely, this conveys the idea that, next to the general priming effects of the mere presentation of reasoning problem containing target words, the process of inhibiting the content will affect access to target words in the subsequent LDT task, which can be registered as prolonged RTs on those words.

---

[1] Please note that word type factor pertains to manipulation of words by whether they are the same or semantically related, or unrelated to beliefs that are to be inhibited during reasoning, and not by lexical categories.

Aim and hypotheses

The present study aimed to observe the impact or cognitive cost of reasoning (ET) on subsequent simpler cognitive tasks (LDT), in a more restrictive, controlled, and parsimonious experimental design, both in terms of the content and the characteristics of stimuli. Directly, our aim was to test if the inhibition of belief is indeed always occurring in syllogistic reasoning.

- In line with this we hypothesized as De Neys and Franssens (2009), with the additional hypothesis, taking into account the possibility of the non-replicated findings:

H1a: if people indeed try to discard beliefs when solving conflict syllogisms, that is – if *everybody always engages in an inhibition process* but not all complete it successfully, RTs to lexical decisions on *cued* words should be *longer after conflict* syllogisms compared to no-conflict syllogisms, regardless of the accuracy of syllogistic reasoning (their access to target words should be distorted regardless of whether they were biased or not). The difference in RTs for *unrelated* words between conflict and no-conflict would not occur.

H1b: if people exhibit belief bias because they do not even initiate belief inhibition due to not detecting that their beliefs conflict with the syllogism validity, those who fail to solve conflict syllogisms should not show longer RTs to stimuli related to beliefs that should have been inhibited during reasoning, but those who solved all conflict syllogisms correctly should.

Method

Participants

A total of 145 first-year psychology students at the University of Belgrade, all native speakers of Serbian language who provided answers to all tasks were included in the study. Data on gender and age was not collected. All participants had normal or corrected to normal eyesight and reported no neurological impairments. Participants received course credit for participation.

Materials

Evaluation Tasks

The ETs used were the type of syllogisms employed by Sá, West, and Stanovich (1999) and Markovits and Nantel (1989), as well as by De Neys and Franssens (2009). Each categorical syllogism consisted of two premises (arguments) and a conclusion that was either valid or invalid. Moreover, each syllogism was either believable or not, meaning that its content, particularly the conclusion, either did or didn't match common beliefs. When these two binary factors - the logical status and the believability of a syllogism are crossed four types of tasks yielded: believable and valid, not believable and invalid, believable and invalid, and not believable and valid. The first two types belong to the higher category of congruent syllogisms as their believability and logical status agree. The last two types - those in which logical status and believability are conflicted, belong to the category of incongruent syllogisms. An example of an incongruent (not believable and valid) syllogism is: 'All plants are oak trees. The root is a plant. Therefore, the root is an oak tree.'

Each evaluation task, i.e., syllogism, was defined by three keywords that pertained to one of the four themes: oak, dove, school, and lemonade. Each of these themes (represented by triplets of keywords) was represented by all four types of tasks. This allowed for all four themes (triplets of keywords) to appear in all four types of syllogisms. For example, the words from the theme "oak" (oak, plants, root) appeared in believable invalid, believable

valid, not believable valid, and believable valid syllogism. This procedure gave a total of 16 syllogisms that were counterbalanced against participants using a Latin square, as each participant was presented with four different types of syllogisms of four different themes. Participants' response times to syllogisms were recorded from the moment a conclusion appeared on the screen until the answer was given.

Lexical Decision Task

In total, 96 letter strings were used, half of which were pseudo-words. From the total of 48 pseudo-words, four sets, each comprising 12 pseudo-words, were randomly assigned to each of the four ET themes. The other half of letter strings – 48 preselected words, was also divided into four sets of 12 words. The words from each of those four sets belonged to two categories: cued and unrelated words. Cued words comprised three keywords that appeared in a syllogism and defined the theme (e.g. oak, plant, root), and three words that were semantically related to those keywords (e.g. tree, trunk, leaf). Unrelated words comprised six words that were semantically unrelated to any of the words from the cued words category (e.g. money, lamp, snake, jacket, etc.). So, to each of the four themes of syllogisms, a set of 24 letter strings was assigned. Letter strings within each set were presented together in one lexical decision task. For the complete list of words see Appendix A.

All used words were nouns matched by frequency and length. To check the stimuli selection, LDT comprising the very same words and pseudo-words as used in the main experiment was administered, but without the preceding ETs. A fixation cross was presented for 300ms at the center of the screen before each word. Words were presented at the center of the screen in a randomized order. Participants' task was to indicate whether each of the presented letter strings was a word or not by pressing one of two buttons on a keyboard.

In comparison to the original study by De Neys and Franssens (2009), we've applied two new solutions to control confounding influences. First, all stimuli in LDTs were nouns

and were of the same length and frequency. Second, we have controlled the themes of reasoning tasks. Namely, we had predefined four themes, each of which was then represented in all four types of syllogisms. Counterbalancing themes and syllogism types allowed for the LDT words to be entirely counterbalanced too. This enabled direct comparison of RTs of the *very same* words in all available conditions, that is - after four different types of syllogisms, instead of comparing reaction times to different words presented in different conditions. In other aspects, we've followed the original DeNeys' and Franssens' (2009) procedure of repeating ET + LDT sequence.

Procedure

Two experimental sessions with six months time gap between were set up. In both experimental sessions, experiments were administered using the OpenSesame software 3.0.7 (Mathôt, Schreij, & Theeuwes, 2012). All stimuli were presented on a black screen and in 1366x768 screen resolution.

In the first session, all participants completed four ET+LDT sequences. Participants were familiarized with both tasks and the procedure before starting the main experiment. The data from the practice were not recorded. Each participant was presented with four types of syllogisms belonging to all four themes. Participants were instructed to a) assume that the information presented in the syllogisms was true and b) focus only on the rules of logical reasoning. In the LDT, participants responded to strings of letters by indicating on the keyboard whether they represented a real word or not. They were instructed to respond as fast as they can. The exact instructions for both tasks are provided in Appendix B.

The ETs started by presenting each of the two premises for three seconds. After six seconds both the premises and the conclusion were presented on the screen until the participant gave their response. Immediately after the response was given, a short lexical

decision task corresponding to the ET theme comprising a total of 24 letter strings began. To avoid the possibility of immediate priming within each LDT stimuli were presented in a pseudo-random order, that is – the cued word was never immediately preceded or followed by a related word. All letter strings were preceded by a fixation cross presented for 300ms at the center of the screen. After completing the LDT, participants had a short break before starting the next ET which differed from the preceding ET by congruency. For the graphical presentation of the procedure employed in each ET+LDT sequence see Figure 1.

Figure 1

The graphical representation of an ET+LDT sequence (Oak theme).
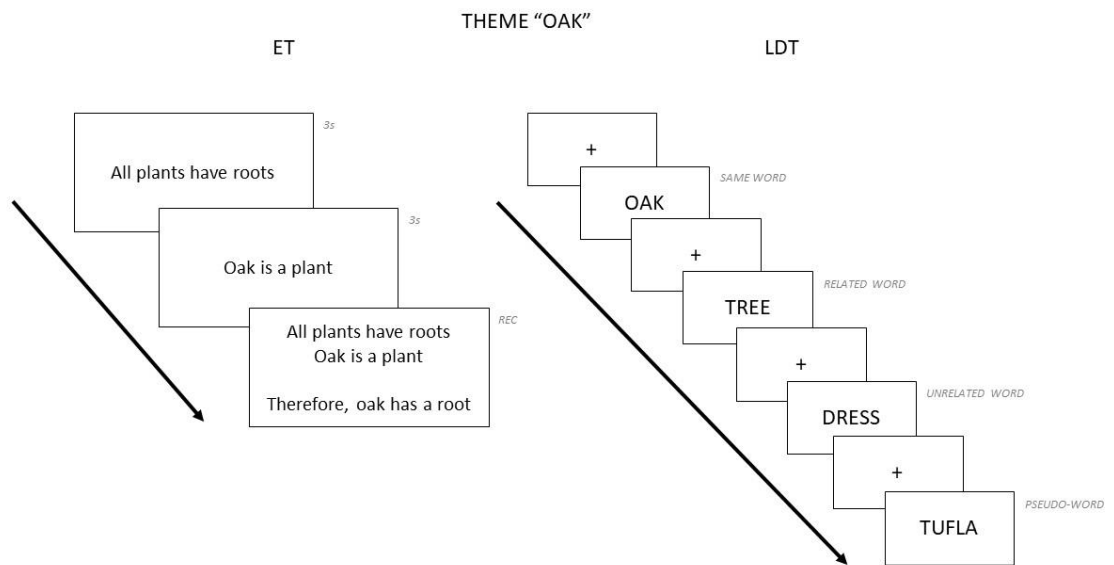
Note. Left: Evaluation Task (congruent syllogism); the labels next to the upper right corners of the boxes containing premises are presentation times of premises, and the label above the upper right corner of the third box indicates that response time was recorded from the moment of the presentation of an entire syllogism. Right: Lexical Decision Task; the labels next to the upper right corners of the boxes containing letter strings used in LDT indicate the word categories. The left to right order of graphical presentation of the two tasks and the corresponding arrows indicate the order of stimuli presentation in a sequence.

For confirmation that there were no differences in RTs for cued and unrelated words (when they were unaffected by preceding syllogistic reasoning tasks), after six months, 78 participants again completed only the LDT. The comparison of RTs showed no significant difference between cued and unrelated words ($F(1, 77) = 1.643$, $p = .204$, ηp2 $= .021$).

Design

The study employed a factorial design. For registering belief bias and potential differences in response times to different types of syllogisms, we used a 2 x 2 factorial design with syllogism believability (2: believable, not believable) and logical status (2: valid, invalid) as within-subjects factors. The dependent variables were accuracy of evaluation (categorical; 2: correct, incorrect), and evaluation task response time, recorded from the moment of conclusion display until the response was given.

For registering whether believability and logical status of evaluation tasks affected reaction times to cued and unrelated words in subsequent lexical decision task we used a 2 x 2 x 2 factorial design with *logical status* (2: valid and invalid) and *believability* (2: believable and not believable) and *word category* (2: cued and unrelated words) as within-subject factors. The dependent variable was reaction time in lexical decision tasks.

Finally, to test whether accuracy on evaluation tasks influenced reaction times in lexical decision tasks after solving different types of syllogisms, we employed 2 x 2 x 2 mixed factorial design with syllogism congruency (2: congruent, incongruent), accuracy (2: correct, incorrect) and word category (2: cued, unrelated). The dependent variable was, again, RT in LD tasks.

Results

Evaluation tasks

Overall accuracy on evaluation tasks was 63%. The percentage of overall correct answers was lower for conflict syllogisms (50%), compared to no conflict syllogisms (75%), with 29 participants who correctly solved all conflict syllogisms and 28 participants who didn't solve any conflict syllogism correctly. Conversely, the numbers of participants who correctly solved all or none no-conflict syllogisms were 74 and 1, respectively. The difference between accuracy (number of correctly solved tasks per participant) on conflict and no-conflict syllogisms was significant (F (1, 144) = 49.384, p < .001, η_p^2 = .255). Further, participants took less time to solve conflict (M=3475.36ms, SD=4109.13ms) than no-conflict (M=4075.33ms, SD=4977.71ms) syllogisms (F (1, 144) = 4.529, p = .035, η_p^2 = .030). The finer-grained analyses including logical status and believability as separate factors are in Appendices C and D.
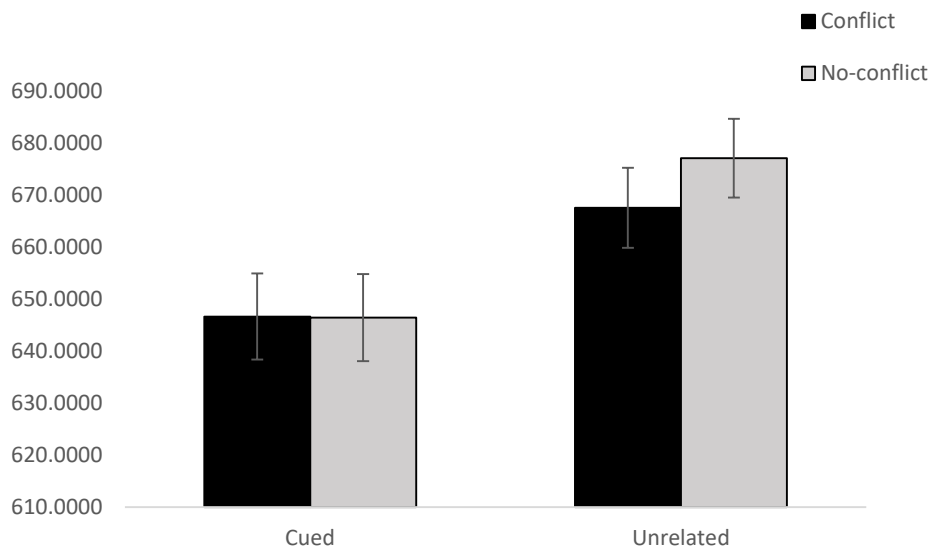
Evaluation tasks and reaction time on lexical decision tasks

To test the effects of word type and belief-logic conflict on lexical decision times we conducted a repeated-measures ANOVA and registered significant interaction between the two factors (F (1, 144) = 5.111, p = .025, η_p^2 = 0.034), as well as significant main effect of word type (F(1, 144) = 74.744, p < .001, η_p^2 = .342). As expected, due to identity and semantic priming effects because of exposure to cued words in preceding evaluation tasks, RTs to cued words were shorter (M=646.579ms) than RTs to unrelated words (M=672.357ms). We registered insignificant main effect of logic-belief conflict on lexical decision times (F(1, 144) = 2.529, p = .114, η_p^2 = .017) as participants took the same time to make lexical decisions about words after conflict and no-conflict syllogisms. Simple main effects analysis showed, as can also be seen in Figure 2, that RTs for *cued* words were statistically the same after solving (whether correctly or incorrectly) conflict and no-conflict syllogisms. The mean difference between these lexical decision times was $M_{conflict} - M_{no-conflict}$ = 0.190ms (F (1, 144) = 0.003, p = .957, η_p^2 = .000).

There was significant simple main effect of belief logic conflict on unrelated words (F (1,144) = 6.340, p = .013, η_p^2 = 0.042) – lexical decisions regarding unrelated words were longer by 9.545ms after no-conflict syllogisms. More importantly, we registered significant simple effects of word type factor on both levels of the belief-logic conflict. Participants' lexical decisions were longer for unrelated words compared to cued by mean difference of 20.910ms (F (1, 144) = 32.537, p < .001, η_p^2 = 0.184) in case of conflict syllogisms and by 30.646ms (F (1, 144) = 68.976, p < .001, η_p^2 = 0.324) in case of no-conflict syllogisms. The analysis of the three-way interaction of believability, logical status, and word type can be found in Appendix E.

Figure 2

The effect of syllogism congruency (2: conflict, no conflict) and word type (cued, unrelated)
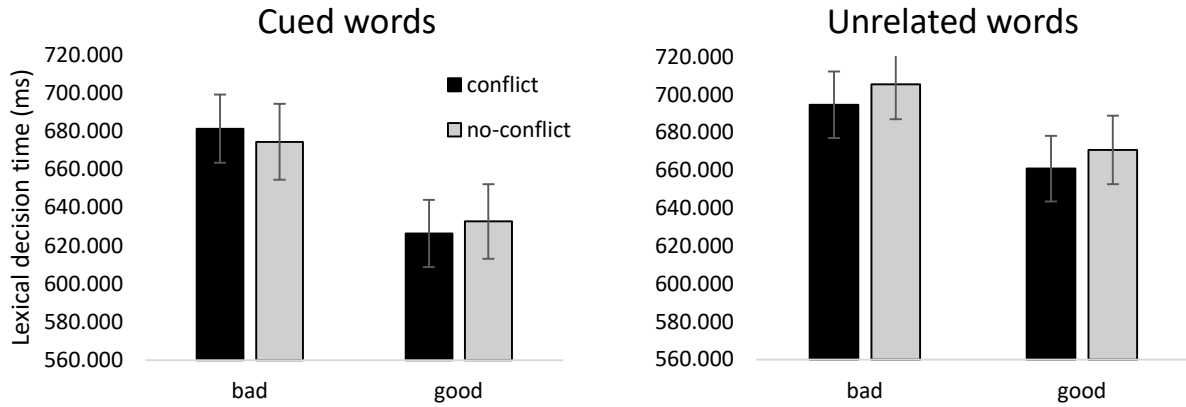
on lexical decision times

Note. Error bars are standard errors of the mean.

Accuracy on evaluation tasks and reaction time on lexical decision tasks

Next, we include the accuracy of syllogistic reasoning to analyses because the absence of the effects of belief inhibition could be due to averaging RTs to cued words after conflict and no conflict syllogisms for participants who were both biased and unbiased by beliefs. Namely, it could be the case that even if not all reasoners detect conflict and initiate inhibition, those that give unbiased validity judgements do. In that case we expect longer reaction times to cued words after solving conflict syllogisms than after solving no-conflict syllogisms, but only in good reasoners group. In bad reasoners group we expect to register only priming effects. As De Neys and Franssens (2009), we split the sample by the reasoning skill (2: good reasoners, bad reasoners). Namely, in the subsequent analysis, we included only participants who solved all conflict syllogisms correctly (good reasoners) and participants who solved all conflict syllogisms incorrectly (bad reasoners). The mixed ANOVA with within-subjects factors word type (2: cued, unrelated) and congruency (2: conflict, no-conflict), and between-subjects factor reasoning skill (2: bad reasoners, good reasoners) showed insignificant main effects of reasoning skill ($F (1, 55) = 2.796$, $p = .100$, $\eta_p^2 = .048$; $\text{Mdiff}_{\text{bad-good}} = 41.305\text{ms}$) and congruency ($F (1, 55) = 0.980$, $p = .326$, $\eta_p^2 = .018$; $\text{Mdiff}_{\text{conflict-noconflict}} = -5.001\text{ms}$). We registered significant main effect of word type ($F (1, 55) = 45.916$, $p < .001$, $\eta_p^2 = .455$; $\text{Mdiff}_{\text{cued-unrelated}} = -29.238\text{ms}$) – reactions to cued words were faster compared to unrelated words,. All two-way interactions were insignificant: word type x reasoning skill ($F (1, 55) = 2.639$, $p = .110$, $\eta_p^2 = .046$), word type x congruency ($F (1, 55) = 1.751$, $p = .191$, $\eta_p^2 = .031$), congruency x reasoning skill ($F (1, 55) = 0.367$, $p = .547$, $\eta_p^2 = .007$). Finally, the three-way interaction was also insignificant ($F (1, 55) = 0.772$, $p = .383$, $\eta_p^2 = .014$) as can be seen in Figure 3.

Figure 3

The word type, congruency, and reasoning skill interaction

Note. Error bars are standard errors of the mean

Results of post-hoc tests showed a few differences and are presented in Tables 1 through 3. First, lexical decisions about cued words were always faster than decisions about unrelated words. This difference was significant for good reasoners regardless of the syllogism congruency, while bad reasoners showed significantly shorter reaction times to cued words only after no-conflict syllogisms. After solving conflict syllogisms, bad reasoners were equally fast to make a lexical decision regardless of the word type (see Table 1). These findings indicate priming effects, which are even stronger in the "good reasoners" group. In line with this is the finding that bad reasoners in general took longer than good reasoners to make a lexical decision, and this difference reached significance in the case of words cued by conflict syllogisms (see Table 3 and Figure 3).

Table 1

Simple main effects of word type on the remaining factors

Reasoning skill	Congruency	Mean Difference (cued-unrelated)	SE	p	CI 95 Lower Bound	CI 95 Upper Bound
Bad	Conflict	-13.416	8.486	0.120	-30.422	3.589
	No-conflict	-31.039*	8.303	0.000	-47.679	-14.399
Good	Conflict	-34.469*	8.338	0.000	-51.179	-17.759
	No-conflict	-38.026*	8.159	0.000	-54.377	-21.676

Note. SE=Standard error.
*significant at p=.05 level

Table 2

Simple main effects of congruency on the remaining factors

Reasoning skill	Word type	Mean Difference (conflict-noconflict)	SE	p	CI 95 Lower Bound	CI 95 Upper Bound
Bad	Cued	6.869	8.614	0.429	-10.394	24.132
	Unrelated	-10.754	9.736	0.274	-30.266	8.759
Good	Cued	-6.280	8.464	0.461	-23.243	10.683
	Unrelated	-9.838	9.567	0.308	-29.010	9.335

Note. SE=Standard error.

Table 3

Simple main effects of reasoning skill on the remaining factors

Word type	Congruency	Mean Difference (bad-good)	SE	p	CI 95 Lower Bound	CI 95 Upper Bound
Cued	Conflict	54.890*	25.074	0.033	4.640	105.139
	No-conflict	41.740	27.842	0.140	-14.055	97.536
Unrelated	Conflict	33.837	24.700	0.176	-15.662	83.336
	No-conflict	34.753	25.831	0.184	-17.013	86.519

Note. SE=Standard error.
*significant at p=.05 level

To summarize, we registered longer reaction times to unrelated than cued words, and longer reaction times in bad reasoners compared to good reasoners, indicating priming effects, that are stronger in good reasoners group.

Discussion

When evaluating the validity of the syllogism reasoners should focus only on the validity of the conclusion, which means that they should not pay any attention to the content of that syllogism. However, not paying attention is not passive nor an automatic process.

Tossing away is not cognitively cheap, especially for cognitive misers, because human reasoning is related and even dependent on the content about which we are reasoning (e.g. Casadio, 2016; Davies, Fetzer & Foster, 1995; Gigerenzer & Hug, 1992, Stanovich, 2018…). Keeping content on hold requires cognitive effort; one should inhibit beliefs and knowledge about that content to override intuitive belief-based answers, and then initiate more demanding type 2 processing, as proposed by DeNeys and Franssens (2009). The aim of the present study was conceptual replication through repeating the procedure comprising of the reasoning task and the lexical decision task, established by DeNeys and Franssens (2009), and to test if the findings of the original study would still hold up after the implementation of more strict control into the design. We've employed two response and lexical access paradigm in repeated experimental design. We have presented participants with four types of categorical syllogisms, in 2 of which believability of the content was conflicted with the logical status, and 2 of which was not. Participants' task was to evaluate if the conclusion logically stems from premises. Each syllogistic ET was followed by the LDT comprised of the word cued by the syllogism (identical and semantically related), non-related words, and pseudo-words as control. The important additions in our repeated design were two. Firstly, there were altogether four themes of syllogisms, which means that each content (or belief) appears in every type of syllogism. This enabled each word to be primed by the four types of syllogisms without changing content, e.g. the word *oak* was primed by the believable-logical, believable-illogical, non-believable-logical, and non-believable-illogical syllogism. Second, all the words in our LDT were controlled for length, frequency, and type.

Results of our study are following findings of a large body of research on belief bias which show that evaluation of the logical validity of the incongruent syllogisms is indeed significantly harder compared to the congruent syllogisms (De Neys et al., 2011; De Neys, & Franssens, 2009; Sá et al., 1999; Stupple & Ball, 2008). Accuracy of the evaluation depended

on the logicality of the syllogism, but to a different degree depending on the believability of the content, which is in accordance with the classical notion that belief bias is more pronounced on invalid syllogisms (Evans et al., 1983). The logicality did not make it any easier to evaluate the logical status of the syllogism when the content of the identically themed syllogism was not believable (eg. *Root has oak* does not differ from *Oak has plant*). When the content is *believable*, the evaluation generally takes a smaller amount of time, and it is easier to correctly assess the logicality of the valid syllogism in comparison to the invalid ones. When the content is *not believable*, differences in the accuracy and the lasting of the evaluation process between valid and invalid syllogisms were *not* registered. The accuracy of evaluation of *believable* syllogisms was higher in comparison to the not-believable on both valid and invalid syllogisms. Even though evaluation of not believable invalid syllogism took a longer time than evaluation of believable ones, time invested did not pay off in terms of accuracy rates. These findings are in accordance with DeNeyes and Franssens' findings, and indicate that the interplay of the logicality and believability works in the following manner: positive logical status helps to reason when content is believable, but if that is not the case – then reasoning requires an additional effort, that is - the inhibition of unbelievable content.

The question was what happened when that very content, conflicted or not, was presented in the subsequent lexical decision task. As expected, every target word was identically or semantically primed by the preceding syllogism which contains those words, as to find in previous studies employing the lexical access paradigm (DeNeys & Franssens, 2009; Svedholm-Häkkinen, 2015). Overall, mean RT for cued words was significantly *shorter* than mean RT for unrelated words. RT for words cued by correct evaluations of non-conflict syllogisms and incorrect evaluations of conflict syllogisms were shorter compared to unrelated words (conflict incorrect, non-conflict correct) after those same syllogisms. However, the registered general priming effect isn't a crucial finding, what was the focus of

both original and our study was the comparison between RTs for cued and uncued words in 4 different *reasoning* conditions. We analyzed the expected inhibitory effects, within priming, registered as the difference in RTs of different types of words after a different type of syllogisms. DeNeys and Franssens (2009) findings showed that RTs of the cued words after conflict syllogisms were *longer* compared to the RTs of the cued words after non-conflict syllogisms. Based on this result, they proposed that inhibition of content is a crucial phase in the evaluation of logical validity of the deductive syllogism, initiated by all reasoners though not always completed. The difference between RTs to words cued by conflict and non-conflict syllogisms in our study was virtually zero, albeit the differences were observed between „good" and „bad" reasoners.

It is proposed that inhibitory processes are triggered by the successful detection of the conflict between content and validity of the conclusion (Stanovich, 2018), which may be observed as longer RTs of the cued words after conflict syllogism compared to the RTs of the cued words after a non-conflict syllogism, regardless of the correctness of the answers. Our findings do not support this notion, since, as stated, RTs to the cued words after conflict syllogisms were not longer than RTs to cued words after no-conflict ones. Moreover, the type of the task did not play a role in reshaping the priming effects, meaning that inhibition or the difference in inhibition was not observed. RTs of the words cued by conflict syllogisms were shorter than for unrelated after conflict syllogism, and the same pattern was observed after non-conflict syllogisms. In short, after introducing methodological restrictions into the original design, priming effects were registered, meaning that the believability of the content, nor the logicality of the conclusion were not factors, the only difference in priming effects came from the type of words.

The interpretation of our findings should take into account theoretical considerations of the nature of the inhibition process. Inhibition may indeed be treated as memory impairment,

meaning that the content of the syllogism should be put aside while judging about logicality, so the subsequent retrieval requires a longer time, as implied in the original study. Yet, inhibition, a form of cognitive control, demands cognitive resources, meaning that the content which is inhibited is being actively processed, even though shallower than computing the content. It could be that the inhibited content was actively processed, which enabled easier access to that content in subsequent LD tasks.

Non-registered effects of the type of the *task* still don't mean that different types of answers based on type 1 and type 2 processes do not influence word access, therefore we analyzed the interplay between syllogism congruency, accuracy on ET, and word type. After correct evaluation of incongruent syllogism, RTs for cued words were *shorter* than after erroneous one, but this was not the case for the same comparison within congruent syllogisms. Both findings dispute our prediction that, because of inhibition, RTs for cued words after correctly solved conflict syllogisms will be longer compared to the RTs of the same words after incorrectly solved conflict syllogisms. This pattern implies that sound evaluation of congruent and evaluation of incongruent syllogisms do differ in terms of cognitive effort, most probably due to the existence of conflict in the latter ones and detection, despite the shorter response times for incongruent ones.

Next to the role of the inhibition, the findings of our study, more specifically those regarding the good and the bad reasoners, could also be discussed in light of the more recent iterations of the dual-process theories (DPT) framework. One plausible assumption is that the good reasoners do not need to engage in inhibitory processing because they simply do not generate a belief based-response. This is supported by the findings of Svedhölm-Hakkinen (2015), based on the observed inhibitory lexical decision effects among less gifted reasoners, while the effect was not registered in a group of highly reflective reasoners. The De Neys & Franssens' median split data also indicated that the effect of belief inhibition was less

23

pronounced among the better reasoners. Findings of these three studies (De Neys & Franssens, 2009; Svedhölm-Hakkinen, 2015 and the present study) direct toward theoretical re-conceptualisation in light of fast logic. Namely, the description of te successful evaluation of the validity of a categorical syllogism was, until recently, based on the traditional dual-process theories (DPT) framework or the "basic and binary" one (Evans 2003; Evans & Over 1996; Smith & DeCoster 2000; Stanovich 1999, Kahneman, 2011). However, as previously pointed out recent iterations of the DPT propose that people can be cognitive misers *and* logicians simultaneously (De Neys, 2012, 2014, 2018; De Neys & Bonnefon, 2013; Bago & De Neys, 2017; Evans & Stanovich; 2013; Pennycook, Fugelsang & Koehler, 2015; De Neys, Thompson & Pennycook, 2018). The simulation of the interplay of reasoning and performance on subsequent simpler cognitive tasks can contribute to the understanding of the cognitive price of the *logical intuitions*. Research suggests that intuitive sensitivity to logical structure arises because logical arguments are more fluent and based on mindware (Howarth, Handley & Walsh, 2018; Klauer & Singmann, 2013; Morsanyi & Handley, 2012; Stanovich, 2018). That would mean that the reaching correct answer on conflict syllogism does not require effort and time, rather it can be the result of implicit and automatic processing. If this is the case, it could be implicated that the reason for our finding of shorter RTs for cued words after correctly evaluated conflict syllogism than after erroneous evaluation, is because there is no need to discard content to give the correct answer (De Neys & Pennycook, 2019).

The main limitation pertains to materials. Rigorous control for type, length, and frequency of words limited the number of ETs, so every participant solved only one ET per syllogism type which made individual differences approach impossible, and analyses were conducted only at the level of the experimental material. On a paradigmatic level, the binary response syllogistic reasoning tasks remove the possibility of registering different heuristic answers and allow guessing the right answer. Further studies should override these

limitations by increasing the number of ETs, and, if possible, include different, more pronounced modes of answering, and other reasoning tasks.

Conclusion. The employed sequence of evaluation and lexical decision task introduced by DeNeys and Franssens (2009) proved to be conducive in the field of experimenting within the dual-process approach. Our replication study imposed methodological restrictions which improved control in the proposed method, and findings, to some extent, shed light on the notion that the inhibition of belief in syllogistic reasoning requires cognitive effort, and, as such, even even though subtle, effects of belief inhibition is one of the markers of the logicality. To summirize, findings of the original study were replicated for the priming effects, but not for the impairment of lexical acces. Our findings suggest that, using the original terminology, the successfully conducted Type 2 reasoning enhances lexical access to the cued content, rather than impairs it, and that, at the same time, heuristic reasoning could be more cognitively pricey than commonly presumed.

References

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90-109. DOI: 10.1016/j.cognition.2016.10.014

Bajšanski, I., & Žauhar, V. (2019). The relationship between consistency and consensuality in syllogistic reasoning. *Psihologijske Teme, 28*(1), 73–91. https://doi.org/10.31820/pt.28.1.4

Balota, D. A. (1994). Visual word recognition: The journey from features to meaning. In M. A. Gernsbacher (Ed.), Handbook of psycholinguistics (p. 303–358). Academic Press.

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. Current Directions in Psychological Science, 27(1), 45-50. https://doi.org/10.1177/0963721417727521

Casadio, C. (2016). Reasoning logically in cognitive domains. Logic Journal of the IGPL, 24(4), 628-638. https://doi.org/10.1093/jigpal/jzw020

Chater, N., Felin, T., Funder, D. C., Gigerenzer, G., Koenderink, J. J., Krueger, J. I., ... & Todd, P. M. (2018). Mind, rationality, and cognition: An interdisciplinary debate. *Psychonomic Bulletin & Review*, *25*(2), 793-826. https://doi.org/10.3758/s13423-017-1333-5

Davies, P. S., Fetzer, J. H., & Foster, T. R. (1995). Logical reasoning and domain specificity. *Biology and Philosophy*, *10*(1), 1-37. https://doi.org/10.1007/bf00851985

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, *7*(1), 28-38. https://doi.org/10.1177/1745691611429354

De Neys, W., & Franssens, S. (2009). Belief inhibition during thinking: Not always winning but at least taking part. Cognition, 113(1), 45-61. 10.1016/j.cognition.2009.07.009

De Neys, W., & Pennycook, G. (2019). Logic, fast and slow: Advances in dual-process theorizing. *Current Directions in Psychological Science*, *28*(5), 503-509. https://doi.org/10.1177/0963721419855658

Evans, J. S. B. (1989). 4. Concepts and Inference. *Mind & Language*, *4*(1-2), 29-34. https://doi.org/10.1111/j.1468-0017.1989.tb00237.x

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223-241. https://doi.org/10.1177/1745691612460685

Evans, J. S. B., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & cognition*, *11*(3), 295-306. doi: 10.3758/bf03196976

Fetzer, J. H. (1996). Biological Adaptations and Evolutionary Epistemology. Adaptive Behavior, 4(2), 201-210. https://doi.org/10.1177/105971239600400204

Gardner, M.K., Rothkopf, E.Z., Lapan, R. et al. The word frequency effect in lexical decision: Finding a frequency-based component. Memory & Cognition 15, 24–28 (1987). https://doi.org/10.3758/BF03197709

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. Cognition, 43(2), 127–171. https://doi.org/10.1016/0010-0277(92)90060-U

Gilmore, C., Göbel, S. M., & Inglis, M. (2018). *An introduction to mathematical cognition*. Routledge. https://doi.org/10.4324/9781315684758

Handley, S. J., Newstead, S. E., & Trippas, D. (2011). Logic, beliefs, and instruction: A test of the default interventionist account of belief bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 28–43. https://doi.org/10.1037/a0021098

Hudson, P. T., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. Journal of Memory and Language, 24(1), 46–58. https://doi.org/10.1016/0749-596X(85)90015-4

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kahneman, D., & Frederick, S. (2002). *Representativeness revisited: Attribute substitution in intuitive judgment*. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (p. 49–81). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.004

Kevin Barton, Jonathan Fugelsang & Daniel Smilek (2009) Inhibiting beliefs demands attention, Thinking & Reasoning, 15:3, 250-267, DOI: 10.1080/13546780902930917

Kostić, A., & Katz, L. (1987). Processing differences between nouns, adjectives, and verbs. Psychological Research, 49(4), 229-236.

Kostić, Đ. (1999). Frekvencijski rečnik savremenog srpskog jezika [Frequency dictionary of contemporary Serbian language]. Institut za eksperimentalnu fonetiku i patologiju govora i Laboratorija za eksperimentalnu psihologiju.

Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in cognitive sciences*, *22*(4), 280-293. doi: 10.1016/j.tics.2018.02.001

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current directions in psychological science*, *21*(1), 8-14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to

complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). Everyday consequences of analytic thinking. *Current Directions in Psychological Science*, *24*(6), 425-432. https://doi.org/10.1177/0963721415604610

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3–22. https://doi.org/10.1037/0033-2909.119.1.3

Stanovich, K. E. (2009). *Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory?* In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (p. 55–88). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230167.003.0003

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking & Reasoning*, *24*(4), 423-444. https://doi.org/10.1080/13546783.2018.1459314

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645–665. https://doi.org/10.1017/S0140525X00003435

Stupple, E. J., Ball, L. J., & Ellis, D. (2013). Matching bias in syllogistic reasoning: Evidence for a dual-process account from response times and confidence ratings. *Thinking & Reasoning*, *19*(1), 54-77. https://doi.org/10.1080/13546783.2012.735622

Stupple, E. J., Ball, L. J., Evans, J. S. B., & Kamal-Smith, E. (2011). When logic and belief collide: Individual differences in reasoning times support a selective processing model. *Journal of Cognitive Psychology*, *23*(8), 931-941. https://doi.org/10.1080/20445911.2011.589381

Svedholm-Häkkinen, A. M. (2015). Highly reflective reasoners show no signs of belief inhibition. Acta psychologica, 154, 69-76. https://doi.org/10.1016/j.actpsy.2014.11.008

Teovanović, P. (2019). Dual Processing in Syllogistic Reasoning: An Individual Differences Perspective. *Psihologijske teme*, *28*(1), 125-145. https://doi.org/10.31820/pt.28.1.7

Tyler, L. K., Bright, P., Fletcher, P., & Stamatakis, E. A. (2004). Neural processing of nouns and verbs: The role of inflectional morphology. Neuropsychologia, 42(4), 512-523. https://doi.org/10.1016/j.neuropsychologia.2003.10.001

Wason, P. C., & Evans, J. (1974-1975). Dual processes in reasoning? *Cognition, 3*(2), 141–154. https://doi.org/10.1016/0010-0277(74)90017-1

Pylyshyn, Z. W., & LePore, E. (Eds.). (1999). *What is cognitive science?*. Blackwell.

Appendix A

Words in LDT

Words used as the stimuli in the LDT were categorized into four groups. The first group consisted of the core words that appeared in syllogisms (e.g. oak, plant, and root). The second group included words that were semantically related to those from the first group, again three for each of the four themes (e.g., for the theme *oak*, those were *trunk, tree,* and *leaf*). The third group consisted of words that were semantically unrelated to any of the words appearing in the syllogisms, six for each theme. And finally, the fourth group included only pseudowords. The overview of the used words and selection of the stimuli are shown in the following table. It is important to note that original words are in Serbian language, and here are given translations. All the original used stimuli were nouns of the Srbian language, balanced according to the frequency and length (Kostić, 1999). A total 48 pseudowords were used and 12 of them were randomly assigned to each theme, they are not presented in the table.

Table 1

THEME	CORE WORDS	SEMANTICALLY RELATED	SEMANTICALLY UNRELATED
Oak	Oak	Tree	Ball
			Money
	Plant	Trunk	Letter
			Milk
	Root	Leaf	Piano
			Cheek
Dove	Bird	Woodpecker	Hotel

			Credit
	Dove	Flight	Lamp
			Belt
	Beak	Wing	Wednesday
			Rope
School	School	Yard	Arrow
			Sister
	Building	Floor	Piece
			Soldier
	Roof	Entrance	Snake
			Net
Lemonade	Lemonade	Drink	Nail
			Crown
	Juice	Glass	Jacket
			Motor
	Liquid	Ice	Radio
			Goat

Appendix B

Task Instructions (both EDS and LDT)

'This experiment consists of four syllogistic reasoning tasks and a lexical decision task. First, you will see two premises on the screen and later a conclusion based on them. Your task is to answer whether the conclusion follows logically from the arguments. Assume that the information given in the syllogisms is true. It is important to base your answer solely on the information given in the syllogisms. Answer YES only if you consider that the conclusion can be logically drawn from the premises, otherwise answer no. After this task, you will be presented with a lexical decision task. Answer YES if the string of letters presented on the screen is a word, or NO if it is not a word. Now you will do a short trial'