Dual Processing in Syllogistic Reasoning:
An Individual Differences Perspective

Predrag Teovanović

University of Belgrade, Faculty for Special Education and Rehabilitation,
Belgrade, Republic of Serbia

Abstract

The study aimed to examine several assumptions of dual process theories of reasoning by employing individual difference approach. A set of categorical syllogisms was administered to a relatively large sample of participants ($N = 247$) along with attached confidence rating scales, and measures of intelligence and cognitive reflection. As expected, response accuracy on syllogistic reasoning tasks highly depended on task complexity and the status of belief-logic conflict, thus demonstrating belief-bias on the group level. Individual difference analyses showed that more biased subject also performed poorer on Raven's Matrices ($r = .25$) and Cognitive Reflection Test ($r = .27$), which is in line with assumptions that willingness to engage and capacities to carry out type 2 processes both contribute to understanding of rational thinking. Moreover, measures of cognitive decoupling were significantly correlated with the performance on conflict syllogisms ($r = .20$). Individual differences in sensitivity to conflict detection, on the other side, were not related to reasoning accuracy in general ($r = .02$). Yet, additional analyses showed that noteworthy correlation between these two can be observed for easier syllogistic reasoning tasks ($r = .26$). Such results indicate that boundary conditions of conflict detection should be viewed as a function of both tasks' and participants' characteristics.

Keywords: dual process theory, individual differences, intelligence, cognitive reflection, conflict detection, cognitive decoupling

✉ Predrag Teovanović, Faculty for Special Education and Rehabilitation, University of Belgrade, Visokog Stevana 2, 11000 Belgrade, Republic of Serbia. E-mail: *teovanovic@fasper.bg.ac.rs*

Introduction

Categorical syllogisms[1] are characterized as one of the fruit flies (De Neys, 2012), key methods (Evans, 2003), or paradigm cases (Evans, 2008) for demonstrating dual processing in reasoning. In standard paradigm, people are asked to evaluate logical validity of given conclusions, with conclusions' validity (whether they logically follow from premises or not) and believability (whether they are consistent with prior beliefs or not) being systematically manipulated across items. As a consequence, some tasks are *non-conflict* (valid-believable and invalid-unbelievable), and some are *conflict* (invalid-believable and valid-unbelievable). The main experimental finding within this paradigm is belief bias – a prevalent tendency to evaluate syllogism based on conclusion's believability rather than on its logical validity (Evans, Barston, & Pollard, 1983).

The central assumption of dual process theories (DPT) is that human reasoning rests on interplay between two distinct types of thinking - type 1 (intuitive) and type 2 (analytical) cognitive processing. Type 1 processes are usually described as fast, effortless, and associative. According to Evans and Stanovich (2013), their defining characteristic is autonomy (type 1 are carried out whenever a triggering stimulus is encountered), along with independence from working memory (WM) capacity. On the other hand, type 2 processes require WM resources and involve cognitive decoupling, which seems to be crucial for mental simulation and hypothetical thinking. This makes type 2 processes relatively slow and resource-demanding.

What makes syllogisms attractive for DPT are differences between conflict and non-conflict tasks. Latter exemplify the situation in which the outcomes of two processes, one supporting belief-based response, and the other leading to logic-based response, are unison. However, in some situations, such as those represented by conflict syllogisms, two processes are supposed to lead to different outcomes, thus creating fertile ground for studying the ongoing competition for control over the response.

Traditional Dual-Process View on Syllogistic Reasoning

Within default-interventionist DPT account (Evans, 2008; Evans & Stanovich, 2013; Stanovich, 2009), two conflicting processes are seen to be of two distinct types. More precisely, type 1 cues a response based on believability of conclusion, the kind

---

[1] Categorical syllogisms are deductive arguments that relate three terms: minor (A), middle (B) and major (C). Syllogisms consist of three statements - two premises and a conclusion. Each statement contains standard logical quantifier (A – all, E – no, I – some, O – some ... are not), and their combination determine mood of syllogism (there are 64 possible combinations, 4^3). There are also four possible figures for syllogisms. Assuming that conclusion can have either A-C or C-A form, the premises may reference the terms as AB-BC, AB-CB, BA-BC, or BA-BC. Each mood can appear in each of four figures, making 512 possible syllogisms.

of response that leads to incorrect response on conflict tasks. In order to override it, one needs to inhibit belief-based intuition, to initiate more demanding type 2 processing, and to successfully perform it through cognitive decoupling and mental manipulation. Such operations are resource demanding, and DPT predicts that success in their performing will depend mainly on WM capacities.

Indeed, previous studies have shown that individual differences in WM capacities predict response accuracy on the conflict tasks, but not on non-conflict ones (Copeland & Radvansky, 2004; Handley, Capon, Beveridge, Dennis, & Evans, 2004; Quayle & Ball, 2000). To experimentally test this relation, De Neys (2006) introduced a secondary dot-memory task which puts some load on WM. In line with the expectation, burdening cognitive resources did not affect performance on non-conflict tasks, but it did markedly decrease response accuracy on conflict items. Further on, considering the high degree of overlap in individual difference of WM tasks and individual differences in measures of intelligence (e.g. Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004), negative correlation between IQ and reasoning performance was both expected (Evans, 2012; Stanovich, 2009), and previously observed (Newstead, Handley, Harley, Wright, & Farrelly, 2004; Sá, West, & Stanovich, 1999; Stanovich & West, 1998, 2000; Torrens, Thompson, & Cramer, 1999).

Nevertheless, relying solely on cognitive ability measures to explain response accuracy on conflict reasoning tasks neglects an aspect of human rationality which concerns a disposition to initiate type 2 processing, i.e. to detect the need to think harder. This faculty is often referred to as reflectivity (Evans & Stanovich, 2013; Stanovich, 2009). While *cognitive ability* refers to the capacity to sustain decoupled representations for purposes of mental simulation (that is, to successfully carry out type 2 processing), *cognitive reflection* is more concerned with mere willingness to engage type 2 processing (that is, to rethink the problem before providing any response). Previous body of research has detected reliable individual differences in syllogistic reasoning tasks performance once intelligence has been controlled for, and showed that cognitive reflection predicts them, measured by both self-rating scales, such as actively open-minded thinking and need for cognition (Kokis, Macpherson, Toplak, West, & Stanovich, 2002; West, Toplak, & Stanovich, 2008), and performance-based tests, such as Frederick's (2005) cognitive reflection test (Toplak, West, & Stanovich, 2011, 2014).

Recent Dual-Process Views on Syllogistic Reasoning

Recently, the classic assumption regarding to belief-logic conflict as a battle between type 1 and type 2 processes has been called into question (De Neys, Cromheeke, & Osman, 2011; De Neys & Franssens, 2009; De Neys & Glumicic, 2008; De Neys, Moyens, & Vansteenwegen, 2010; De Neys, Rossi, & Houdé, 2013). What is rather the case, according to De Neys' group, is that conflict occurs on the

intuitive level, between two type 1 processes. One is the traditional, i.e. *heuristic* intuitive response based on believability of conclusion. The other one, termed *logical* intuitive response, is grounded on the basic apprehension of logical principles. Traditionally considered to be an outcome of effortful reasoning, logic-based response is now assumed to be cued implicitly and automatically. Such claim that people are *intuitive logicians* (De Neys, 2012, 2014, 2018; De Neys & Bonnefon, 2013) is certainly bald, yet well-founded. Considerable amount of evidence, based on studies designed to contrast various behavioral and physiological measures (such as response latencies, confidence ratings, skin conductance, eye movements, activation of specific brain regions, etc.) on incorrectly solved conflict tasks in comparison to correctly solved non-conflict tasks, strongly indicate that people are generally sensitive to conflict between competing intuitive responses, even when they fail to provide correct solution. Consistent results observed on different reasoning tasks (including bat-and-ball problem and tasks typically employed to demonstrate base-rate neglect, conjunction fallacy, and ratio bias) and reported by independent research groups (for literature review, see De Neys, 2014) suggest that people show metacognitive awareness of a failure to conform to logic when responding incorrectly to conflict items.

Sensitivity to conflict between two intuitive responses in the presented study was examined by using confidence rating measures. In line with previous research within the area that employed the same measures (Brisson, Schaeken, Markovits, & De Neys, 2018; De Neys et al., 2011, 2013), it was expected that confidence ratings will be lower for incorrectly solved conflict syllogisms relative to correctly solved non-conflict syllogisms. This expectation and corresponding findings can be also viewed from the perspective of wider meta-reasoning framework as an evidence that people are able to identify if they have made a mistake on reasoning tasks which contain conflict between correct and misleading response (see e.g. Ackerman & Thompson, 2017).

Although primarily observed on a group level, results on logical intuition have recently been explored within the paradigm of individual differences. This line of research is still in its early phase, and at least two questions need to be distinguished here. The initial one was whether there are any individual differences in sensitivity to conflict detection? Empirical evidence unequivocally lead to the positive answer (for the first wave of empirical demonstrations, see Mevel et al., 2015, and Pennycook, Fugelsang, & Koehler, 2015). The second question is whether those who detect conflict also show a higher probability of responding correctly to conflict items, that is, is there a positive correlation between conflict detection and reasoning performance. Findings regarding this question are rather mixed, with some studies showing correlation (e.g. Frey, Johnson, & De Neys, 2018, Study 1 and Study 3b; Mevel et al., 2015; Pennycook et al., 2015; Swan, Calvillo, & Revlin, 2018, Study 1; see also Mata, Schubert, & Ferreira, 2014 and Mata, Ferreira, Voss, & Kollei, 2017 for evidences on relation between conflict detection and response accuracy by using

somewhat different paradigm), and the others that failed to reveal such relation (e.g. Frey et al., 2018, Study 3b; Swan et al., 2018, Study 2).

Pennycook's group was among the firsts ones which provided evidence on conflict detection failures (Pennycook, Fugelsang, & Koehler, 2012) and also who considered conflict detection as one of the sources of type 2 processing (Pennycook et al., 2015). In addition to conflict detection, these authors proposed another measure which can be derived from indirect measures (such as response times and confidence ratings), and it concerns cognitive decoupling. Specifically, they expressed this measure as the additional time needed to provide a correct response to conflict items as compared to the time spent on non-conflict items. Although it seems plausible to suppose how prolonged response time on conflict items for the aim of reaching correct response might reflect additional effort that participant puts in order to override the intuitive response, it remains unclear why the response time on non-conflict items should be used as a baseline. Also, scores derived in such a way correlated *negatively* with reasoning performance (Pennycook et al., 2015) or showed no significant correlation (Swan et al., 2018).

In the present study, the measures of cognitive decoupling were expressed as differences in confidence ratings for correctly solved conflict items and incorrectly solved conflict items. Such scores are supposed to reflect the additional effort needed to inhibit heuristic intuitive response after detecting a conflict between two responses. Accordingly, higher difference scores should reflect greater cognitive decoupling and they should be *positively* related to response accuracy.

Research Aims and Hypotheses

Study was designed with aim to explore whether response accuracy on conflict syllogistic reasoning tasks could be predicted by measures hypothesized in the assumptions of default-interventionist account (Evans, 2007; Evans & Stanovich, 2013; Stanovich, 2009) and more recent models which assume intuitive quality of belief-logic conflict (De Neys, 2012, 2014, 2018; Pennycook et al., 2015).

Following the De Neys' (2006) seminal experimental research, but also some correlational studies (Copeland & Radvansky, 2004; Handley et al., 2004; Newstead et al., 2004; Quayle & Ball, 2000; Sá et al., 1999; Stanovich & West, 1998, 2000; Torrens et al., 1999), it was hypothesized that measures of cognitive abilities, such as Raven's matrices or vocabulary test, should be related to the performance on conflict tasks, but not to the performance on non-conflict tasks. Further on, it was expected that cognitive reflection, typically seen as measures of propensity to engage type 2 processing, can contribute to our understanding of individual differences in reasoning on conflict tasks, over and above intelligence (Toplak et al., 2011, 2014).

Also, considering mixed results of recent studies (Frey & De Neys, 2017; Frey et al., 2018; Pennycook et al., 2015), present research was aimed to examine if there is a correlation between conflict sensitivity, measured through confidence ratings,

and response accuracy. Finally, it was expected that measures of cognitive decoupling, also derived from corresponding confidence ratings, will correlate positively with performance on conflict items, despite the fact that Pennycook and colleagues (2015) reported negative correlation, although for a somewhat different measure of decoupling.

Method

Participants

The study was part of a wider research on cognitive biases (see Teovanović, 2013; Teovanović, Knežević, & Stankov, 2015). It involved 247 undergraduate students (22 males) from the University of Belgrade who participated in research and earned partial course credit in return. Their mean age was 19.82 (SD = 1.29). Participants signed informed consent before taking part in the study.

Reasoning Tasks

Four types of reasoning task used in the present study are categorical versions of modus ponens (MP), modus tollens (MT), denial the antecedent (DA), and affirmation of the consequent (AC) from the propositional logic. Their formal structure is presented in the first three columns of Table 1.

For each task type, four items were derived, with some of them being based on examples from previous research (De Neys & Franssens, 2009; Kokis et al., 2002; Sá, West, & Stanovich, 1999). Two of these were *conflict items*, in which empirical status of conclusion was inconsistent with logical validity of the argument. Other two were *non-conflict items*, in which believability was congruent with the validity. This resulted in a total of 16 syllogistic reasoning items, which were presented to participants in a predetermined randomized order. Two practice items were administered first to ensure participants fully understood the task.

Participants were asked to evaluate syllogisms, i.e. to indicate whether the conclusion follows logically from the two premises. Instruction emphasized that all premises should be assumed to be true. No time limit for providing answers was imposed.

Nearly fair level of internal consistency was observed across 16 items (α = .69). However, reliability of individual differences in accuracy on conflict (α = .61) and congruent (α = .56) items was somewhat lower.

Confidence Ratings

After each submitted response, participants were asked to rate how confident they were that their response was correct. Confidence ratings were indicated on the percentage scale ranging from 50 ("*just guessing*") to 100 ("*absolutely certain*") in steps of 10. Depending on task conflict status and response accuracy, confidence rating scores were used to calculate measures of sensitivity to conflict detection, and the amount of cognitive decoupling.

Conflict Detection. As previously noted, logical intuition account emerged in the results evidencing that participants exhibit lower confidence for heuristic intuitive answers on conflict items as compared to non-conflict items (De Neys, 2012, 2014). To ensure that higher scores indicate more pronounced conflict detection, conflict incorrect confidence ratings were subtracted from non-conflict correct ones.

Bearing in mind considerable noisiness of individual measures of conflict detection (see e.g. De Neys, 2018; Frey & De Neys, 2017; Frey et al., 2018; Pennycook et al., 2015), absolute difference scores for each participant were divided by observed variability of his/her confidence ratings across all items, irrespective of task conflict and response accuracy. In such way, they were transformed into a measure which holds a resemblance to Cohen's *d*, and the weight was put on differences between corresponding confidence ratings for participants who generally showed less variability in confidence ratings. As an additional consequence, participants who showed no variability in their confidence ratings at all ($n = 16$) were automatically excluded from further analysis, since their differences scores could not be divided by zero.

Cognitive Decoupling. Cognitive decoupling scores were also calculated in such a way to ensure that higher scores indicate a larger amount of cognitive decoupling (conflict correct – conflict incorrect). To account for individual differences in confidence scores, raw (absolute) differences were divided by intra-individual SD of confidence ratings.

Other Measures

Raven's Matrices (Raven, Court, & Raven, 1979) consist of 18 items. Participants' task was to identify the missing symbol which completes the 3x3 matrix in the most logical manner by choosing from among five options. The time limit was restricted to six minutes. A fair level of internal consistency was observed ($\alpha = .79$).

Vocabulary Test (Knežević & Opačić, 2011) has 56 items. Subjects were asked to characterize the different words by choosing from among six options. No time limit for the completion of this test was imposed. On average, the participants completed this test in 13.11 minutes ($SD = 2.09$). Cronbach's alpha was .73.

Cognitive Reflection Test (CRT; Frederick, 2005) consists of only three questions, with each of them triggering most of the participants to give an immediate

and incorrect answer. Due to the small number of items, a low level of internal consistency was registered ($\alpha = .40$)

Procedure

Measures were administered in two sessions, one week apart. Personal identification numbers were used for matching participants' data. In the first session, participants completed categorical syllogisms in paper and pencil format. In the second session, a battery of cognitive ability tests was computer-administered.

Results

Four participants showed no variability in neither answering (accepted conclusions on each item) nor in providing confidence ratings (always expressed 100% certainty level) on syllogistic reasoning tasks. Their data were discarded from further analyses making the final sample consist of 243 participants.

Reasoning Accuracy

Performance on syllogistic reasoning tasks was analyzed first. Results, presented in detail in Table 1, indicate that non-conflict MP was the easiest task ($M = 98.6\%$, CI95 [97.1 - 99.3]), while conflict AC had the lowest rate of correct responses (e.g. only 10.3% of participants concluded that *Catfish is a fish* do not follow logically from *All fish have grills* and *Catfish has grill*).

Table 1

Mean Response Accuracy and Confidence Ratings (with 95%CI) on Syllogistic Reasoning Items

Structure of task				Content	Accuracy	Confidence
Figure 1 (B-C; A-B; A-C) AAA mood	Modus ponens (valid)	Believable	Non-conflict	All humans are mortal. I am a human. I am mortal.	99.2 [97.1 - 99.8]	97.86 [96.97 - 98.75]
				All birds have feathers. Robin is a bird. Robin has feathers.	98.0 [95.4 - 99.1]	93.79 [92.25 - 95.33]
		Unbelievable	Conflict	All mammals walk. Whales are mammals. Whales walk.	41.7 [35.7 - 47.9]	88.31 [86.22 - 90.40]
				All dogs are tame. Pit bulls are dogs. Pit bulls are tame.	73.3 [67.4 - 78.4]	87.24 [85.2 - 89.28]

Structure of task			Content	Accuracy	Confidence
Figure 1 (B-C; A-B; A-C)	AEE mood — Denying the antecedent (invalid)	Unbelievable — Non-conflict	Everything wooden is fuel. Gas is not wooden. Gas is not fuel.	71.7 [65.7 - 76.9]	84.69 [82.49 - 86.89]
			Every vegetable is edible. Banana is not a vegetable. Banana is not edible.	74.9 [69.1 - 79.9]	87.48 [85.43 - 89.53]
		Believable — Conflict	All the African countries are poor. Switzerland is not an African country. Switzerland is not poor country.	36.0 [30.3 - 42.2]	87.90 [86.00 - 89.80]
			All seas are salty. Danube is not a sea. Danube is not salty.	19.0 [14.6 - 24.4]	89.92 [87.94 - 91.90]
Figure 2 (C-B; A-B; A-C)	AAA mood — Affirming the consequent (invalid)	Unbelievable — Non-conflict	All trolleybuses use power. Boilers use power. Boilers are trolleybuses.	83.4 [78.3 - 87.5]	92.15 [90.37 - 93.93]
			All dogs go to heaven. Believers go to heaven. Believers are dogs.	75.4 [68.7 - 79.5]	86.13 [84.03 - 88.23]
		Believable — Conflict	All living beings need water. Roses need water. Roses are living beings.	22.7 [17.9 - 28.3]	90.08 [88.31 - 91.85]
			All fish have gills. Catfish has gills. Catfish is fish.	10.9 [7.6 - 15.4]	91.85 [90.18 - 93.52]
	AEE mood — Modus tollens (valid)	Believable — Non-conflict	All fruits are edible. Cigarettes are not edible. Cigarettes are not fruits.	83.8 [78.7 - 87.9]	90.99 [89.27 - 92.71]
			Medit. countries have a coastline. Hungary has no coastline. Hungary is not a Medit. country.	82.2 [76.9 - 86.5]	86.94 [84.78 - 89.10]
		Unbelievable — Conflict	All birds can fly. Ostrich cannot fly. Ostrich is not a bird.	47.4 [41.2 - 53.6]	85.72 [83.55 - 87.89]
			All capital cities have subway. Belgrade has no subway. Belgrade is not a capital city.	41.3 [35.3 - 47.5]	87.28 [85.16 - 89.40]

Note. 95% confidence intervals for mean accuracy scores are calculated using Wilson formula.

Two-way ANOVA for repeated measures was run to examine the effects of task type and believability-validity conflict. Results, descriptively presented in Figure 1, indicate that both task type ($F_{3,726} = 72.95$, $p < .001$, $\eta^2 = .23$) and task conflict status

($F_{1,242}$ = 525.46, p < .001, η^2 = .69) significantly determined response accuracy[2]. As expected, performance dropped rapidly when conflict between believability and validity of conclusion was introduced, confirming reliable findings on belief bias (Evans et al., 1983). Additionally, valid arguments (MP and MT) were generally easier to evaluate in comparison to invalid ones (AC and DA), which is in line with previously reported results (Brisson et al., 2018).
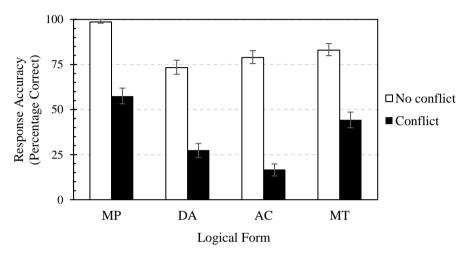

Figure 1. Performance on syllogisms as a function of task type and task conflict.

Response Confidence

As results presented in the last two columns of Table 1 show, mean confidence ratings across items were consistently higher than response accuracy, except for the easiest tasks (non-conflict MP). Nevertheless, confidence ratings were analyzed in relation to the task conflict status and response accuracy. For each participant who had at least one appropriate data in corresponding cell (that is, who did not give all belief-based or all logic-based answers), individual confidence rating scores were computed for four conditions that result from crossing the conflict status (conflict vs. non-conflict) and response accuracy (correct vs. incorrect). These measures were further used as a basis for calculating conflict detection and cognitive decoupling measures. Descriptive statistics are presented in the last three rows of Table 2.

---

[2] Two repeated factors were in low-intensity interaction ($F_{3,726}$= 24.25, p < .001, η^2 = .09). Effect of conflict for MP and DA tasks (η^2 = .51 and η^2 = .53, respectively) was to a certain degree weaker in comparison to the same effect for AC task (η^2 = .71), but it was stronger than conflict effect for MT task (η^2 = .39). Such results shed some light on classical finding that belief bias is more pronounced on invalid syllogisms (Evans et al., 1983).

Table 2

Correlations among Confidence Rating (CR) Measures as a Function of Task Conflict and Response Accuracy, Measures of Cognitive Detection and Cognitive Decoupling (Raw Differences), and Response Accuracy (Percentage Correct), Including Means and Standard Deviations

Confidence rating	1	2	3	4	5	6	7	8
1. Conflict correct CR	-	.54***	.56***	.60***	-.06	.71***	.05	.13*
2. Conflict incorrect CR		-	.71***	.42***	-.52***	-.21**	.09	.02
3. Non-conflict correct CR			-	.47***	.23**	.06	.08	.04
4. Non-conflict incorrect CR				-	.01	.35**	-.08	.12
5. Conflict Detection (3-2)					-	.36**	-.02	.00
6. Cognitive Decouple (1- 2)						-	-.01	.11
7. Non-conflict accuracy							-	-.26**
8. Conflict accuracy								-
N	222	240	243	156	240	219	243	243
M	85.16	89.98	91.51	80.68	1.47	-4.79	83.45	36.54
SD	13.92	9.96	8.74	15.82	7.19	12.06	17.75	23.20

$^*p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Conflict Detection

Group measures. As it was expected, average participant showed lower confidence on incorrectly solved conflict items ($M = 89.98$, $SD = 9.96$) in comparison to correctly solved non-conflict items ($M = 91.51$, $SD = 8.74$), as predicted by conflict detection account (De Neys, 2012, 2014; De Neys & Bonnefon, 2013). Although the difference was statically significant ($F_{1, 239} = 10.17$, $p = .002$), it should be noted that the effect was relatively small in size ($\eta^2 = .04$).

Individual measures. A total of 16 participants were excluded from following analyses since they showed no intra-individual variability in confidence rating scores. Besides that, three participants gave all correct answers on conflict items, thus showing no belief bias. Among 224 participants with valid data, majority ($n = 128$; $P = 57.6\%$) showed expected decrease in the response confidence for conflict incorrect items as compared to confidence ratings for non-conflict correct items, with the average decrease of 6.02% ($SD = 5.93$). Nevertheless, there were also 41 ($P = 33\%$) of biased participants who showed higher confidence (mean increase = 5.66, $SD = 4.61$), and 21 ($P = 9.4\%$) who provided the same rating for both classes of items. The last two groups indicate that some participants do not show sensitivity to conflict as measured by their confidence scores, which replicates earlier findings (Mevel et al., 2015; Pennycook et al., 2015).
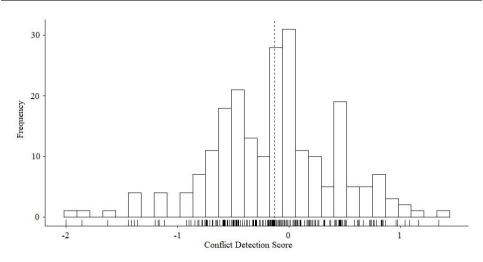
Figure 2. Distribution of conflict detection measures.

Distribution of individual measures of sensitivity to conflict detection is presented in Figure 2. In the whole biased sample, reasoning accuracy on conflict syllogisms could not be predicted neither by individual conflict detection measures ($r = .04$, $p = .57$) nor by categorical three-level group factor ($F_{2, 221} = 2.35$, $p = .10$)[3]. Also, numerical conflict detection measures were not related to response accuracy on non-conflict items ($r = -.03$, $p = .62$), nor to scores on Raven's matrices ($r = .08$, $p = .25$), vocabulary test ($r = .04$, $p = .55$) and CRT ($r = .03$, $p = .68$). The very same pattern of results is observed when raw difference scores were used as measures of conflict detection.

Cognitive Decoupling

A total of 207 participants gave at least one correct answer to conflict items, while three of them had all correct responses (which disallowed the computation of difference score). Among 204 subjects, only the minority ($n = 67$, $P = 32.8\%$) showed an increase in confidence after correctly solved conflict items in comparison to incorrectly answering them (average increase 6.97, $SD = 6.75$). On the other hand, 14 participants (6.9%) showed no difference between two confidence ratings, while 123 (60.3%) showed a decrease of confidence ($M = 12.49$, $SD = 9.89$). These three

---

[3] Within two conflict detection groups (increased vs. decreased confidence), marked correlations between measures of conflict detection and reasoning performance were observed. Specifically, for the group that showed positive conflict detection, the magnitude of conflict detection positively correlated with accuracy ($r = .38$, $p = < .001$), which is in line with results of previous studies (Mevel et al., 2015). However, for the group that showed "inverse" conflict detection, the magnitude of conflict detection correlated *negatively* with accuracy ($r = -.23$, $p = .054$).

groups ("increase", "same" and "decrease") did not differ in respect to response accuracy on conflict items ($F_{1, 201} = 0.29$, $p = .75$). However, within both "increase" and "decrease" cognitive decoupling group, significant relation with performance was observed ($r = .25$, $p = .047$; $r = .33$, $p < .001$; respectively).
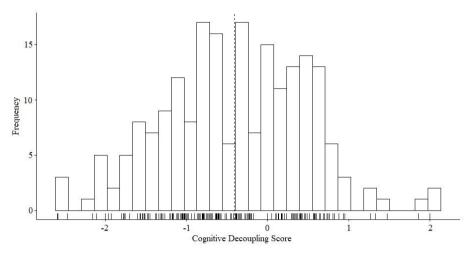

Figure 3. Distribution of cognitive decoupling measures.

Numerical measures of cognitive decoupling were related to both response accuracy ($r = .20$, $p = .004$) and conflict detection ($r = .38$, $p < .001$), marginally related to scores on Raven's matrices ($r = .13$, $p = .07$), and showed no significant relation to scores on vocabulary or CRT ($rs < .10$, $ps > .30$). Distribution of cognitive decoupling measures is presented in Figure 3.

Predictors of Reasoning Accuracy

Final set of analyses aimed to examine if measures of cognitive abilities, cognitive reflection, conflict detection, and cognitive decoupling are related to biased reasoning.

Separate bivariate correlations of these measures with performance scores on conflict and non-conflict tasks, as well as results of multiple regression analyses are presented in Table 3. Results indicate that scores on Raven's matrices, vocabulary test, CRT and conflict decoupling were indeed related to achievement on conflict items (rs ranged from .18 to .27, $ps < .01$), but they were not associated with performance on non-conflict items ($rs < .10$, $ps > .20$). Tests of the difference between two dependent correlations with one variable in common (Lee & Preacher, 2013) was run. One-tailed levels of significance were used considering unidirectional expectation that predictors are related to performance on conflict, but not on non-conflict syllogistic reasoning task. Differences between corresponding correlation

coefficients were significant in the case of Raven's matrices ($Z = 1.72$, $p = .04$), CRT ($Z = 2.02$, $p = .02$), and cognitive decoupling ($Z = 2.73$, $p = .003$).

Table 3

Results of Performance Prediction for Conflict and Non-Conflict Items

	Conflict		Non-Conflict		Test of r difference	
Predictors	r	β	r	β	Z	p (one-tailed)
Raven's matrices	.25***	.15*	.08	.07	1.71	.04
Vocabulary rest	.18**	.08	.06	.06	1.19	.12
Cognitive reflection	.27***	.13†	.05	.05	2.22	.01
Conflict detection ($n = 224$)	.02	-.07	-.01	.02	0.66	.25
Cognitive decoupling ($n = 207$)	.20**	.20**	-.10	-.12	2.73	.003
Adjusted R^2	8.6%		0.1%		1.71	.04

Note. †$p < .06$; *$p < .05$; **$p < .01$; ***$p < .001$.

In general, cognitive measures accounted for only 0.1% of variance of non-conflict items score ($F_{5,198} = 1.01$, $p = .41$), yet their predictive capacity was non-negligible when predicting scores on conflict items ($R^2 = 8.6\%$, $F_{5, 198} = 4.81$, $p < .001$). Significant partial contributions to regression model in the case of conflict response accuracy were registered for cognitive decoupling (β $= .20$, $p = .008$), Raven's matrices (β $= .15$, $p = .036$) and also marginally for cognitive reflection (β $= .14$, $p = .054$), but not for vocabulary (β $= .08$, $p = .28$) and conflict detection (β $= -.07$, $p = .33$).

Finally, hierarchical regression analysis was performed in order to examine if there is reliable variance in reasoning, over and above what can be predicted by traditional intelligence measures, which can be explained by individual differences in cognitive reflection. In the first step, performance scores on eight conflict items were regressed on Raven's matrices and vocabulary test, and they accounted for 7.3% of the variance ($F_{2, 240} = 9.40$, $p < .001$). After that, CRT measure was entered, and it accounted for additional 4.4% of the variance ($\Delta F_{1, 239} = 11.91$, $p = .001$).

Discussion

This study was aimed to examine predictors of individual differences in reasoning which can be hypothesized by following the basic assumptions of dual process theories. To this end, a set of categorical syllogisms was administered, along with confidence rating scales and several standard psychometric measures of cognitive functioning. Some syllogisms were worded in a way that made believability of conclusion consistent with argument validity (so-called control, i.e. non-conflict tasks), while some others included belief-logic conflict, either by using empirically unbelievable statement as a conclusion of logically valid syllogism or by

coupling a believable statement with invalid conclusion. As expected, the conflict between conclusion's validity and believability accounted for as much as 71% of response accuracy variability. This confirms reliability of belief bias finding, firstly reported by Evans et al. (1983) and replicated many times ever since (e.g. De Neys et al., 2011; De Neys, & Franssens, 2009; Sá et al., 1999; Stupple & Ball, 2008).

According to standard default-interventionist DPT account (De Neys, 2006; Evans, 2007; Evans & Stanovich, 2013; Stanovich, 2009), when believability and validity of conclusion are in accordance, two types of cognitive processes lead to correct response, which explains consistently higher performance on non-conflict items. However, these two are supposed to cue different responses on conflict syllogisms. More precisely, type 1 processes provide a *default* intuitive response (based on believability of conclusion), on which subsequent type 2 might *intervene* in order to override it with more thoughtful reasoning (based on logic rules).

There are two aspects of type 2 intervention, both amenable to measurement of individual differences. The first one is concerned with capability of central executive to perform demanding analytical operations, including inhibition of intuitive response, cognitive decoupling, mental simulation and hypothetical thinking. Individual differences in this capacity are usually expressed through psychometric measures of intelligence. In previous studies, higher rates of correct responses on conflict syllogisms were indeed related to both measures of WM capacity (Copeland & Radvansky, 2004; Handley et al., 2004; Quayle & Ball, 2000), and intelligence (Newstead et al., 2004; Sá et al., 1999; Stanovich & West, 1998, 2000; Torrens et al., 1999). In the present study, scores on Raven's progressive matrices correlated with the performance on conflict syllogisms, but not on the non-conflict ones, and this difference was found to be statistically significant.

The same pattern of results was observed in the case of CRT – correlation with response accuracy was significantly higher for conflict items in comparison to the non-conflict ones. This finding is directly related to the second aspect of presupposed type 2 intervention, concerned with the probability of such an intervention. Individual differences in detection of the need to engage type 2 processing, expressed both through self-rating and performance-based measures, has been shown to predict reasoning performance, over and above intelligence (Kokis et al., 2002; Toplak et al., 2011, 2014; West et al., 2008). The very same result was observed in the presented study, thus confirming the claim that individual differences in rational thinking are not reducible to IQ (Stanovich, 2009).

Probability of type 2 intervention can also be manipulated experimentally, e.g. reduced by limiting time allowed for providing response (Evans & Curtis-Holmes, 2005), and by putting the load on working memory capacities (De Neys, 2006), or it can be enlarged through presentation of tasks with difficult-to-read font (Alter, Oppenheimer, Epley, & Eyre, 2007). Besides, it has been recently proposed that bottom-up (stimulus-related) factors of type 2 processing should be taken into account (Pennycook et al., 2015). Within hybrid (De Neys, 2014) and three-stage

(Pennycook et al., 2015) DPT models, the conflict between responses has been conceptualized as a clash between intuitions, rather than between an intuition and a thought. Implicit awareness of the belief-logic conflict was demonstrated by showing how even biased reasoners implicitly activate basic normative principles, which was evidenced by lower confidence or increased response time on incorrectly solved conflict items in comparison to correctly answered non-conflict items (Brisson et al., 2018; De Neys & Franssens, 2009; De Neys & Glumicic, 2008; De Neys et al., 2010, 2011, 2013; Frey et al., 2018). This finding is validated through different indirect measures on various reasoning tasks (see e.g. De Neys, 2014, 2018). In the present study, group-level conflict detection was also observed - mean confidence ratings were somewhat lower for conflict incorrect in comparison to non-conflict correct responses.

Recently, calls for exploration of potential benefits of individual difference perspective on conflict detection have emerged (see e.g. De Neys, 2014; De Neys & Bonnefon, 2013), mainly driven by findings that conflict detection is not ubiquitous (e.g. Pennycook et al., 2012), and that biased reasoners are less sensitive to conflict detection (e.g. Mata et al., 2014; Mevel et al., 2015; Pennycook et al., 2015). However, asking if there are individual differences in conflict detection (i.e. is conflict detection indeed flawless/perfect) is not the same as asking whether those who miss to detect conflict also fail to provide a correct answer. In general, it is possible that individual differences in conflict detection do exist, while the most biased reasoners still show some sensitivity to conflict. The results of the present study seem to be in accordance with such possibility. Although a considerable variability of conflict detection scores was registered, these variations were not related to variability in response accuracy. However, a null result could also be due to other reasons.

Variability between participants concerning intra-individual fluctuations of confidence rating scores could be seen as a potential source of error variance. It could be argued that the same nominal decrease (or increase) in confidence brings different information depending on the general stability of confidence for a given participant. In other words, the difference should weigh more when a subject had relatively stable confidence ratings than when s/he showed a greater variation of confidence ratings across items. For this reason, raw difference scores were divided by standard deviation of individual confidence ratings across items. As an added benefit, participants who showed no variability were excluded from further analyses (instead to be classified as showing no conflict detection). Nevertheless, not even "cohen-d-ized" conflict detection measures were related to response accuracy.

Null result could also be due to differences in logic complexity between the tasks employed in the present study. Logical intuitions are hypothesized to be bounded to non-complex conditions, meaning that they are expected to arise only for relatively simple problems which can be solved by using basic normative principles (De Neys, 2012, 2014). As Stanovich (2018) argue, probability of successful

detection strongly depends on mindware instantiation. Recently, Brisson et al. (2018) demonstrated group-level conflict detection only for MP and MT syllogisms (easy problems), but not for DA and AC syllogisms (hard problems). Our data confirm such finding - conflict was implicitly detected in the case of valid unbelievable MP and MT items (*non-conflict correct M* = 93.05, *conflict incorrect M* = 85.56, *t* = -7.68, *df* = 183, *p* < .001), but not in the case of invalid believable DA and AC items, where reversed situation was detected (*non-conflict correct M* = 88.23, *conflict incorrect M* = 91.11, *t* = 3.68, *df* = 227, *p* < .001). Additionally, measures of conflict detection were not related to response accuracy on corresponding items in the case of hard tasks (*r*(211) = -.06, *p* = .38), but they were in the case of easy ones (*r*(171) = .26, *p* < .001). In other words, not only that group-level conflict detection findings are dependent on task complexity, but it seems that also is the case with individual-difference results. When the underlying principle is relatively simple, individual differences in activation of logical intuitions about the given problem might arise leading to differences in sensitivity to conflict between logic and intuition, which serves as a signal for initiating type 2 processing, which then affects their performance on conflict syllogisms. Such moderating effect of task complexity can be used to explain inconsistencies in results of previous studies (Frey et al., 2018; Swan et al., 2018), but also to enhance our understanding of conditions in which meta-cognitive monitoring operates (Ackerman & Thompson, 2017).

Potential predictive capacity of indirect measures of cognitive decoupling was tested as well. Two differences in regard to previous operationalization of this capacity (Pennycook et al., 2015; Swan et al., 2018) should be noted. First, confidence ratings for *conflict incorrect* responses (and not for *non-conflict* tasks irrespective of response accuracy) were used as a baseline. Consequently, differences between implicit measures related to successful and unsuccessful overriding of heuristic intuitive response were captured. Similarity of proposed cognitive decoupling measures and measures of monitoring resolution (Koriat, 2012) should be noted. This measure indicates a degree of metacognitive sensitivity, i.e. one's ability to distinguish between correct and incorrect responses, and as such it should reflect the ability to sustain decoupled representations of problem task in order to accomplish required mental operations. Moreover, results have shown that these measures are *positively* (and not negatively) related to response accuracy, even after controlling for individual differences in intelligence and cognitive reflection. Further on, cognitive decoupling and conflict detection measures were positively related indicating that implicit apprehension of clash between intuitive responses was in relation to a more successful override of heuristic response.

It should be noted that the generalizability of the results reported in this study is limited, considering that only syllogistic reasoning tasks were employed. Moreover, only four out of the 512 possible tasks were translated into items. Besides that, individual measures of conflict detection are known to be fairly noisy (De Neys, 2018; Frey & De Neys, 2017; Frey et al., 2018), and relatively low level of internal

consistency of response accuracy scores should be noted. One of possible solution for these problems is to collect various indirect measures (e.g. response time, measures of skin conductance, time of fixation of critical parts of task, etc.) on several reasoning tasks on which conflict between normative rule and "stronger" intuitive response is pronounced (e.g. bat-and-ball, base-rate, and ratio-bias tasks). First attempts in that direction are already made. The results of these studies indicate a certain level of convergence of multiple conflict detection indexes across several tasks (Frey et al., 2018). However, additional research is needed in order to reach a more conclusive understanding of their generalizability (cf. Frey & De Neys, 2017). If the results turn out to be positive, it could be additionally examined whether sensitivity to conflict detection is correlated with traditional psychometric constructs (De Neys, 2018). It seems that at least some of the future individual difference studies will trace that lead.

References

Ackerman, R., & Thompson, V. A. (2017). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences*, *21*(8), 607-617.

Alter, A., Oppenheimer, D., Epley, N., & Eyre, R. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*(4), 569-576.

Brisson, J., Schaeken, W., Markovits, H., & De Neys, W. (2018). Conflict detection and logical complexity. *Psychologica Belgica*, *58*(1), 318-322.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, *32*(3), 277-296.

Copeland, D., & Radvansky, G. (2004). Working memory and syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*, *57*(8), 1437-1457.

De Neys, W. D. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science, 17*(5), 428-433.

De Neys, W. D. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, *7*(1), 28-38.

De Neys, W. D. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning*, *20*(2), 169-187.

De Neys, W. D. (Ed.). (2018). *Dual process theory 2.0*. Routledge.

De Neys, W. D., & Bonnefon, J. F. (2013). The 'whys' and 'whens' of individual differences in thinking biases. *Trends in Cognitive Sciences*, *17*(4), 172-178.

De Neys, W. D., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PloS One*, *6*(1), e15954.

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

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*(3), 1248-1299.

De Neys, W. D., Moyens, E., & Vansteenwegen, D. (2010). Feeling we're biased: Autonomic arousal and reasoning conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 208-216.

De Neys, W. D., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: Cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, *20*(2), 269-273.

Evans, J. S. B. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454-459.

Evans, J. S. B. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, *13*(4), 321-339.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255-278.

Evans, J. S. B. (2012). Questions and challenges for the new psychology of reasoning. *Thinking & Reasoning*, *18*(1), 5-31.

Evans, J. S. B., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, *11*(3), 295-306.

Evans, J. S. B., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning, 11*(4), 382-389.

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.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25-42.

Frey, D., & De Neys, W. (2017). Is conflict detection in reasoning domain general? *Proceedings of the Annual Meeting of the Cognitive Science Society, 39*, 391-396.

Frey, D., Johnson, E. D., & De Neys, W. (2018). Individual differences in conflict detection during reasoning. *The Quarterly Journal of Experimental Psychology, 71*(5), 1188-1208

Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. S. B. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking & Reasoning*, *10*(2), 175-195.

Knežević, G., & Opačić, G. (2011). *Vocabulary test.* Unpublished Material.

Kokis, J., Macpherson, R., Toplak, M., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, *83,* 26-52

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, *119*(1), 80-113.

Lee, I. A., & Preacher, K. J. (2013). *Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software].* Available from http://quantpsy.org

Mata, A., Ferreira, M. B., Voss, A., & Kollei, T. (2017). Seeing the conflict: An attentional account of reasoning errors. *Psychonomic Bulletin & Review, 24*(6), 1980-1986.

Mata, A., Schubert, A. L., & Ferreira, M. B. (2014). The role of language comprehension in reasoning: How "good-enough" representations induce biases. *Cognition, 133*(2), 457-463.

Mevel, K., Poirel, N., Rossi, S., Cassotti, M., Simon, G., Houdé, O., & De Neys, W. (2015). Bias detection: Response confidence evidence for conflict sensitivity in the ratio bias task. *Journal of Cognitive Psychology*, *27*(2), 227-237.

Newstead, S. E., Handley, S. J., Harley, C., Wright, H., & Farrelly, D. (2004). Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology*, *57*(1), 33-60.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2012). Are we good at detecting conflict during reasoning?. *Cognition*, *124*(1), 101-106.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80*, 34-72.

Raven, J. C., Court, J. H., & Raven, J. (1979). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* London: H. K. Lewis & Co.

Quayle, J., & Ball, L. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *The Quarterly Journal of Experimental Psychology, 53A*(4), 1202-1223.

Sá, W., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, *91,* 497-510.

Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55-88). Oxford: Oxford University Press.

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

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*, 161-188.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645-665.

Stupple, E. J., & Ball, L. J. (2008). Belief-logic conflict resolution in syllogistic reasoning: Inspection-time evidence for a parallel-process model. *Thinking & Reasoning*, *14*(2), 168-181.

Swan, A. B., Calvillo, D. P., & Revlin, R. (2018). To detect or not to detect: A replication and extension of the three-stage model. *Acta Psychologica, 187*, 54-65.

Teovanović, P. (2013). *Susceptibility to cognitive biases* (Doctoral disertation). Retrieved from National Repository of Dissertations in Serbia. (Accession No. 3303).

Teovanović, P., Knežević, G., & Stankov, L. (2015). Individual differences in cognitive biases: Evidence against one-factor theory of rationality. *Intelligence, 50*, 75-86.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*(7), 1275-1289.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, *20*(2), 147-168.

Torrens, D., Thompson, V., & Cramer, K. (1999). Individual differences and the belief bias effect: Mental models, logical necessity, and abstract reasoning. *Thinking and Reasoning*, *5*(1), 1-28.

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*(4), 930-941.