Studia Psychologica, Vol. 65, No. 4, 2023, 320-336
https://doi.org/10.31577/sp.2023.04.883

Acquired Knowledge and Bias Susceptibility: Mindware Automatization Measured with a Two-Response Paradigm and Its Relationship with Conflict Detection

Roman Burič
Institute of Experimental Psychology, Centre of Social and Psychological Sciences, Slovak Academy of Sciences, Bratislava, Slovak Republic

According to a recent model of mindware automatization, the extent to which the mindware is automatized should play a crucial role in preventing cognitive biases, as it should lead to very easy detection of a conflict between a misleading intuition and the logical structure of a situation. To examine the model in the present study, mindware automatization was measured as intuitive accuracy in neutral tasks commonly used as a measure of mindware instantiation. Participants also solved two types of conflict reasoning tasks, with response time and confidence used as measures of conflict detection. The results indeed showed the relationship of mindware automatization and most of the conflict detection measures with reasoning accuracy; however, mindware automatization was not related to conflict detection measures, except in one of the detection indices. Mindware automatization also emerged as a significant predictor of reasoning accuracy but lost its predictive power when other variables were added to the model. Overall, the results provide little support for the mindware automatization model. However, whether the relationship between the measures is linear and whether conflict detection is even necessary if the mindware is automatized remain open questions.

Key words: mindware automatization, two-response paradigm, conflict detection, individual differences, bias susceptibility

Introduction

Since the early studies of Kahneman and Tversky (1972, 1973; Tversky & Kahneman, 1974, 1983) it became evident that people often fall for cognitive biases in situations that induce compelling, intuitive conclusions. The most common explanation of these biases (e.g., Kahneman, 2011) is overreliance on

Correspondence concerning this article should be addressed to Roman Burič, Institute of Experimental Psychology, Centre of Social and Psychological Sciences, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovak Republic. E-mail: roman.buric@savba.sk
The data for this study are publicly available at: https://osf.io/6kxsr/. Supplementary material is available at https://journals.savba.sk/index.php/studiapsychologica/article/view/1344/485

Received December 1, 2022

the quick, autonomous, heuristic-based type 1 processes, instead of slow, deliberate, and logic-based type 2 processes.

As De Neys and Bonnefon (2013) sum it up, there are three possible failures that can lead to such biases – *mindware, detection, and inhibition failures*. *Mindware* represents the knowledge structures needed to solve a problem – people cannot solve logical syllogisms if they lack knowledge of formal logic (Stanovich, 2018). It allows them to *detect the conflict* between compelling Type 1 response and logical principles of the tasks and to *inhibit the Type 1 response*.

As Stanovich (2018) points out, very little attention is drawn to the interaction of these elements. This led the author to come up with his mindware automatization model[1] (Figure 1). To put it simply, if mindware is missing, the conflict detection is not possible, as there are no instantiated rules to interfere with the automatic type 1 response. If mindware is present, conflict detection is at least possible and can lead to overriding intuition. Once mindware is fully *automatized*, it has become a part of type 1 processes, and can be triggered autonomously. Stanovich (2018) proposes that people with automatized mindware should be able to detect the conflict very easily and even to generate logically correct type 1 responses.

[1] In this paper, the terms *mindware automatization* and *mindware instantiation* are used separately. I refer to *mindware instantiation* as the extent to which mindware is adopted, in line with previous research (e.g., Burič & Konrádová, 2021; Burič & Šrol, 2020; Šrol & De Neys, 2020). *Mindware automatization* is referred to as the "final stage" of mindware adoption, meaning the mindware is acquired to such extent that it is fully automatized. Even though the automatized mindware can be perceived as extremely instantiated, they are qualitatively different. The distinction between the two is that if mindware is to be considered *automatized*, it needs to be learned to the extent it is a part of type 1 processes and thus can be channeled autonomously, without drawing upon working memory resources.

The main aim of this paper is to examine Stanovich's hypothesis, which directly proposes that the better mindware is instantiated or even automatized, the more likely it will lead to conflict detection. Some of the recent studies have already shown that mindware predicts bias susceptibility (Burič & Šrol, 2020; Klaczynski, 2014; Klaczynski & Felmban, 2020) and conflict detection (Frey et al., 2018; Šrol & De Neys, 2020), however, none of these studies measured the extent to which mindware is automatized.

Even though Stanovich (2018) states that conflict detection should be very easy in cases when mindware is automatized, there are also contradictory proposals. The hybrid model of dual processes of Bago and De Neys (2017) postulates that reasoners generate *two types of intuitive type 1 response* – the heuristic type 1 response, and the type 1 response based on stored mindware. Whether participants provide logically correct intuitive response depends on the intensities of both intuitions. If their relative difference is high, meaning the logical type 1 response is much more intense, participants provide logically correct type 1 response. However, if the mindware is fully automatized and the logical type 1 response is so dominant that the participants provide logically correct intuitive response, the question arises whether there is any interference with heuristic type 1 response at all and thus whether conflict detection even takes place.

The present study has two goals. The main one is to test Stanovich's (2018) claim about the relationship between mindware automatization and conflict detection. Despite my argument above, I hypothesize that participants with better automatized mindware will be better at detecting the conflict in conflict tasks. Also, mindware automatization should be a stronger predictor of performance in all reasoning tasks than the final mindware ac-

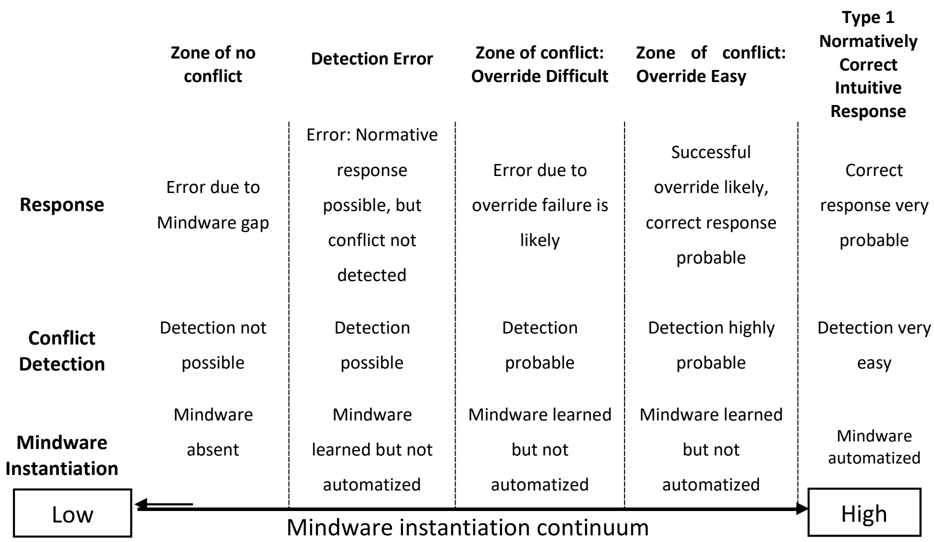

Figure 1 Processing states on the mindware continuum (Stanovich, 2018, p. 435).

curacy and the time needed to answer as two separate variables. The secondary goal is to examine the intuitive accuracy of neutral task as a measure of mindware automatization. I want to achieve this by using a two-response paradigm, which is commonly used to separate type 1 and type 2 responses.

Method

Pre-registration for this study is available at https://osf.io/zgdnk

Participants and Data Collection

The data were collected through an online survey hosted on Qualtrics. Participants were recruited by an external research agency in accordance with predetermined criteria.

To determine the sample size, I applied a priori power analysis for linear regression using the G*Power 3.1.9 software (Faul et al., 2007). The analysis was conducted using an alpha of 0.05, a power of 0.80, and an f^2 of 0.10, suggesting a sample size of 200 participants.

Overall, data were collected from 277 participants. After excluding those who did not finish the whole survey or who responded incorrectly to the control questions, 196 participants remained (100 female, 96 male, age: M = 48.67, SD = 15.93), 34.7% of whom reported having at least some college or a university degree and 64.3% reported having finished a high school education.

Materials

Conflict and no-conflict tasks manipulation. All the participants solved the conflict and no-conflict tasks of each reasoning problem type – syllogistic problems and base-rate neglect problems. Due to a technical error, conjunction fallacy problems were dropped from the analyses.

In the conflict versions, the problems themselves cue an intuitive response that conflicts with the logical structure of the task; therefore, it is incorrect. In the no-conflict versions, the intuitive answer is also the logically correct one (see Supplementary material). These two types of problems were used to show to what extent participants fell to the cognitive biases and subsequently used to compute the conflict detection indices that were used in the following analyses.

Syllogistic reasoning problems: Participants were presented with two premises and a conclusion and were asked to decide whether the conclusion logically follows from the premises. I used four conflict and four no-conflict items used in previous studies (Bago & De Neys, 2017; Burič & Šrol, 2020; De Neys et al., 2010), based on work of Sá et al. (1999) and Markovits and Nantel, 1989. These items were shown to reliably cue heuristic-based intuitive response and to be related to multiple other measures such as cognitive reflection, analytic and intuitive thinking dispositions, or cognitive capacity (Burič & Šrol, 2020; Klaczynski, 2014; Šrol & De Neys, 2020; Thompson & Johnson, 2014; Toplak et al., 2011). This points to conflict syllogisms to be a valid measurement of belief bias. Believability of the conclusion in these conflict syllogisms conflicts with its logical validity. The no-conflict syllogisms are as similar as possible in their content to the conflict ones, but the logical validity of the conclusion is in line with its believability. The four conflict items showed relatively good reliability (ω = .75).

Base-rate neglect problems: Participants were presented with two pieces of information about an individual who was randomly drawn from a sample that consisted of two groups (e.g., lawyers and dustmen). Participants were then given a single stereotypical trait pertaining to the individual, which cued one of the groups ("Person 'A' is rich"). Finally, they were given a proportion of the two groups in the original sample ("There are 995 dustmen and 5 lawyers"). The task was to decide to which group the imaginary person more likely belongs to. Again, I used four conflict and four no-conflict items from previous studies (Bago & De Neys, 2017; Burič & Šrol, 2020). The four no-conflict problems contained base-rate information which favored the group with the stereotypical trait. In contrast, the four conflict items were constructed by simply changing the base rates to favor the group contrary to the stereotypical characteristic. The conflict items showed a reliability of ω = .56.

Conflict detection indices. Conflict detection ability was captured by two measures (e.g., Bago & De Neys, 2017; Burič & Konrádová, 2021; Burič & Šrol, 2020; Šrol & De Neys, 2020; Thompson et al., 2011; Thompson & Johnson, 2014). The first was response time, measured from the onset of the problem to the submitting of the answer. The second was the response confidence. After each problem, participants were asked to indicate on a scale from 0 to 10 how confident they are that their response is correct (0 = "I'm not sure at all", 10 = "I'm completely sure"). The conflict detection indices for each participant were then calculated as the difference between average time and confidence of the incorrectly solved conflict problems and correctly solved no-conflict problems (Frey, Johnson, & De Neys, 2018; Šrol & De Neys, 2020). The difference was then divided by the total number of cases from which the index could be calculated (i.e., the sum of incorrectly solved conflict problems and correctly solved no-conflict problems) to obtain an index of conflict detection capabilities for participants across the problems. The indices based on time and confidence were used as separate variables in further analyses.

Mindware instantiation and automatization. As in other studies (Frey, Johnson, & De Neys, 2018; Šrol & De Neys, 2020), the accuracy of neutral problems was used to measure *mindware instantiation*. They are called neutral because, unlike conflict and no-conflict problems, they do not elicit any heuristic answer (for an example, see online supplement).

Overall, five items were used to estimate the presence of mindware for each type of problem. If participants solved a given neutral problem correctly, it meant they possessed the basic knowledge structures needed to solve the problem.

For each neutral problem, the time needed to solve it was measured. How fast participants solved the problems was used previously by Burič and Konrádová (2021) as a rough estimate of mindware automatization, as automatization means autonomously triggered knowledge which requires a minimum time. Although the validity of this measurement was questioned even by the authors themselves, it yielded significant results in predicting correct intuitive responses in the cognitive reflection test. That is why it was also used in this study, so the new mindware automatization measurement described below could be examined in a model along with this variable.

In the current study, a different approach to measure mindware automatization was used – the neutral problems were presented via two-response paradigm.

Two-response paradigm: In this paradigm, participants were asked to solve each problem twice.

The first answer should be intuitive – the first answer that comes to mind after reading the problem. To achieve that, one needs to restrict the defining characteristics of type 2 processes – that is, the response needs to be autonomous and without engaging in working memory processing (see De Neys, 2018, for summary). Multiple restrictions, previously shown to be an effective way of knocking out type 2 processing (Bago & De Neys, 2017; De Neys, Moyens, & Vansteenwegen, 2010; Thompson et al., 2011), were adopted: instruction, a time limit (5 seconds for syllogisms and 7 seconds for the base rate tasks) resulting from the reading pretest (Burič & Šrol, 2020), and a secondary cognitive task. As a secondary cognitive task, the participants had to memorize a pattern of dots in a matrix while solving the neutral problems, a task that should burden their working memory and thus limit type 2 processing. After solving the problem, participants were asked to pick which matrix they had been previously presented out of four options. In their second response, participants were not restricted by either the time limit or the secondary cognitive tasks, so they could think freely about their final response. Unlike neutral problems, conflict and no-conflict problems were presented without any constraints.

The correct intuitive answer on a neutral problem should mean that the mindware needed to solve it is automatized. The sum of the correct intuitive responses on the neutral tasks is used as a measure of *mindware automatization*. The sum of the correct final responses is used as a measure of overall *mindware instantiation.*

Correction tendency in neutral tasks. Presenting neutral tasks under the two-response paradigm allowed me to examine the tendency to correct the erroneous first response, meaning participants submitted an incorrect intuitive answer but corrected it in the second response. I calculated this as a percentage of cases in which the erroneous answer was corrected out of the cases in which the first response was incorrect.

Procedure

First, every participant started with Block 1, which involved informed consent, general instruction, and demographic questions.

In the instruction, participants were told we are interested in the ways in which people deal with different decision-making and inference tasks and what differences exist between them in their approach to these tasks. They were informed about how much time it will take and what types of tasks they will face.

In Block 2, participants encountered the reasoning problems along with a practice problem to get familiar with the process. First, they solved syllogistic tasks, then base-rate neglect problems. This order was fixed, however, the order of items was randomized. All the participants solved both conflict and no-conflict problems.

The content of the conflict and no-conflict problems was counterbalanced, and each participant was randomly assigned to solve either Set A or Set B, in which the conflict and no-conflict items were switched (e.g., a conflict item from Set A contained the same content in Set B but was no-conflict).

After that, participants solved neutral tasks presented under the two-response paradigm in random order in Block 3. The order of the Blocks was fixed to prevent priming the analytic thinking with the neutral problems that could affect the performance in the conflict tasks (Frey et al., 2018).

Results

The main aim of this study was to analyze the relationship between mindware automatization and conflict detection. First, I present the results of reasoning accuracy and detection efficiency analysis to see whether the participants were indeed biased and whether they were still able to detect the conflict. Then I explore the correlations between the measures, and finally, I present two regression analyses to examine the predictors of reasoning accuracy and conflict detection.

Accuracy and Conflict Detection Analyses

The results of group-level analysis of reasoning accuracy and conflict detection are presented in Table 1. As predicted, participants' accuracy was higher in the no-conflict tasks when compared to the conflict ones ($t(195) = -24.8$, $p < .001$, $d = -1.77$), confirming that the conflict tasks indeed cued biased responses. Importantly, both the conflict detection measures in base-rate tasks and the confidence measure in the syllogisms show that participants were able to detect the conflict between compelling intuition and the logical structure of the tasks. The only exception is the response time for syllogisms, as it seems that participants took similarly long to solve conflict and no-conflict tasks.

Examining Reliability of Mindware Automatization Measure

Here I present the results of the analyses examining the secondary goal of the study, that is the possible use of intuitive response on neutral tasks as a reliable measure of mindware automatization. Table 2 contains the estimations of internal consistency of all the used measures.

Most of the measures reached the level of what is commonly considered to be sufficient internal consistency or were slightly below that level. Measures of mindware automatization for both syllogisms and base-rate neglect tasks reached very satisfying values of both McDonald's ω and Cronbach's α, suggesting it could be subsequently used as a reliable estimate of mindware automatization in the following analyses. Although the measures of overall mindware instantiation showed a bit lower

Table 1 *Group level analysis of reasoning accuracy and conflict detection*

	Accuracy	Response time	Response confidence
Syllogisms			
no-conflict *(SD)*	3.54 (0.69)	8.42 (4.13)	10.16 (1.29)
conflict *(SD)*	0.86 (1.21)	8.31 (3.92)	9.12 (1.64)
difference	***t(195)* = - 24.8**	*t(163)* = - 0.13	***t(195)* = 10.77**
p	**< .001**	.62	**< .001**
Cohens *d* (95% CI)	**- 1.77 (- 2; - 1.55)**	- 0.03 (- 0.18; 0.13)	**0.77 (0.61; 0.93)**
Base-rates neglect			
no-conflict	3.84 (0.47)	8.19 (3.33)	9.88 (1.63)
conflict	1.1 (1.45)	8.94 (4.02)	9.47 (1.81)
difference	***t(195)* = - 24.9**	***t(170)* = 2.59**	***t(195)* = 5.40**
p	**< .001**	**.005**	**< .001**
Cohens *d* (95% CI)	**- 1.78 (- 2; - 1.55)**	**0.20 (0.04; 0.35)**	**0.39 (0.24; 0.53)**

Note. The values reflect participants' average times of answers, confidence and average number of correctly solved problems. The reported differences represent the pairwise comparison of the average time, response confidence or accuracy between the conflict and the no-conflict version of the task. Cohen's *d* is reported as a measure of effect size. Significant paired differences are presented in bold ($p < .05$).

Table 2 *Reliability measures of all variables*

	McDonald's ω	Cronbach's α
Mindware instantiation for syllogisms	.55	.25
Mindware automatization for syllogisms	.93	.92
Mindware instantiation for base rates	.64	.63
Mindware automatization for base rates	.77	.76
Conflict detection in syllogisms – time	.68	.66
Conflict detection in syllogisms – confidence	.85	.83
Conflict detection in base rates – time	.62	.62
Conflict detection in base rates – confidence	.88	.88

internal consistency, it is important to note that these values are higher than in previous studies (Burič & Šrol, 2020; Šrol & De Neys, 2020). Finally, conflict detection indices showed relatively good internal consistency, again higher than in the previous, above-mentioned studies.

Relationship between Mindware Automatization, Conflict Detection Measure and Reasoning Accuracy

To examine the relationship of the mindware automatization, conflict detection measures,

Table 3 *Correlations between mindware measures, conflict detection measures, and the reasoning accuracy*

		1	2	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1.	Accuracy syllogisms	1												
2.	Accuracy base rates	**.45**	1											
3.	Mindware automatization – syllogisms	**.19**	.10	1										
4.	Mindware automatization – base rates	**.19**	**.29**	.08	1									
5.	Mindware instantiation – syllogisms	**.20**	.03	**.49**	**.19**	1								
6.	Mindware instantiation – base rates	**.25**	**.34**	**.16**	**.71**	**.24**	1							
7.	Syllogisms – average time of neutral tasks	-.01	**-.14**	-.02	-.02	**.20**	.06	1						
8.	Base rates – average time of neutral tasks	-.04	-.01	-.03	.05	.03	.09	**.58**	1					
9.	Syllogisms – correction tendency	**.31**	**.15**	**.54**	.13	**.86**	**.23**	**.21**	.03	1				
10.	Base rates – correction tendency	-.19	.12	.19	**.42**	.21	**.74**	**.30**	.17	.13	1			
11.	Syllogisms – conflict detection – time	**.36**	.11	.05	-.01	.12	.10	.09	.09	.15	-.12	1		
12.	Base rates – conflict detection – time	**-.15**	.02	.01	.04	.01	.10	.01	-.05	.00	.28	.02	1	
13.	Syllogisms – conflict detection – confidence	**.16**	-.01	.05	.04	**.13**	.00	-.06	-.11	.09	.07	**.38**	.08	1
14.	Base rates – conflict detection – confidence	**.13**	**.15**	.10	**.14**	**.16**	**.23**	.08	.04	.14	.27	.03	**.24**	.06

Note. Significant correlations are presented in bold ($p < .05$).

and cognitive biases, I first conducted a correlation analysis.

As can be seen from Table 3, mindware automatization measures for both syllogisms and base rates were positively related to the reasoning accuracy. However, the association between mindware instantiation and reasoning accuracy was in both cases even stronger. I also observed positive, fairly strong correlations between mindware automatization and the tendency to correct the first erroneous response. These results suggest that participants with better automatized mindware were not only more accurate in their intuitive response in neutral tasks but were also more likely to inhibit the intuitive response when it was incorrect. On the other hand, mindware automatization correlated only with the detection index for base rates based on confidence. This was also the case for conflict detection measure significantly associated with mindware instantiation, along with the index for syllogisms based on confidence. Such unclear results are in line with previous studies of Šrol and De Neys (2020) and Burič and Konrádová (2021), who also observed a relationship of mindware instantiation with just one of the multiple conflict detection indices. Despite such results, I observed a positive relationship between the conflict detection measures and reasoning accuracy, which is again in line with previous work (Burič & Konrádová, 2021; Burič & Šrol, 2020; Šrol & De Neys, 2020; Thompson et al., 2011). Namely, both indices based on time and confidence correlated with reasoning accuracy for syllogisms, even though the relationships were weak to moderate. For base rates, I observed a positive correlation just between the reasoning accuracy and the detection measure based on confidence.

Mindware Automatization as a Predictor of Conflict Detection

To examine the predictive power of mindware automatization on the conflict detection and reasoning accuracy, I conducted several hierarchical regression analyses. As can be seen in Table 4, correction tendency was not included in the model, as too few cases of correction of erroneous intuition for base rates ($N = 45$) were observed, which could decrease the statistical power of the analysis.

First, I conducted hierarchical regression analyses to test the predictive power of mindware automatization on conflict detection indices, separately with an index based on time and confidence, for both syllogisms and base rates. To test the effects of mindware automatization separately, I added the variable in the first step of regressions and the remaining variables in the second step of the regression. As can be seen in Table 4, none of the variables significantly predicted conflict detection based on time. However, the results slightly differ with the confidence index as the dependent variable. Even though mindware automatization did not emerge as a significant predictor, mindware response time did ($\beta = -.18$, $p = .039$), suggesting the less time participants needed to solve the neutral tasks, the better they were at detecting the conflict. However, the model explained only a negligible 1% of the variance with the index based on time as the dependent variable and 2.3% with the index based on confidence.

When predicting conflict detection in base-rate neglect tasks (Table 5), none of the variables emerged as significant predictors of conflict detection based on time. The only predictor worth noting is mindware automatization in the first step of the model, which predicted the detection based on confidence

Table 4 *Regression analyses with conflict detection capabilities in syllogisms as dependent variables*

	Conflict detection: time				Conflict detection: confidence			
	b (SE)	*β*	*t*	*p*	*b (SE)*	*β*	*t*	*p*
Step 1								
Constant	-.97 (.79)		-1.22	.225	.94 (.25)		3.86	<.001
Mindware automatization	1.39 (1.10)	.11	1.27	.207	.24 (.33)	.06	.71	.481
	R^2 = .01, *F* (1, 122) = 1.61, *p* = .207				R^2 = -.003, *F* (1, 144) = 0.50, *p* = .481			
Step 2								
Constant	-2.35 (1.24)		-1.90	.060	.91 (.37)		2.45	.015
Mindware automatization	.31 (1.29)	.03	.24	.809	-.14 (.39)	-.03	-.35	.724
Mindware instantiation	3.26 (2.01)	.18	1.63	.107	1.07 (.60)	.18	1.79	.076
Mindware response time	-.004 (.13)	-.003	-.03	.977	**-.08 (.04)**	**-.18**	**-2.09**	**.039**
	Δ R^2 =.02, *F* (2, 120) = 1.40, *p* = .250				Δ R^2 = .04, *F* (2, 142) = 2.92, *p* = .057			

Note. The table contains unstandardized (*b*) and standardized regression coefficients (*β*) with their respective *t*-ratio and significance. R^2 and ΔR^2 indicate adjusted *r*-square for the initial model and change in *r*-square at the 2nd step of the regression. Significant regression coefficients are presented in bold.

Table 5 *Regression analyses with conflict detection capabilities in base rates as dependent variables*

	Conflict detection: time				Conflict detection: confidence			
	b (SE)	*β*	*t*	*p*	*b (SE)*	*β*	*t*	*p*
Step 1								
Constant	.65 (.85)		.77	.446	.04 (.23)		.16	.875
Mindware automatization	.24 (1.10)	.02	.24	.809	**.54 (.27)**	**.15**	**1.98**	**.050**
	R^2 = -.01, F (1, 151) = 0.06, p = .809				R^2 = .02, F (1, 167) = 3.90, p = .050			
Step 2								
Constant	.72 (1.12)		.64	.520	-.25 (.30)		-.82	.411
Mindware automatization	-.21 (1.52)	-.02	-.14	.890	.17 (.41)	.05	.42	.679
Mindware instantiation	.78 (1.87)	.05	.42	.677	.59 (.50)	.14	1.18	.239
Mindware response time	-.09 (.14)	-.05	-.64	.521	.03 (.04)	.05	.70	.484
	Δ R^2 =.003, F (2, 149) = 0.26, p = .774				Δ R^2 = .01, F (2, 165) = 1.09, p = .338			

Note. The table contains unstandardized (*b*) and standardized regression coefficients (*β*) with their respective *t*-ratio and significance. R^2 and ΔR^2 indicate adjusted *r*-square for the initial model and change in *r*-square at the 2nd step of the regression. Significant regression coefficients are presented in bold.

Table 6 *Regression analyses with reasoning accuracy in syllogisms and base-rate neglect tasks as dependent variables*

	Syllogisms				Base rates			
	b (SE)	*β*	*t*	*p*	*b (SE)*	*β*	*t*	*p*
Step 1								
Constant	.30 (.23)		1.32	.189	.04 (.32)		.14	.893
Mindware automatization	**.92 (.32)**	**.25**	**2.88**	**.005**	**1.47 (.38)**	**.30**	**3.82**	**< .001**
	R^2 =.06 , F(1, 122) =8.31 , p =.005				R^2 = .08, F(1, 151) = 14.60, p < .001			
Step 2								
Constant	.17 (.36)		.47	.638	-.09 (.42)		-.22	.828
Mindware automatization	.58 (.36)	.16	1.63	.106	.49 (.57)	.10	.87	.385
Mindware instantiation	.60 (.57)	.11	1.07	.289	**1.51 (.70)**	**.25**	**2.17**	**.032**
Mindware response time	-.01 (.04)	-.03	-.30	.763	-.09 (.05)	-.14	-1.77	.080
Conflict det.: time	**.09 (.03)**	**.31**	**3.48**	**.001**	-.04 (.03)	-.09	-1.17	.243
Conflict det.: confidence	.02 (.08)	.02	.24	.814	.15 (.11)	.11	1.37	.174
	ΔR^2 = .12, F(4, 118) = 4.25 , p =.003				ΔR^2 = .06, F (4, 147) = 2.38, p = .054			

Note. The table contains unstandardized (b) and standardized regression coefficients (β) with their respective t-ratio and significance. R^2 and ΔR^2 indicate adjusted r-square for the initial model and change in r-square at the 2nd step of the regression. Significant regression coefficients are presented in bold.

(β = .15, p = .050), but even this variable lost its predictive power in the next steps of the regression.

Incremental Validity of Mindware Automatization when Predicting Reasoning Accuracy and Conflict Detection

Finally, in the following two regressions, I tested incremental validity of mindware automatization with reasoning accuracies as dependent variables.

First, as I wanted to test the predictive power of mindware automatization separately from the other variables, I entered the variable in the first step of regression. In both regressions, mindware automatization emerged as a significant predictor, explaining 6% of the variance in syllogisms and 8% of the variance in base-rate neglect tasks. In the second step of the regressions, I added the mindware instantiation measure, the time needed to solve the neutral tasks, and the conflict detection measures. In syllogisms, only the conflict detection index based on time emerged as a significant predictor (β = .31), as the model explained 15% of the variance. In base-rate neglect tasks, the only variable that passed the conventional significance level was mindware instantiation, which showed relatively good predictive power (β = .25) with p = .032. Overall, the model explained 11% of the variance for base-rate neglect tasks.

Discussion

The main goal of this study was to examine Stanovich's (2018) hypothesis of automatized mindware influencing subsequent conflict detection. A necessary step to achieve this goal was to find a reliable measure of mindware automatization, which was a secondary goal of the study. I tried to achieve this by isolating the type 1 response in neutral problems via a two-response paradigm. I am fully aware that the measure of internal consistency may not be sufficient to draw conclusions and that more measures should be used. However, the measures of both McDonald's ω and Cronbach's α can serve as a good first estimate of reliability. The internal consistency scores for mindware instantiation were higher than in previous studies (Burič & Šrol, 2020; Šrol & De Neys, 2020) and overall, it seems that increasing the number of items and measuring mindware separately for each type of reasoning problem has indeed improved the reliability of the measure.

After estimating the measures' reliability, I examined whether conflict detection indeed took place in the conflict problems. Traditional group-level analysis of the detection confirmed the results of previous studies (Bago & De Neys, 2017, 2019; Bonner & Newell, 2010; Burič & Šrol, 2021; De Neys et al., 2010, 2011, 2013; Thompson & Johnson, 2014). The measures of conflict detection showed that participants indeed detected the conflict, as they reported significantly lower confidence levels, and in case of base-rate neglect tasks also took more time to solve the conflict problems, when compared to no-conflict ones. After confirming the presence of conflict detection, I moved to examining the relationship between the measured variables.

Mindware as a Key Element in Bias Susceptibility

The first step to examine the hypothesis about the relationship between mindware automatization, conflict detection, and reasoning accuracy was the correlation analysis: the results showed a significant positive association between the reasoning accuracies of both types of reasoning problems, mindware automatization, and mindware instantiation. The results suggest that the better instantiated and automatized mindware indeed led to higher reasoning accuracy, which confirms Stanovich's (2018) model of mindware automatization as well as the results of previous studies (Burič & Šrol, 2020; Frey et al., 2018; Šrol & De Neys, 2020). Both instantiation and automatization also strongly correlated with the correction tendency, meaning the better instantiated and automatized the mindware was, the more likely participants were to correct their erroneous intuitive response. This supports the proposal of the logical intuition model (Bago & De Neys, 2017; De Neys, 2012), according to which reasoners generate multiple intuitions based both on heuristics and logical principles. If one of them dominates the other, but the relative difference of intensities is small, the response may change, as there will be doubt, leading to rethinking. As better adopted mindware should contribute to higher intensity of logical intuition, the relative difference should decrease, which should lead to correction of the initial response.

The subsequent regression analyses did not confirm the results of the correlation analysis. Even though mindware automatization did significantly predict reasoning accuracy in the first steps of the regressions, the predictive power decreased in the second step and the significance level did not meet the conventional criteria. In the second step, only the conflict detection based on time emerged as a significant predictor of reasoning accuracy in syllogisms. For base-rate neglect tasks, only mindware instantiation predicted the reasoning performance. These results not only do not support my initial hypothesis but they also differ from previous studies showing the predictive power of both mindware instantiation and detection capabilities (Burič & Šrol, 2020; Frey et al., 2018; Šrol & De Neys, 2020). There are exceptions, as the study of Burič and Konrádová (2021) also showed no predictive power of the detection measures. However, one has to keep in mind that the studies using two-response paradigm (e.g., Bago & De Neys, 2017, 2019; Burič & Konrádová, 2021; Raoelison et al., 2020) also showed that logically correct initial response will usually lead to a logically correct final response, while deliberate correction could also improve initially incorrect responses. This could lead to mindware instantiation being more potent than mindware automatization in predicting the reasoning accuracy. The results of correlation analysis also support this, as mindware instantiation showed slightly stronger correlations with the reasoning accuracy.

Automatized Mindware and Conflict Detection

As a next step, I examined the relationship between mindware measures and conflict detection. Firstly, not all the detection indices correlated with mindware measures. Mindware automatization correlated positively only with the detection index based on confidence in base-rate neglect tasks, but the relationship was quite weak. Mindware instantiation correlated only with confidence indices, which is in line with studies of Burič and Konrádová (2021) and Šrol and De Neys (2020), but again, the correlations were weak. The authors argued that this could be caused by low reliability of the used measures, but this was not the case of the present study, as the measures showed sufficient internal consistency. But again, as mentioned above, the internal consistency analyses might not be enough to conclude the measure is reliable. However, as follows from the model of Bago and De Neys (2017), when the intensity of logical intuition is much higher than the intensity of heuristic intuition, these two will not interfere, thus the chance of conflict detection can in fact decrease. This could be exactly the case for the participants who had the mindware fully automated.

Finally, I examined predictors of detection capabilities, as Stanovich (2018) states that the better the mindware is instantiated, the easier the conflict detection should be. For syllogisms, this study provides very weak support for such a claim, as neither mindware instantiation nor automatization predicted conflict detection. Only the mindware response time emerged as a significant predictor, but only of detection based on confidence. Similar results were found when predicting the detection capabilities in base-rate neglect tasks, as mindware automatization was the only significant predictor of detection based on confidence, but only in the first step of the analysis. I can again compare these results to study of Šrol and De Neys (2020), where the results showed the predictive power of mindware instantiation on conflict detection capabilities, but only in the measure based on confidence. However, the point mentioned above also applies here. That is, the conflict detection might decrease due to the dominant logical intuition.

Overall, the results provide very little support for Stanovich's (2018) model. It should be stressed that the author proposes that in cases when the mindware is automatized, conflict detection should be the easiest. To sum up the point from above, this is partially in line with the hybrid models of dual processes, for example, the logical intuitions model suggesting that the small relative difference between heuristic and logical intuition should lead to easier conflict detection and possible rethinking of the response. However, only Stanovich's model explicitly specifies the role of conflict detection for reasoners with fully automatized mindware. Translated into the jargon of the logical intuitions model, the question remains whether conflict detection is possible at all when the logical intuition is way more dominant than the heuristic one. Conflict detection itself means that there is interference between heuristic intuition and the logical principles of the given tasks that one has acquired (e.g., Bago & De Neys, 2017; Frey et al., 2018). However, if the principles are adopted to such an extent that they are fully automatized and the autonomous response is consequently logically correct, it is questionable whether interference occurs at all, as the intuition based on logic fully dominates the other. To be clear, there are studies showing that conflict detection takes place even while responding intuitively (e.g., Bago & De Neys, 2017; Burič & Šrol, 2020; Thompson & Johnson, 2014); these studies, however, did not examine its relationship with mindware, nor the detection of reasoners with fully automatized mindware.

On that note, one needs to also consider validity of conflict detection measures themselves. The indices are computed as a difference in confidence and response time between conflict and no conflict items, however, just from the cases when the conflict item response was incorrect, i.e., participant fell for the belief bias. This approach is used due to the assumption that unbiased response includes not only the conflict detection, but also inhibition of incorrect response, which could be reflected, for example, in an even longer response time. However, as one of the reviewers pointed out, this means there might be some variance that is not captured, especially for reasoners with higher overall accuracy in neutral tasks. As these reasoners with highly automated mindware are assumed to only make few mistakes in conflict tasks, it might be the case the variance from these reasoners is not captured with the measures of conflict detection. This also implies a strong relationship between mindware automatization and reasoning accuracy, which was not the case in this study. However, the weak correlations could be caused by various reasons. With this in mind, it would be worth trying alternative ways to capture conflict detection, such as ECG (Mevel et al., 2019) or skin conduction measures (De Neys et al., 2010).

I propose that future studies should focus on whether the relationship between the mindware and conflict detection is linear or not. According to Stanovich, the mindware is a set of cognitive rules and strategies that can help people overcome biases and heuristics in reasoning. He also suggests that the chance of conflict detection, which is the ability to recognize when one's intuitive response is incorrect, increases with the extent of mindware adoption. However, as I argue above, the question is whether detection is necessary or even possible for reasoners who have fully automatized their mindware. In other words, if the mindware becomes so ingrained that it overrides the intuitive response automatically, then there may be no need or opportunity for conflict detection. Also, as suggested by one of the reviewers, when mindware is not acquired, conflict cannot be detected and as a result one can be more confident with a heuristic response. When mindware is acquired, reasoners could actually become less confident as this time they are able to detect the conflict. However, when the mindware is automated, reasoners can be (rightfully) very confident with their responses again, since the problem is perceived as an easy one. Thus, the relationship between mindware and conflict detection could be U-shaped, rather than linear. This possibility of a non-linear nature of the relationship between conflict detection and mindware instantiation would explain why the results of the present study do not support the hypothesis that the mindware automatization would lead to easier conflict detection. Therefore, I suggest that future research should investigate this issue further and use also non-linear methods to test such hypothesis.

In addition, it should be noted that the present study did not examine the possible effects of other known correlates of bias susceptibility and conflict detection, such as cognitive ability (Šrol & De Neys, 2020; Thompson & Johnson, 2014), numeracy (Klaczynski, 2014; Šrol & De Neys, 2020), reasoning strategies (Markovits et al., 2020) or cognitive reflection (Burič & Šrol, 2020; Šrol & De Neys, 2020). All of these can impact the relationship between mindware measures, conflict detection, and bias susceptibility. Therefore, I suggest that the present study should be replicated with all of these measures to get a clearer picture of such a relationship.

The field of cognitive biases is indeed very popular, yet our knowledge of the processes leading to such biases is still very limited. This study aimed to broaden that knowledge by examining the relationship between mindware automatization and conflict detection. Even though the results do not fully support the assumptions of Stanovich's (2018) model, mindware automatization still needs to be examined in a broader context of cognitive biases predictors.

Acknowledgmnets

The research presented here is based on an unpublished doctoral dissertation of the author. The study was supported by the scientific grant agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic as part of the project VEGA 2/0053/21: *"Examining unfounded beliefs about controversial social issues".* The data for this study are publicly available at: https://osf.io/6kxsr/.

Author's ORCID
Roman Burič
https://orcid.org/0000-0002-3762-3354

References

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

Bago, B., & De Neys, W. (2019). The Smart System 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning*, *25*(3), 257–299. https://doi.org/10.1080/13546783.2018.1507949

Burič, R., & Konrádová, Ľ. (2021). Mindware instantiation as a predictor of logical intuitions in the Cognitive Reflection Test. *Studia Psychologica*, *63*(2), 114–128. https://doi.org/10.31577/sp.2021.02.822

Burič, R., & Šrol, J. (2020). Individual differences in logical intuitions on reasoning problems presented under two-response paradigm. *Journal of Cognitive Psychology*, 1–18. https://doi.org/10.1080/20445911.2020.1766472

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., & Bonnefon, J.-F. (2013). The 'whys' and 'whens' of individual differences in thinking biases. *Trends in Cognitive Sciences*, *17*(4), 172–178. https://doi.org/10.1016/j.tics.2013.02.001

De Neys, W., Moyens, E., & Vansteenwegen, D. (2010). Feeling we're biased: Autonomic arousal and reasoning conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 208–216. https://doi.org/10.3758/CABN.10.2.208

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Frey, D., Johnson, E. D., & De Neys, W. (2018). Individual differences in conflict detection during reasoning. *Quarterly Journal of Experimental Psychology*, *71*(5), 1188–1208. https://doi.org/10.1080/17470218.2017.1313283

Klaczynski, P. A. (2014). Heuristics and biases: Interactions among numeracy, ability, and reflectiveness predict normative responding. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00665

Klaczynski, P. A., & Felmban, W. (2020). Effects of thinking dispositions, general ability, numeracy, and instructional set on judgments and decision-making. *Psychological Reports*, *123*(2), 341–370. https://doi.org/10.1177/0033294118806473

Markovits, H., de Chantal, P.-L., Brisson, J., Dubé, É., Thompson, V., & Newman, I. (2020). Reasoning strategies predict use of very fast logical reasoning. *Memory & Cognition*. https://doi.org/10.3758/s13421-020-01108-3

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, *17*(1), 11–17. https://doi.org/10.3758/BF03199552

Mevel, K., Borst, G., Poirel, N., Simon, G., Orliac, F., Etard, O., Houdé, O., & De Neys, W. (2019). Developmental frontal brain activation differences in overcoming heuristic bias. *Cortex*, *117*, 111–121. https://doi.org/10.1016/j.cortex.2019.03.004

Raoelison, M., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition*, *204*, 104381. https://doi.org/10.1016/j.cognition.2020.104381

Sá, W. C., 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. https://doi.org/10.1037/0022-0663.91.3.497

Šrol, J., & De Neys, W. (2020). Predicting individual differences in conflict detection and bias susceptibility during reasoning. *Thinking & Reasoning*, 1–31. https://doi.org/10.1080/13546783.2019.1708793

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

Thompson, V. A., & Johnson, S. C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning*, *20*(2), 215–244. https://doi.org/10.1080/13546783.2013.869763

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140. https://doi.org/10.1016/j.cogpsych.2011.06.001

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. https://doi.org/10.3758/s13421-011-0104-1