Supplementary material

Section A

Conflict and no-conflict versions of the CRT items – English version:

Conflict versions	No-conflict versions
10 Christmas elves take 10 minutes to wrap 10 presents. How long would it take 50 Christmas elves to wrap 50 presents? *Answers: 50/10/5*	The printer takes 5 minutes to print 5 magazines. How many minutes would it take the printer to print 100 magazines? *Answers: 10/5/100*
The number of virus victims doubles every day. If a virus infects an entire city in 50 days, in how many days will it infect half of the city? *Answers: 25/40/49*	The amount of bacteria in the dish grows evenly every day. If it takes them 48 days to cover the whole dish, in how many days will they cover half the dish? *Answers: 24/47/12*
A man fell into a 20-meter well. During the day he climbs 5 meters, at night he slides 4 meters back. How many days will it take him to climb out of the well? *Answers: 10/16/20*	A man weighing 100kg wants to get to 90kg. If he loses 1kg in a week, how many weeks will it take him to get to 90kg? *Answers: 10/9/8*
An apple and an orange weigh 160 grams in total. The apple weighs 100 grams more than the orange. How much does the orange weigh? *Answers: 60/30/130*	The handyman and the electrician work for 240 days in total. The electrician works for 200 days. How many days does the handyman work? *Answers: 160/120/40*

Jakub is 4 years old and his sister Miška is half his age. How old will Miška be when Jakub is 40?

Answers: 32/30/38

When the old man was 20 years old, his future wife was 10 years younger. How old is his wife when the old man is 100 years old?

Answers: 130/100/90

Conflict and no-conflict versions of the CRT items – Slovak version:

Conflict versions	**No-conflict versions**
10 vianočným škriatkom trvá 10 minút zabaliť 10 darčekov. Koľko minút by trvalo 50 vianočným škriatkom zabaliť 50 darčekov? *Odpovede: 50/10/5*	Tlačiarni trvá 5 minút vytlačiť 5 časopisov. Koľko minút by trvalo tlačiarni vytlačiť 100 časopisov? *Odpovede: 10/5/100*
Počet obetí vírusu sa každý deň zdvojnásobí. Ak vírus za 50 dní nakazí celé mesto, za koľko dní nakazí polovicu mesta? *Odpovede: 25/40/49*	Množstvo baktérii v miske rastie každý deň rovnomerne. Keď im trvá 48 dní pokryť celú misku, za koľko dní pokryjú polovicu misky? *Odpovede: 24/47/12*
Muž spadol do 20 metrovej studne. Počas dňa vylezie 5 metrov, v noci skĺzne 4 metre späť. Za koľko dní zo studne vylezie? *Odpovede: 10/15/20*	Muž vážiaci 100kg chce schudnúť na 90. Ak za týždeň schudne 1kg, koľko týždňov mu zaberie, kým schudne na 90kg? *Odpovede: 10/9/8*
Jablko a pomaranč spolu vážia 160 gramov. Jablko váži o 100 gramov viac, ako pomaranč. Koľko váži pomaranč? *Odpovede: 60/30/130*	Údržbár a elektrikár pracujú dokopy 240 dní. Elektrikár pracuje 200 dní. Koľko dní pracuje údržbár? *Odpovede: 160/120/40*
Jakub má 4 roky a jeho sestra Miška je od neho o polovicu mladšia. Koľko rokov bude mať Miška, keď	Keď mal starec 20 rokov, jeho budúca manželka bola o 10 rokov mladšia. Koľko rokov má jeho manželka,

Jakub bude mať 40 rokov?

Odpovede: 32/30/38

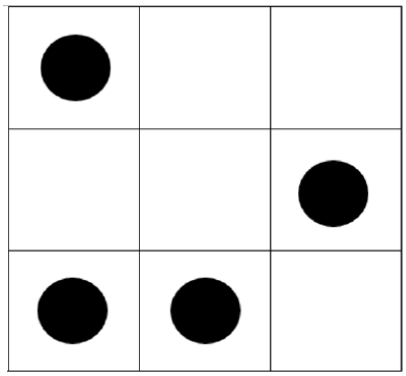
keď má starec 100 rokov?

Odpovede: 130/100/90

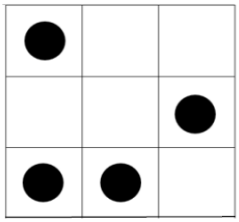
Section B

Example of the secondary cognitive task

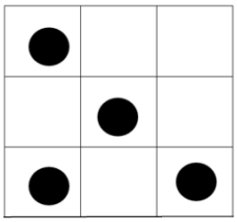
Participants had to remember patterns like this during the initial response.



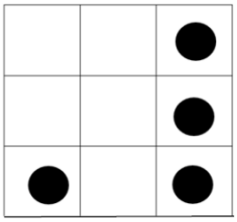
After the response, they were asked to pick the correct pattern from four options.



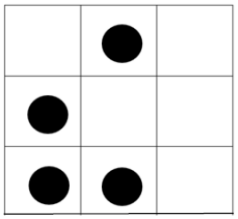
A



B



C



D

Section C

Table S1. Examples of neutral problems used as mindware instantiation measure

CRT items	Neutral problem based on the item
10 Christmas elves take 10 minutes to wrap 10 presents. How long would it take 50 Christmas elves to wrap 50 presents?	Calculate *y,* if: 10 * 10x = 10 50y * x = 50 y = ?
A man fell into a 20-meter well. During the day he climbs 5 meters, at night he slides 4 meters back. How many days it will take him to climb out of the well?	Calculate *x:* (5 – 4)x = (20 – 4)
The number of virus victims doubles every day. If a virus infects an entire city in 50 days, in how many days will it infect half of the city?	Mindware for this item was not measured
An apple and an orange weigh 160 grams in total. The apple weighs 100 grams more than the orange. How much does the orange weigh?	Calculate *x*, if: 100 + 2x = 160
Jakub is 4 years old and his sister Miška is half his age. How old will Miška be when Jakub is 40?	Calculate *y*, if *x* = 40 x = y + 13

Section D

Calculation of the mindware automatization index

To put the number of correctly solved neutral items and the time needed to correctly solve them in a relationship, we came up with an equation *(1 - t) * n*, where:

t = average time of correctly solved neutral problems,

n = number of correctly resolved problems

For example, if a participant had an average time of correctly solved neutral problems of 16 seconds and he solved 10 problems correctly, the calculation would be (1 - .16) * 10 = 8.4. The index would thus rise if the average time decreased and vice versa. In this case, the index representing the degree of mindware automatization equals 8.4. The mindware automatization index significantly correlated with the number of correct responses (r = .72; p < .001), as well as with the time required to respond (r = - .63; p < .001). In this respect, the index met the criteria - it increased depending on the growing number of correct mindware responses and the decreasing time required to respond. However, after consultation with Dr. Jakub Šrol, we finally decided not to use it in further analyses for the following reasons. The role of correctly solved problems is clear in terms of mindware instantiation - the more correctly solved tasks the better instantiated mindware. However, its relationship with the response time does not seem to be so clear. The response time of mindware problems correlated positively with their accuracy, even though this relationship was very weak (r = .10; p = .04). Also, in regression analyses, the index enabled us, significantly, to explain the lower amount of variance of correctly solved CRT problems (β = .32; p < .001 for intuitive and β = .34; p < .001 for final responses), than if we had used the accuracy and the time needed to solve the neutral problems as two separate predictors.

Section E

More information about the data distribution

Table S4. Kolmogorov – Smirnov test of normality, skewness, and kurtosis of data distribution of measured variables

	D	*p*	*Skewness*	*Kurtosis*
Initial accuracy	.113	< .001	.248	- .146
Final accuracy	.152	< .001	- .784	2.391
Initial response time	.041	.200	- .106	.792
Initial response confidence	.083	.002	.676	.058
Final response time	.246	< .001	1.869	6.626

Final response confidence	,129	< .001	.401	- .752
Accuracy on neutral problems	.137	< .001	1.327	2.48
Time of correctly solved neutral problems	.197	< .001	2.278	6.304
Conflict detection index – time of initial response	.370	< .001	- .863	- .960
Conflict detection index - confidence of initial response	.381	< .001	.799	- 1.113
Conflict detection index – time of final response	.147	< .001	- .286	- 1.024
Conflict detection index - confidence of final response	.268	< .001	.696	- .627

Section F

Individual differences in conflict detection

According to Stanovich (2018), the degree of mindware automatization is in direct relationship with conflict detection ability and ability to inhibit the type 1 answer, postulating that the better instantiated mindware the easier conflict detection and inhibition of type 1 answer and that the fully automatized mindware can even lead to the normatively correct type 1 response. His model differentiates five degrees of mindware automatization resulting in various possibilities of cognitive failures.

As we gathered all the data needed to test this model, we decided to examine the model directly and compared multiple categories based on the degree of mindware instantiation. However, the first degree of the Stanovich's model includes reasoners with missing mindware and just two of our participants failed in all of the neutral problems. Therefore, we have decided to divide our sample into quartiles based on the number of correctly solved neutral problems and the time needed to correctly solve them, and to compare the first and the fourth quartile in terms of conflict detection ability at initial and final response stage, to get at least a hard estimate of the fit with Stanovich's model.

The results showed, that participants with more correctly solved neutral problems (n = 33) had significantly higher conflict detection index based on the time at the initial response stage (M = 0.82; SD = 0.35), than participants with fewer correctly solved problems (n = 59; M = 0.64; SD = 0.43; t (78.8) = -2.25; p = .027; d = 0.47). Similary, participants who needed less time to correctly solve neutral problems (n = 55) showed higher conflict detection ability based on the time of the initial response (M = 0.78; SD = 0.39) than participants with higher average time (n = 56; M = 0.55; SD = 0.43; t (109) = - 2.99; p = .003; d = 0.57). Significant differences in other detection indices were not observed.

The results indicate, that participants with better instantiated mindware achieved higher indices of conflict detection. However, these were only intuitive responses and only a detection index based on response time, not confidence. Our ambiguous results are thus consistent with regression analyses in which only insignificant predictive ability of conflict detection was observed. These findings lead us once again to the need for a better examination of conflict detection itself and potentially to the need for more reliable measurement.

Section G

Item – level analysis: Number of initial responses exluded from the analysis and percentage of correctly solved conflict problems per item for initial and final response

Table S2. The number of valid answers to conflict problems at the initial response stage per item that passed through all of the exclusion criteria and the number of excluded answers.

	CRT 1	CRT 2	CRT 3	CRT 4	CRT 5
Valid *(n)*	175	173	216	311	251
Excluded *(n)*	733	735	692	597	613

Note. A response could be excluded for several reasons - participants did not respond within the time limit, responded incorrectly to the secondary cognitive task associated with the item, or simply did not provide an answer at all.

Table S3. Percentage of correctly solved conflict problems per item for the initial and final response

	CRT 1	CRT 2	CRT 3	CRT 4	CRT 5
Initial	44 %	31.8 %	41.2 %	17.7 %	50.2 %
Final	52 %	45.1%	15.9%	50 %	83.9 %

Supplementary references:

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

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

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