

# 16

## The Impact of the Length and Solvability of Anagrams on Performance and Metacognitive Judgments

**Pavle Valerjev**

*Department of Psychology, University of Zadar, Zadar, Croatia*

**Marin Dujmović**

*School of Psychological Science, University of Bristol, UK*

### Abstract

*Anagrams are commonly used in the field of problem solving since they provide a number of possible experimental manipulations such as length, difficulty, and solvability. Recent trends in cognitive psychology emphasize the importance of metacognitive processes which accompany human reasoning, decision making and problem solving. In this study, our goal was to measure performance and metacognitive judgments while manipulating the length and solvability of anagrams. Participants were in general faster for shorter and for solvable anagrams. Additionally, the difference between shorter and longer anagrams was significantly larger for solvable compared to unsolvable anagrams. The only significant effect on accuracy was the length with a higher accuracy for shorter anagrams. For judgments of solvability length was again the only significant factor with shorter anagrams initially judged as more solvable. Finally, the analysis of post-trial judgments of difficulty showed both effects were significant. Solvable and shorter anagrams were judged as easier in general. However, the difference in judgments of the difficulty between shorter and longer anagrams was significant only for solvable anagrams. In total, participants seem to rely on extremely salient cues (length) when making initial metacognitive judgments while post-trial judgments are impacted by more factors. Further experiments should provide a more in-depth study of the differences depending on solvability and accuracy (e.g. differences in metacognitive judgments between correctly solved solvable anagrams and correctly recognized unsolvable anagrams).*

*Keywords : problem solving, anagrams, metacognition, unsolvable problems*

## Introduction

Metacognition is traditionally described as cognition about cognitive processes, or as “knowledge about knowledge” (Flavell, 1979). Furthermore, the concept of metacognition can be divided into metacognitive knowledge and metacognitive regulation. Metacognitive knowledge is described as knowledge and beliefs which have an impact on specific cognitive processes. For example, knowledge about particular abilities and strategies which can be helpful while performing a specific task. On the other hand, metacognitive regulation is a concept which includes both metacognitive monitoring and metacognitive control. Metacognitive monitoring includes processes which provide information about the efficiency of cognitive processes involved in some task. Many circumstances and factors, both related and not related to the task, may influence the accuracy of the monitoring feedback when compared to the actual efficiency while solving a specific task. It is important to note that metacognitive monitoring includes processes which occur before, during and after the solving of a particular cognitive task. Finally, metacognitive control is related to different types of active control of cognitive processes based on metacognitive knowledge and monitoring. More specifically, it usually includes planning, choosing, the use and changing of strategies involved in a particular kind of cognitive task (Nelson & Narens, 1990).

One of the key early endeavors of metacognitive scientists was to investigate and describe the mechanisms which underlie metacognitive monitoring and control, as well as the development of useful research paradigms. This kind of research can target different cognitive processes. Traditionally, impactful early metacognitive research was conducted in the area of memory (meta-memory) where research paradigms and metacognitive parameters were developed (Nelson & Narens, 1990). Following this tradition, by adapting the methodology and parameters adequately, the metacognitive approach was incorporated into reasoning research, as well as research of other forms of thinking. Consequently, the meta-reasoning framework was developed in order to investigate the metacognitive processes which occur in parallel and/or in relation to different reasoning and, more generally, thinking processes. For a more detailed review of the meta-reasoning framework, see work by Ackerman & Thompson (2015; 2017). In meta-reasoning research, metacognitive parameters can be measured before (e.g. judgment of solvability), during (e.g. feeling of rightness; initial judgment of confidence) or after (e.g. final judgment of confidence) the completion of reasoning processes. A brief summary of the main meta-reasoning findings includes the following:

1. Actual performance is rarely predictive of confidence and other metacognitive ratings, and in addition participants tend to show systematic overconfidence (Bajšanski, Močibob, & Valerjev, 2014; Dujmović & Valerjev, 2018; Thompson & Johnson, 2014; Thompson, Prowse Turner, & Pennycook, 2011; Valerjev & Dujmović, 2017).
2. Metacognitive judgments are often better predicted by the so-called indirect metacognitive cues such as answer fluency and detection of conflict between two or more answers. Response times, which are the measure of answer fluency, usually moderately to highly correlate with metacognitive judgments (Ackerman & Zalmanov, 2012; Thompson, Evans, & Campbell, 2013; Thompson, Prowse Turner, et al., 2013).
3. Conflict detection and resolution is another important cue that may significantly impact metacognitive judgments independently of fluency (Dujmović & Valerjev, 2018).
4. Lower initial and on-going metacognitive judgments are correlated with a higher probability of answer change when prompted to rethink the response (Shynkaruk & Thompson, 2006; Thompson & Johnson, 2014; Thompson, Prowse Turner et al., 2013).
5. It has been demonstrated that significant inter-individual differences in sensitivity to conflict detection and the modus of its resolution exist (Dujmović & Valerjev, 2018; Frey, Johnson, & De Neys, 2018; Mevel et al., 2015).

These findings can be considered separate from broader models of reasoning, but they also fit well into the dual-process approach to reasoning (for this approach see Evans, 2008; Evans & Stanovich, 2013) according to which, our reasoning is governed by one or more rapid, heuristic, automatic, intuitive processes which may (if there is more than one of them), end in conflicted outcomes. Sometimes this conflict is resolved by another, more analytical, deliberate and slower process.

Many reasoning tasks in meta-reasoning research are constructed in a way to exploit the conflict between the so-called intuitive (or heuristic) and analytical answers which are found in dual-process research. Typical dual-processing tasks include formal logic tasks such as categorical syllogisms (Thompson & Johnson, 2014; Thompson & Morsanyi, 2012), base rate neglect task (De Neys & Glumicic, 2008; Dujmović & Valerjev, 2018; Pennycook, Fugelsang & Koehler, 2015), items from Cognitive Reflection Test (Toplak, West & Stanovich, 2014) etc. Most of these can traditionally be categorized as reasoning tasks while some fall within the field of judgment and decision-making (JDM). The distinction between reasoning and JDM tasks is not always clear since some of them are considered as members of both categories, but JDM tasks always incorporate making a choice based on some aspect of statistical reasoning and/or subjective probabilities (e.g. base rate task or Linda problem). Reasoning and JDM tasks are usually solved in one or two steps after one or more answers are generated by means of heuristic and/or analytical processes. On the other hand, Ackerman (2014) described goal-driven tasks which were investigated in the rich tradition of problem-solving research in psychology. Goal-driven tasks include various problem tasks such as the *hobbits and orcs* problem, the Tower of Hanoi problem, chess problems, anagrams etc. These problems cannot be solved in a single step. Problem tasks are usually solved in a series of steps during which the solver compares his current state with the goal state. The goal-driven nature of problem solving is derived from Newell & Simon's (1972) cognitive theory of problem solving according to which problem solving may be illustrated as the search for a path through the problem space. The path links the initial state and final (goal) state of the problem space through a series of interstates. Moving from one state to another is achieved through the use of allowed operators. An example would be the movement of discs in the Tower of Hanoi problem by adhering to the rule that a larger disc cannot be placed on top of a smaller one and that only one disc at a time is allowed to be moved from peg to peg. The problem solver has the mental representation of his current state and its neighborhood, as well as the mental representation and the estimated distance to the final state. General problem-solving strategies, such as *hill climbing* and *means-ends analysis*, are based on the solver choosing the next move by evaluating which action would lead to a state closest to the goal. When choosing the operator, and by proxy, the next state, both heuristic and analytical processes may be activated, as well as more complex cognitive processes such as planning and evaluation of strategies.

### **Metacognition in problem-solving**

In an early paper which dealt with metacognitive aspects of problem solving, Metcalfe (1986a) presented subjects with various insight problems and asked them to estimate the feeling-of-knowing judgment for each problem after a short presentation. After that, the participants tried to solve the presented set of problems. For comparison, in Experiment 2 participants estimated feeling-of-knowing judgments for memory tasks. Participants could predict their memory performance well, but their metacognitive predictions in the insight problem solving were nonexistent. In another paper, Metcalfe (1986b) asked participants to provide on-going metacognitive judgments which are labeled feeling-of-warmth. Participants had to estimate how warm, or how close they are to the solution in regular time intervals during their work on a given problem. Among five experiments, two used various problem tasks and three used anagrams. On-going feelings of warmth showed a very slow increase and then suddenly jumped to maximum values when the solution was reached, and this pattern is typical for insight problem solving. Metcalfe and Wiebe (1987)

compared on-going feelings of warmth for insight and non-insight problems. The results indicated that warmth judgments increased incrementally for non-insight problems, but not for insight problems. This is because insight problems are usually solved in a moment of sudden illumination and it is hard to estimate the closeness to a solution before that moment.

Important work on metacognitive judgments that accompany thinking processes during problem-solving behavior was done by Ackerman (2014). She proposed the Diminishing Criterion Model which elegantly describes how both on-going metacognitive judgments and the final criterion that problem solvers strive to reach change during goal-driven thinking tasks. As more and more time is invested in problem solving, on-going metacognitive judgments (such as intermediate judgments of confidence) increase. However, at the same time, the compromise on the acceptable final goal (the criterion) also increases, which in turn decreases the final criterion. When the criterion is reached the problem-solving behavior stops. Because of that, final judgments of confidence in goal-driven tasks are still negatively correlated with solving time which was also obtained in a second study (Ackerman & Zalmanov, 2012).

## **Anagrams and metacognitive judgments**

Anagrams are scrambled nonsense strings of letters which have to be rearranged in order to form a word. Anagram solving tasks have been used in studies spanning different areas of cognition and language from research on the mappings between visual and lexical representations (Witte & Freund, 2003) to memory processes (Bernstein, Rudd, Erdfelder, Godfrey, & Loftus, 2009) and problem-solving strategies (Novick & Sherman, 2003). It has been shown that anagrams typically get solved in one of two ways. First, the solution presents itself through what seems like insight, people are not aware of the stepwise process (Novick & Sherman, 2003), although this depends on the structure of the anagram itself. The second type of solution results from what would be a classical problem-solving process in which participants rearrange the letters in a serial manner until they solve the problem. Studies have found that many different aspects influence the success in anagram solving tasks. These aspects include both the characteristics of the anagram and the traits of the solvers, as well as their interaction. For example, reading fluency has been shown to predict success of anagram solving in children (Deloche, Ott, & Tavella, 1995; Sarris & Panagiotakopoulos, 2013). Anagram length seems to have a large influence on the success but also on the process of solving anagrams. Longer anagrams are generally more difficult to solve (Kaplan & Carvellas, 1968; Gilhooly & Johnson, 1978) and a review has shown that the number of syllables seems to be at the root of the influence length has on difficulty, rather than simply varying the number of characters (Muncer & Knight, 2010). Some characteristics of the anagrams have counter-intuitive effects on difficulty. For example, everything else held equal, anagrams easier to pronounce are more difficult to solve than anagrams which are harder to pronounce (Mayzner & Tresselt, 1958; Novick & Sherman, 2008). This implies strategies which involve reorganizing chunks rather than single letters, which draws similarities with heuristics in reasoning and meta-reasoning research.

The metacognitive aspects of anagram solving have been a target of a small number of studies so far. Metcalfe (1986b) measured ongoing ratings of warmth during anagram solving in three experiments. She found that correctly solved anagrams usually have a pattern of ratings which indicate an insight-based solution. The warmth ratings do not increase sharply until the final solution is given. On the other hand, incorrectly solved anagrams have a larger proportion of patterns which suggest incremental increase when compared to the correctly solved anagrams. Additionally, two of the experiments showed that warmth ratings for incorrect solutions were higher towards the end of the problem-solving process when compared to the correct ones. This was attributed to participants using a satisficing strategy, and in the experiment where that was encouraged the effect was larger. The strategy allows for *good enough*, inelegant solutions

where the solution is gradually accepted to a greater degree showing an incremental pattern. Indeed, when participants were encouraged to use this strategy, they had a lower proportion of correct responses and the warmth ratings showed an incremental pattern rather than an insight pattern.

Kelley and Jacoby (1996) measured metacognitive ratings in an indirect manner. They asked the participants in various experimental conditions to indicate how difficult an anagram would be to solve for other people. The results showed that familiarity with the solution word, recognition and an effect of personal experience on ratings. Higher familiarity and recognition lead to lower estimates of difficulty and solving the anagrams or seeing the solution also reduced difficulty estimates.

Finally, Topolinski, Bakhtari and Erle (2016) conducted the only large-scale study of metacognition during anagram solving. Their goal was to determine the influence of anagram length, pronounceability, and actual solvability on metacognition. They did not ask the participants to solve the anagrams, just to give estimates of whether the anagrams were solvable, the effort needed to solve them, and the estimated time needed to solve them. The results showed that shorter anagrams were rated as easier to solve when looking at the different metacognitive estimates participants gave. This was also true for anagrams easier to pronounce, showing that pronounceability is a metacognitive cue even if it actually predicts lower success rates as shown in other studies. Some of the results in this study also show that metacognitive estimates differentiated solvable from unsolvable anagrams, with higher ratings of effort and required time as well as a lower proportion rated as solvable for unsolvable anagrams.

As this short review shows, the metacognition of anagram solving has not been studied to a great extent in the current literature. Even the most comprehensive study does not actually include the process of solving anagrams, just their assessment based on short presentations. The goal of this experiment was to determine how two distinct features of anagrams (length and solvability) influence main performance indicators (proportion of correct solutions and solution time) and metacognitive ratings given before (judgments of solvability), and after the solution has been given (judgment of difficulty). We expected a strong effect of length on both performance and metacognitive indicators. Shorter anagrams were expected to be accompanied with higher judgments of solvability, a larger proportion of correct solutions, shorter solution times and lower judgments of difficulty when compared to longer anagrams. Solvability was expected to impact solution times and the proportion of correct responses (the correct response for unsolvable anagrams was recognizing it was in fact unsolvable). However, previous studies do not offer a clear prediction of actual solvability on metacognitive ratings. Topolinski et al. (2016) found a difference between solvable and unsolvable anagrams, but with a small effect size and without the participants having experience in solving the actual anagrams. If solvable and unsolvable anagrams showed a difference in judgments of solvability and difficulty, we expected the effect to be larger for shorter anagrams when compared to longer ones.

## Method

A total of 27 participants was recruited among undergraduate psychology students from the Department of Psychology at the University of Zadar.

The experiment was a 2 (anagrams/non-anagrams)  $\times$  2 (two-syllable length/three-syllable length) fully within-subject design. For the purposes of the current study, twelve items were chosen from Ostojić (2016), three items per experimental condition. All the anagram solutions were in the singular nominative, while non-anagrams were created by modifying otherwise solvable anagrams. The modification was a single letter change. While analysing responses it became apparent that one two-syllable non-anagram was solvable with a solution being in the plural genitive. This item was removed from further analysis.

The single trial procedure can be seen in Figure 1.

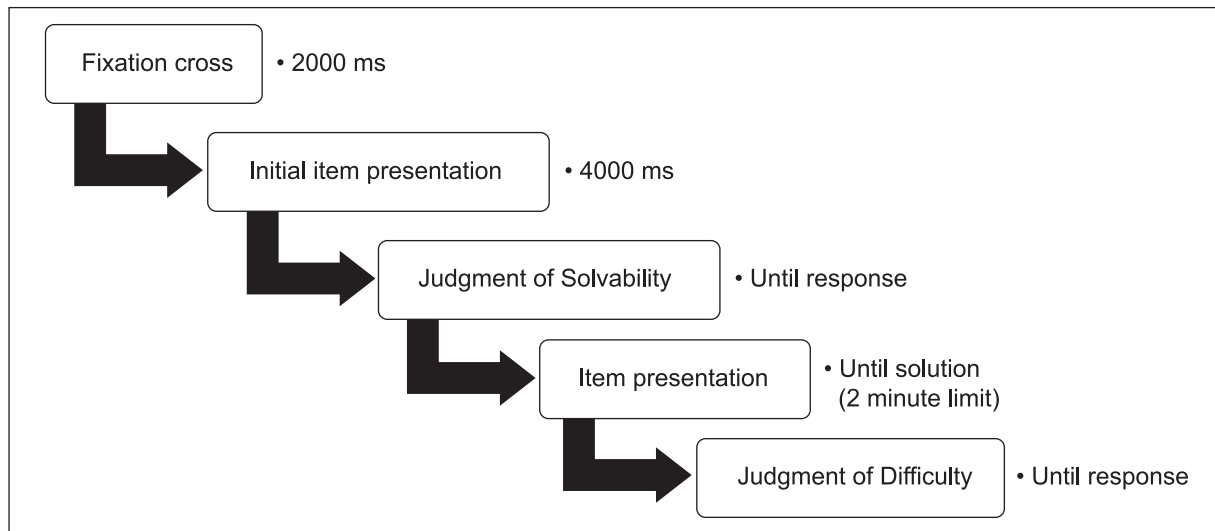


Figure 1 Single trial procedure diagram

Participants were initially shown each item for a duration of 4000 ms after which they gave judgments of solvability on a 7-point Likert scale ranging from *definitely unsolvable* to *definitely solvable*. Following the judgment of solvability, the item was presented again with a text box beneath it. Participants were given a maximum of two minutes to type their response. They were instructed to type “n” if they concluded there was no valid solution, and “o” if they thought there was a solution but that they would not be able to find it and made the decision to give up on that item. Finally, participants made a judgment of the difficulty on a 7-point Likert scale ranging from *extremely difficult* to *extremely easy*. The order of the scale points was from difficult towards easy in order to make it congruent with judgments of solvability which ranged from unsolvable to solvable. The item order was randomized for each participant, solutions, solution times and judgments were recorded for analysis.

## Results

The mean percentage of correct responses, solution times and metacognitive judgments were calculated for the final analysis. Mean judgments of solvability may be observed in Figure 2, percentage of correct solutions in Figure 3, solution times in Figure 4 and judgments of difficulty in Figure 5.

In order to determine the effect of solvability and length on the dependent variables, four  $2 \times 2$  repeated measures analyses of variance were conducted. Results of these analyses can be seen in Table 1.

Table 1 *Solvability by length ANOVA results*

	Solvability		Length		Interaction	
	$F(1, 26)$	$\eta_p^2$	$F(1, 26)$	$\eta_p^2$	$F(1, 26)$	$\eta_p^2$
Judgments of solvability	0.98	.04	10.05**	.28	0.70	.03
Percent correct	4.06†	.13	24.05**	.48	3.51‡	.12
Solution time	25.05**	.49	18.71**	.42	10.07**	.28
Judgments of difficulty	54.10**	.68	12.62**	.33	8.97**	.26

\*\* $p < .01$ ; † $p = .055$ ; ‡ $p = .072$

The analysis shows only a main effect of item length on judgments of solvability. Participants gave significantly higher initial judgments of solvability for two-syllable items when compared to three-syllable items (Figure 2).

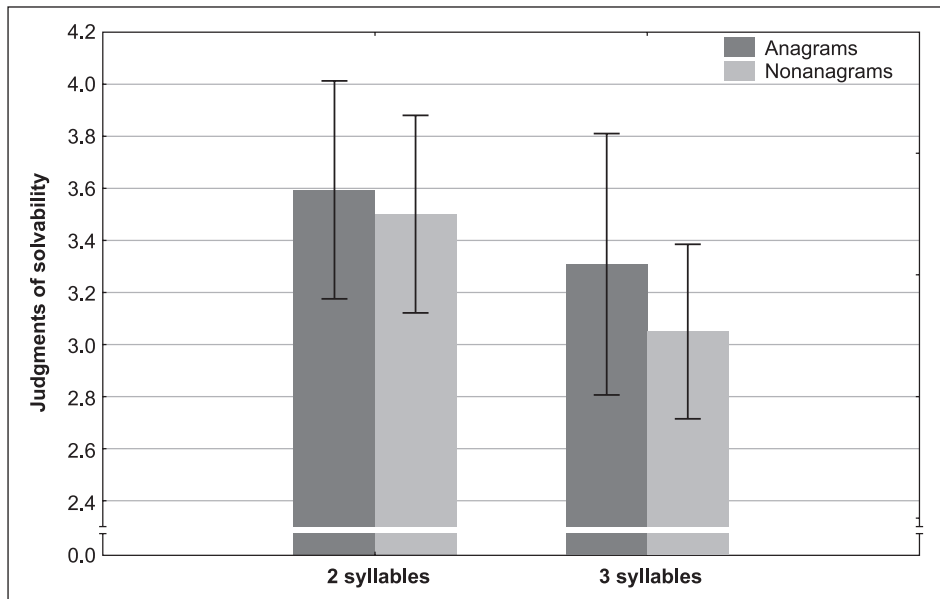


Figure 2 Judgments of solvability as a function of solvability and item length (spreads represent 95% CI)

Before reviewing the results for the percentage of correct solutions it is important to note, once more, that the correct response for non-anagrams was the recognition the item was indeed unsolvable. The main effect of length was significant for the percentage of correct solutions while the main effect of solvability and the solvability by length interaction were marginally significant ( $p = .055$  and  $p = .072$ ).

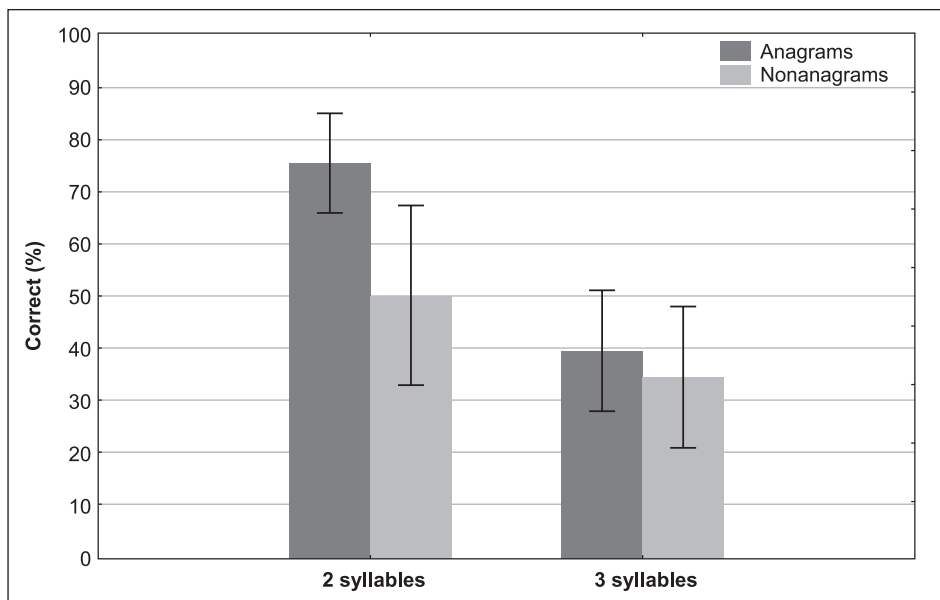


Figure 3 Percent correct as a function of solvability and item length (spreads represent 95% CI)

Post-hoc analysis (Tukey HSD) shows an interaction effect. Participants generally had a larger percentage of correct solutions for shorter items as well as for anagrams. However, the effect of solvability was only significant for the shorter items (Figure 3).

The analysis of solution times revealed that both of the main effects as well as the interaction effect were significant.

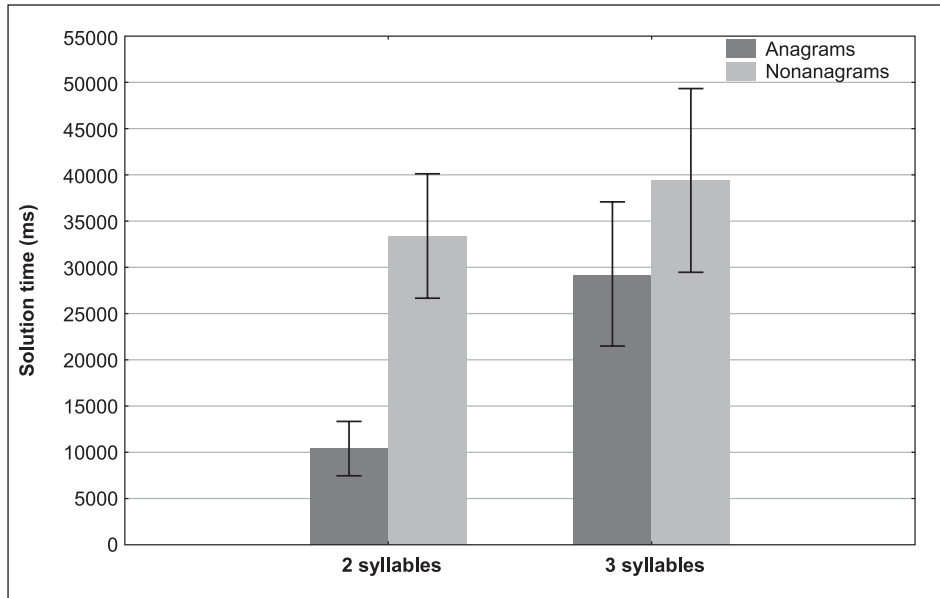


Figure 4 Solution times as a function of solvability and item length (spreads represent 95% CI)

Participants were generally faster in anagrams when compared to non-anagrams and in short when compared to longer items. Post-hoc analysis showed solution times were significantly shorter for short anagrams compared to all other conditions. Additionally, the effect of solvability was stronger for short items, and the effect of length was not significant for non-anagrams.

The final analysis of variance determined significant main and interaction effects on judgments of difficulty. Shorter items and anagrams were generally judged as easier (Figure 5). Post-hoc analysis revealed that the effect of solvability was stronger for short when compared to long items, resulting in a significant solvability by length interaction.

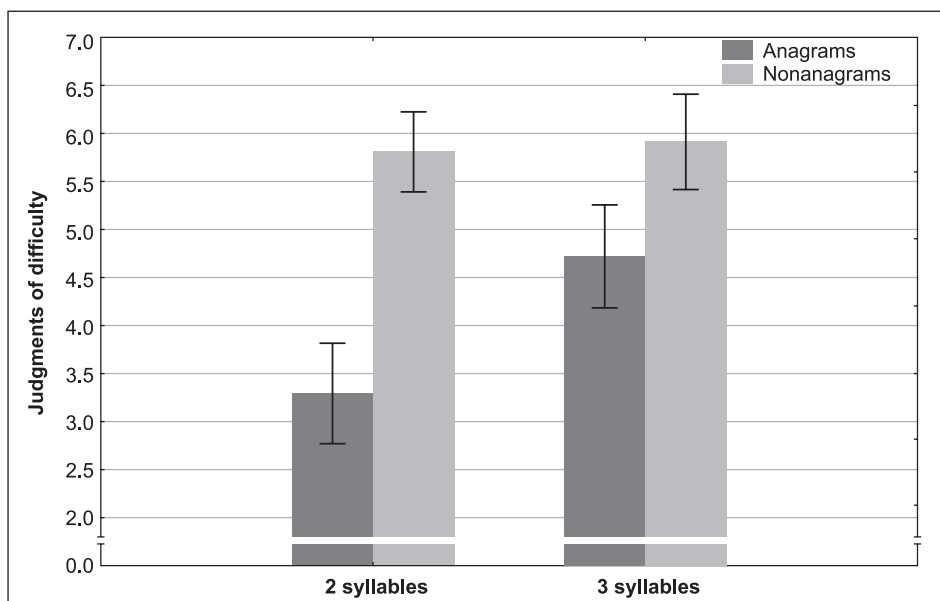


Figure 5 Judgments of difficulty as a function of solvability and item length (spreads represent 95% CI)

In order to determine correlations between metacognitive judgments and performance variables, inter-correlations, seen in Table 2, were calculated. Pearson correlation coefficients were calculated independently for each experimental condition.

The results showed that higher initial judgments of solvability predicted the percentage of correct responses only for longer anagrams, even though this may have proven the case for shorter anagrams as well with a larger sample size and/or item pool.

Lower initial judgments of solvability were a strong predictor of higher judgments of difficulty in every condition except the short non-anagrams. However, the correlation was established in the expected direction and may have reached significance with a larger sample size and/or item pool.

As could be expected, the percentage of correct solutions was predictive of judgments of difficulty for anagrams but not for non-anagrams. The participants with a higher number of correct responses gave lower judgments of difficulty.

The most interesting result was that solution time was not correlated with either of the metacognitive judgments. Faster participants usually provide higher metacognitive judgments of confidence, so it was expected that judgments of solvability would predict time on task and that the time would predict judgments of difficulty.

Table 2 *Pearson correlation coefficients between dependent variables*

2 syllable anagrams	2.	3.	4.
1. JOS	.21	-.24	-.40*
2. Percent correct	-	-.54**	-.73**
3. Solution time		-	.37
4. JOD			-
3 syllable anagrams	2.	3.	4.
1. JOS	.62**	.01	-.83**
2. Percent correct	-	-.16	-.81**
3. Solution time		-	.13
4. JOD			-
2 syllable non-anagrams	2.	3.	4.
1. JOS	.09	.34	-.30
2. Percent correct	-	-.04	.30
3. Solution time		-	.04
4. JOD			-
3 syllable non-anagrams	2.	3.	4.
1. JOS	.06	.32	-.59**
2. Percent correct	-	-.31	.01
3. Solution time		-	-.23
4. JOD			-

\* $p < .05$ ; \*\* $p < .01$ ;  $df = 25$

JOS = judgment of solvability

JOD = judgment of difficulty

Participant-level correlational analysis is more of an indication of individual differences, so it is common in meta-reasoning research to conduct an item-level analysis. The analysis was conducted and is reported in Table 3 though it should be interpreted with caution since the total item pool consists of only eleven items.

Table 3 *Item-level correlations*

	2.	3.	4.
1. JOS	.78**	-.75**	-.81**
2. Percent correct	-	-.89**	-.79**
3. Solution time		-	.94**
4. JOD			-

\*\* $p < .01$

JOS = judgment of solvability

JOD = judgment of difficulty

As is usual in meta-reasoning research all of the correlations is highly significant, with the unusually high coefficients probably a result of the item pool size. Items for which participants gave higher judgments of solvability were solved correctly in a higher percentage of trials, were accompanied with shorter solution times and lower judgments of difficulty. Items solved correctly in a higher percentage of trials were accompanied by shorter solution times and lower judgments of difficulty. Finally, items with longer solution times were judged as more difficult.

## Discussion and conclusion

The goal of this experiment was to determine the influence of anagram solvability and length on performance and metacognitive judgments. As expected, we found a strong overall effect of length on both performance and metacognitive indicators. This is in line with the review by Muncer and Knight (2011) which shows a sharp decline in performance from two-syllable to three-syllable anagrams. For judgments of solvability this effect proved to be simple, regardless of actual solvability. This result means that short exposition was not enough to distinguish between anagrams and nonanagrams, although while not reaching significance, the tendency to judge anagrams as more solvable was evident, especially for longer anagrams. There is a need to expand on these results by increasing item count and sample size in future studies. This pilot study thus does not replicate the findings of Topolinski et al. (2016) for initial metacognitive judgments. It is interesting that while length remained a strong effect for solution correctness, the difference between anagrams and nonanagrams was significant only for longer items. This shows how large an impact length has on the anagram solving process. Participants correctly identified nonanagrams at the same rate as they solved anagrams for the three-syllable items. The decreased effect of solvability was evident on solution times where the difference between anagrams and nonanagrams was significantly stronger for two-syllable when compared to three-syllable anagrams. The most interesting result comes from judgments of difficulty where the effect of solvability was weaker for longer items but was still highly significant. Even though performance was very similar for anagrams and nonanagrams in terms of success rate, the nonanagrams were judged as more difficult. This however, may be due to the difference in solution times rather than a direct distinction between solvable and unsolvable items.

Correlational analysis revealed that judgments of solvability represent the strongest predictor of

difficulty judgments on a participant level. An interesting result is a failure to determine correlations of solution times with either metacognitive judgment. Faster participants did not judge the items as less difficult and those who gave higher judgments of solvability did not spend more time solving the anagrams. There is a trend for nonanagrams which shows that higher judgments of solvability may be correlated with the time on task (Table 2). On the other hand, success rate was expectedly negatively correlated for anagrams while the same was not found for nonanagrams. This is expected due to the nature of problem solving tasks which often have easily confirmable solutions. Anagrams had a solution which lead to lower judgments of difficulty. Nonanagrams, however, retain a high level of uncertainty even when correctly identified as unsolvable thus the pattern of correlations between success rate and metacognitive judgments of difficulty mirrors those found in meta-reasoning research. The item level correlations provide only limited information due to the low number of items used for this experiment but reveal expectedly strong correlations between the dependent variables. Items which were accompanied by higher judgments of solvability were solved at a higher rate, faster and had lower judgments of difficulty. Since the intercorrelations are quite high two suggestions for further research include the increase of the item pool and a regression analysis to better understand mediation effects for item-level relationships.

When considered within the dual-process approach and meta-reasoning framework we can emphasize a few main findings. First, the participants choose the most salient feature of the items to provide initial judgments of solvability. Length seems to be a key and salient feature for generating these early metacognitive representations of the anagram solving task. Further research should follow-up on the anagram solving literature which has already identified many features which influence performance as other possible metacognitive cues for early representations (e.g. pronounceability, syllabic structure, solution word frequency, etc.). These early metacognitive representations seem to play a dominant role in the formation of post-processing metacognitive judgments. Indeed, early metacognitive representations and cues influencing them may have further implications on intermediate metacognitive judgments as well as the choice of problem-solving strategies and other processes. Thus, further research of initial metacognitive representations, which has not been a focus in problem-solving research so far, should feature more heavily in the future. Second, it remains unclear if participants are able to directly distinguish between solvable and unsolvable problems, or if the distinction is mediated by the differences in time on task. From this perspective, an increased item and participant pool would enable in-depth analysis base do on item and response type (for an example from meta-reasoning research see Dujmović & Valerjev, 2018). Just as for reasoning research problem-solving tasks need to be analysed in a more complex manner in order to uncover the relationships among underlying processes. Overall, the results show that manipulating with the heuristics on which participants base their decision has an impact on problem-solving performance and metacognition.

The main limitations of the present study are a limited sample size and number of trials per experimental condition and therefore limited statistical power. We would expect marginal effects to be fleshed out in a more comprehensive study. Future improvements might include a larger sample, additional items, four-syllable anagrams, longer intervals for solvability judgments and manipulating the amount and type of anagram training. As the results of this study suggest, the meta-reasoning processes that are involved in problem solving have its own specific characteristics when compared to other meta-reasoning findings obtained with classical reasoning tasks. For a more complete picture of meta-reasoning processes it is important to continue research on different types of thinking. Multistep problem solving research may provide more insights into similar processes which unite reasoning, problem solving and decision making.

To summarize, both anagram length and anagram solvability proved significant effects on response times and judgments of difficulty. Anagram length also demonstrated a significant effect on accuracy and judgment of solvability. We can conclude that this study shows that meta-reasoning research paradigms may be applied to problem-solving tasks to expand the rich tradition and provide further insight into prob-

lem-solving. The study also highlights the necessity for large item pools and sample sizes in order to go beyond the superficial relationships. This deeper step is required to truly investigate metacognition in problem-solving. Basic findings of mechanisms influencing both metacognitive and problem-solving processes will surely have important implications in numerous settings from metacognitive self-regulation in education to on-line feedback in the performance of many different tasks.

## References

- Ackerman, R. (2014). The diminishing criterion model for metacognitive regulation of time investment. *Journal of Experimental Psychology: General*, *143*(3), 1349-1368.
- Ackerman, R., & Thompson, V.A. (2015). Meta-reasoning: What can we learn from meta-memory? In A. Feeney & V.A. Thompson (Eds.), *Reasoning as Memory* (pp. 164-182), New York: Psychology Press.
- Ackerman, R., & Thompson, V. A. (2017). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences*, *21*(8), 607-617. <https://doi.org/10.1016/j.tics.2017.05.004>
- Ackerman, R., & Zalmanov, H. (2012). The persistence of the fluency-confidence association in problem solving. *Psychonomic Bulletin & Review*, *19*, 1187-1192. <https://doi.org/10.3758/s13423-012-0305-z>
- Bajšanski, I., Močibob, M., & Valerjev, P. (2014). Metacognitive judgments and syllogistic reasoning. *Psychological Topics*, *23*(1), 143-165.
- Bernstein, D. M., Rudd, M. E., Erdfelder, E., Godfrey, R., & Loftus, E. F. (2009). The revelation effect for autobiographical memory: A mixture-model analysis. *Psychonomic Bulletin & Review*, *16*, 463-468. DOI: 10.3758/PBR.16.3.463
- De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1248-1299. <https://doi.org/10.1016/j.cognition.2007.06.002>
- Deloche, G., Ott, M., & Tavella, M. (1995). Anagram solving and reading abilities in children. *Reading and Writing: An Interdisciplinary Journal*, *7*, 395-406. <https://doi.org/10.1007/BF01027726>
- Dujmović, M., & Valerjev, P. (2018). The influence of conflict monitoring on meta-reasoning and response times in a base rate task. *Quarterly Journal of Experimental Psychology*. Online first. <https://doi.org/10.1177%2F1747021817746924>
- Evans, J.St.B.T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255-278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Evans, J.St.B.T., & Stanovich, K.E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223-241. <https://doi.org/10.1177/1745691612460685>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906-911. <http://dx.doi.org/10.1037/0003-066X.34.10.906>
- 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>
- Gilhooly, K., & Johnson, C. E. (1978). Effects of solution word attributes on anagram difficulty: A regression analysis. *Quarterly Journal of Experimental Psychology*, *30*, 57-70. <https://doi.org/10.1080/14640747808400654>
- Kaplan, I.T., & Carvellas, T. (1968). Effect of word length on anagram solution time. *Journal of Verbal Learning and Verbal Behavior*, *7*, 201-206. [http://dx.doi.org/10.1016/S0022-5371\(68\)80189-6](http://dx.doi.org/10.1016/S0022-5371(68)80189-6)
- Kelley, C.M., & Jacoby, L.L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language*, *35*, 157-175. <http://dx.doi.org/10.1006/jmla.1996.0009>

- Mayzner, M. S., & Tresselt, M. E. (1958). Anagram solution times: A function of letter order and word frequency. *Journal of Experimental Psychology*, *56*(4), 376–379. <http://dx.doi.org/10.1037/h0041542>
- Metcalfe, J. (1986a). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(2), 288-294. <http://dx.doi.org/10.1037/0278-7393.12.2.288>
- Metcalfe, J. (1986b). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 623-634. <http://dx.doi.org/10.1037/0278-7393.12.4.623>
- Metcalfe, J. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, *15*(3), 238-246. doi: 10.3758/BF03197722
- 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. <https://doi.org/10.1080/20445911.2014.986487>
- Muncer, S.J., & Knight, D. (2011). The syllable effect in anagram solution: Unrecognised evidence from past studies. *Journal of Psycholinguistic Research*, *40*, 111-118. <https://doi.org/10.1007/s10936-010-9159-6>
- Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower. (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 26, pp. 125-173). San Diego: Academic Press.
- Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Novick, L. R., & Sherman, S. J. (2003). On the nature of insight solutions: Evidence from skill differences in anagram solution. *Quarterly Journal of Experimental Psychology*, *56A*(2), 351-382. <https://doi.org/10.1080/02724980244000288>
- Novick, L. R., & Sherman, S. J. (2008). The effects of superficial and structural information on online problem solving for good versus poor anagram solvers. *The Quarterly Journal of Experimental Psychology*, *61*(7), 1098–1120. <https://doi.org/10.1080/17470210701449936>
- Ostojić, A. (2016). *Metakognitivne procjene pri rješavanju anagramskih problema [Metacognitive judgments during anagram problem solving]*. Unpublished graduate thesis. Mostar: Filozofski fakultet.
- 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. <https://doi.org/10.1016/j.cogpsych.2015.05.001>
- Sarris, M.E., & Panagiotakopoulos, C.T. (2013). Linguistic effects on anagram solution: The case of a transparent language. *World Journal of Education*, *3*(4), 41-51. DOI: <https://doi.org/10.5430/wje.v3n4p41>
- Shynkaruk, J.M., & Thompson, V.A. (2006). Confidence and accuracy in deductive reasoning. *Memory & Cognition*, *34*(3), 619-632. <https://doi.org/10.3758/BF03193584>
- Thompson, V.A., Evans, J.St.B.T., & Campbell, J.I.D. (2013). Matching bias on the selection task: It's fast and feels good. *Thinking & Reasoning*, *19*(3), 431-452. <https://doi.org/10.1080/13546783.2013.820220>
- 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., & Morsanyi, K. (2012). Analytic thinking: Do you feel like it? *Mind & Society*, *11*, 93-105. <https://doi.org/10.1007/s11299-012-0100-6>
- Thompson, V.A., Prowse Turner, J.A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*, 107-140. <https://doi.org/10.1016/j.cogpsych.2011.06.001>
- Thompson, V.A., Prowse Turner, J.A., Pennycook, G., Ball, L.J., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, *128*, 237-251. <https://doi.org/10.1016/j.cognition.2012.09.012>

- Toplak, M.E., West, F., Stanovich, K.E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, 20(2), 147– 168. <https://doi.org/10.1080/13546783.2013.844729>
- Topolinski, S., Bakhtiari, G., & Erle, T.M. (2016). Can I cut the Gordian knot? The impact of pronounceability, actual solvability, and length on intuitive problem solving of anagrams. *Cognition*, 146, 439-452. <https://doi.org/10.1016/j.cognition.2015.10.019>
- Valerjev, P., & Dujmović, M. (2017). Instruction type and believability influence on metareasoning in a base rate task. In: G. Gunzelmann, A. Howes, T. Tenbrink, E. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 3429-3434.). Austin, TX: Cognitive Science Society
- Witte, K. L., & Freund, J. S. (2001). Single-letter retrieval cues for anagram solution. *Journal of General Psychology*, 128(3), 315-328. <https://doi.org/10.1080/00221300109598914>