THE LEVEL OF METACOGNITIVE MONITORING WITH LEARNERS AT HIGHER ELEMENTARY GRADE IN PROCESSING INFORMATIONAL TEXTS IN THE SUBJECT OF SCIENCE DEPENDING ON THE TYPE OF EDUCATIONAL CURRICULUM

Jaroslav Říčan[1], Vlastimil Chytrý[1]

[1]*Univerzity of Jan Evangelista Purkyně in Ústí nad Labem, Czech Republic*

Abstract

The paper consists of finding the level of metacognitive monitoring in processing informational texts in Science with 5th grade learners attending (a) general elementary schools, (b) Dalton education elementary schools (c) RWCT programme elementary schools. The goal of our empirical research was to verify a presumption that educational programme and the strategies of controlling the curriculum derived from it do influence metacognitive monitoring development. The methodology of the research has a quasi-experimental design. The results confirm the presumption that education programme types influence the learner´s metacognitive development. The RWCT pupils together with Dalton learners are significantly more accurate in their judgement of (in)correctness of answers linked to processing informational texts and scored vastly better in informational text processing tests compared to learners at general schools. The conclusion has drafted potential reasoning for the differences analyzed as a result of the executed curriculum (teaching effect) and offers recommendations for theoretical and professional practice

INTRODUCTION

Systematical empirical research in education has shown that children differ in the ways of using their memory. Some file information in their memory and recall it from there with no care and effort, "passively". Conversely, some exercise active and purposed usage of strategies while handling data saved in their memory (Schneider & Bjorklund, 1998; Schneider & Pressley, 1997). The continuum of such behaviour differs vastly with individuals depending on the situation as well as its context. Regardless of the fact, the frequency along with the quality of strategies used monitored by teaching and teaching tools organized by the teacher or other agents entrusted with education (Bjorklund & Coyle, 1995; Waters, 2000). It proves justified to assume that the quality of the pedagogues, education institutions as well as entire education curricula strongly determines the learner´s ability to think strategically. The study of metacognition has lent many an educationalist the key to a new dimension of a learner´s potential, which is to be carefully analyzed and researched. So far, we have found only very little on how individuals move from rather passive unintended use of strategies to the purposed one, which we more often notice with older and more experienced children and especially adults (Waters & Schneider, 2009). The use of metacognition has proved to be a promising direction to unveiling suspected and unsuspected options in education. As contemporary empirical finds show, the level of metacognitive development explains a large number of varieties in test scores (Dunlosky & Metcalfe, 2009; Schneider & Artelt, 2010; Schneider, Schlagmüller, & Visé, 1998) and its predictive potential in its relation to successful school performance is higher than the level of intelligence mapped through standard finding methods (Veenman & Spaans, 2005; Wang, Haertel, & Walberg, 1990). Herewith we express our presumption which at

the same time is the goal of this study, that a specific type of educational programme may be the cause of the subtleties at the level of metacognitive monitoring in learners. Since according to Winne and Hadwin (1998) metacognitive aspect of monitoring occupies the very core of metacognition and the models of self-regulated learning, our study focuses on this regulation component of metacognition.

Metacognitive monitoring as a theoretical and empirical construct

Metacognitive monitoring is a partial element in a process component to metacognition (metacognitive regulation) and is largely defined as an individual´s ability to evaluate their own cognitive performance; it is a specific way of self-evaluation (for reasons of finding metacognitive regulation) to be grasped at various moments of the learning-teaching process (before, during and after) and they are linked to a variety of cognitive processes (Leonesio & Nelson, 1990). From a theoretical perspective metacognitive monitoring is to be filed in the regulation component of metacognition (together with planning and evaluation) and we may state that the importance of the regulation component is in some ways larger than the knowledge-based component of metacognition (metacognitive knowledge composed of declarative, procedural and conditional knowledge), as to master the strategic knowledge does not equal its proper practical use. As we read in Winne (1995) and Zimmerman (1990), metacognitive monitoring is the most complicated as well as the longest-developing metacognitive competence, which not necessarily has to reach an adequate level even when adults.

The current trend in world scientific literature stresses metacognitive strategies which vary in the sense of in what domain we apply them (reading comprehension, Mathematics, History, Biology and others). Generally expressed, there are two phases to be recognised where metacognition contributes to a learner´s success at a set task. In its first phase it is the successful grasp of a set task and taking all information in, in the second phase it is the finding of a correct solution to the task and its application. The first phase refers to metacognitive strategies leading to understand the task (comprehension strategies), to be followed by metacognitive strategies leading to finding a correct solution (problem solving strategies). The learner primarily needs to gain vital information and to understand it as well as point at what the problem is and what is expected of them. In the second phase potential solutions and their outcomes are handled. Both aspects are present at any type of task with no regard to the domain they cover, and both are key elements in reaching a successful solution. Nevertheless, the nature of a task influences which of the two aspects is stressed to reach a plausible solution. The directions in research may be differentiated in two major directions: the first direction deals with the issue of using metacognitive strategies in terms of problem-solving, where we place mathematics tasks and other branches constructed upon exact research (Biology, Physics, Chemistry and others - Listiana, Susilo, Suwono, & Suarsini, 2016; Sadi & Uyar, 2013; Sen, 2016); the second direction analyses metacognitive strategies in the context of literacy and especially reading comprehension and writing Listiana, Susilo, Suwono, & Suarsini, 2016; Sadi & Uyar, 2013; Sen, 2016). In the mutual consonance of both directions, researches in the field of intended intervention clearly show that systematic exercise develops a learner´s metacognitive potential (the intervention leads to significantly positive changes in the participating subjects - Kurtz & Borkowski, 1984; Pressley, Harris, & Marks, 1992; Schleifer, 2009; Schneider & Pressley, 1997; Schraw, 1998). In the context of the research intention of this study we have pointed out the latter direction analysing the level of reading comprehension.

In terms of text processing we focus on e.g. attention (Serafini, 2013a; 2013b), motivation (Gambrell, 2011; Roeschl-Heils, Schneider, Kraayenoord, 2003), the ability to solve problems (Braselton & Decker, 1994) or a scope of cognitive effort exerted in processing (Maki, Foley, Kojer, Thompson, & Willert, 1990). The study in the field of research reading with comprehension and metacognition bring rather consistent conclusions: metacognition plays a significant part in the quality of comprehenstion (Bransford, Brown, & Cocking, 2000). Efficient readers characterized with high level of metacognitive development - while processing reading with comprehension - construct especially situational text representation (mental representation, fully dependent of the text contents) and propositional representation (mental representation formatting the intention and major ideas of the text). They permanently interact with the text (extrapolate major ideas, notice discrepancies, express the level of trust in their understanding, connect just-read to the information gained before etc.) and take notice of their reading as a process of a sense-quest. Such process is complemented with metacognitive thinking (handling the process of reading as an object designed to be planned at a certain meta-level, to be monitored, regulated and reflected). Conversely, less able readers will form verbatim representation based on a lexical meaning and on-the-surface "literal" structures of the text (Kintsch, 1988) and the process of reading is understood as a mere act of its deciphering (with an absence of higher cognitive processes of cogitation). We see that to understand textual material deeply, (a) background knowledge and (b)

learner´s metacognitive level are of a vital importance. Not only children, but also adults – who are to be labelled efficient readers – find it demanding to understand certain types of texts with technical and professional terms. Deep understanding of such texts is a challenge as readers commonly offer a low level of professional vocabulary relevant or other forms of previous experience, which disables them to understand the text fully (Best, Rowe, Ozuru, & McNamara, 2005). The lack of experience linked to a particular task may well be compensated through implementing metacognitive processes (Prins, Veenman, & Elshout, 2006; Schraw, 1998).

The construct of metacognitive monitoring and its place in the pedagogical psychology research

Independent of the moment (before, during, after) metacognitive monitoring provides the learner with a feedback of their own progress in learning and enables them to adapt learning processes. In this respect, the core importance of metacognitive monitoring lies in (1) unravelling the hardship throughout the teaching-learning process (and implementing a learning strategy could adequately react) and (2) adjusting the time to be used for learning to the learning requirements or ceasing the learning process in time if it has been successfully completed or remains with no reachable solution (Dunlosky & Metcalfe, 2009). Metacognitive monitoring is commonly defined as a rate of concordance between self-evaluation of one´s own learning, reading and memory performance and the real performance. Maki, Shield, Wheeler a Zacchilli (2005) poignantly formulate, that "… the ability of people to evaluate their own performance is one of the aspects in metacognition, mainly metacognitive monitoring." (p. 723). The acts of evaluating are known as metacognitive judgments (Dunn, 2004; Nelson 1996), and as such they are to be considered indicators of metacognitive monitoring levels. On a more general level, they may be a sign of evaluating the difficulty of a task (if it is to be simple or complicated – ease of learning judgments), the evaluation of results gained through learning (judgements of learning referring to the prediction: e.g. evaluating "the number of points I score in a test", "which passages I master and which I don´t", the evaluation of if the problem has been successfully solved (confidence in the retrieved answer/confidence judgments, referring to the postdiction and in its specific sense of the word if they stand on the other side in relation to judgments of learning), or the evaluation if something that cannot be recalled just now belongs to the knowledge base and therefore may be recalled or not (feeling of knowing judgements). The evaluation of one´s own certainty in giving a correct answer, so called confidence judgments, refer to evaluating one´s own performance (postdiction) and serves the role of an inner feedback that can incite a repeated learning activity along with the act of reading. Confidence judgements may refer to either a general judgement ("How many points do I think I could have scored in the test?" or the local judgement ("How self-confident am I in judging if my answer to task has been correct?")

The research has shown (Nietfeld, Cao, & Osborne, 2005) that the local judgement rates higher relation to the performance measured than the general judgement; given that, our paper focuses at researching the meanings of predictive and postdictive judgements in textual processing to reach a proximate impact on the educational practice and pointed out especially predictive judgements as they instantly refer to the very learning efficiency, meaning with the estimation of the quality of acquiring particular area in the learning material and behaviour (devotion of time or the lack of it) corresponding with it. In terms of prediction, as has been verified, even gifted learners and adults tend to overestimate their performance and postdiction is a considerably more precise act compared to prediction (Pressley & Ghatala, 1990; Hacker, Bol, Horgan, & Rakow, 2008; Narens, Nelson, & Scheck 2008), which led us to finding data (and do so more accurately) postdictive judgements instead of the predictive ones. The data gathered on the base of confidence judgements may be used for pinning down such indicators as bias, dicrimination index, scatter index and relative accuracy index. The most frequently used indicator to determine the level of metacognitive monitoring is index of absolute accuracy (Burson, Larrick, & Klayman, 2006; Nelson, 1996). Absolute accuracy refers to accuracy within the confidence judgements in comparison to the performance at a particular task and may therefore be interpreted as a discrepancy between the confidence judgement and the performance in a particular task (Maki, Shield, Wheeler, & Zacchilli 2005). For example, if an individual´s confidence reaches 90 % in judging the correctness of their answer, we may assume that the individual is highly accurate. Oppositely, if the individual has answered correctly and their estimation only reaches 10 % in believing so, the individual´s judgement is largely inaccurate. The result of this is known as absolute accuracy index (also referred to as calibration index) and refers to the precision of judgements reaching the value of 0 (absolute accuracy) to 1 (absolute inaccuracy):

Absolute Accuracy Index $= \frac{1}{N}\sum_{i=1}^{N}(c_i - p_i)^2$

, where there c_i is a confidence judgement value rate (10 %, 20 %, … as 0,1; 0, 2) and p_i refers to the performance in the tested item (in dichotomy; 0 = inaccurate response, the value of 0; 100 % = correct response, the value of 1).

METHODOLOGY OF RESEARCH

Research design and Sample

The research design has been inspired by the work of a research team at an American university in Idaho (Thiede, Redford, Wiley, & Griffin, 2012). The researcher followed the differences in reading comprehension in its context to metacognitive monitoring in learners of the 7th and 8th grades and compared them to their peers at the same school grades. At the school where the research was carried out, the curriculum explicitly stressed the importance of reading comprehension (quest for textual sense and judgement-making process) and metacognitive strategies were implemented (in relation to the regular evaluation of understanding through summarizing, discussion and reasoning in dialogues; the creation and reasoning for mental maps; activities for predication). The school enlarged classes and created parallel classes at the 7th and 8th grades earlier than 2012, meaning that a half of the learners had not faced the specific curriculum of the school given. These new-comers arose from rather conventional schools with no trace of such activities. The team was trying to find if facing a particular curriculum would influence their reading comprehension and metacognitive monitoring. The data proved remarkable differences in textual processing as well as metacognitive monitoring to the advantage of the learners facing the specific curriculum of the particular school (in comparison to the newly-coming learners who had had no experience with a specific curriculum). We have aimed our research intent in a similar style. We wanted to know if there are differences to be noticed in metacognitive development of learners if they are under long-term influence of different curricula. According to the RVP ZV[1] in the Czech Republic the autonomy of schools strengthened and offered an individual focus of the school, we may calculate many a school as institutions with specific interests (elementary language schools, IT-oriented schools, …), and an elementary school may be differentiated according to a specific point of view. For research reasons, the authors of the paper have stated the three categories to follow:

- Non-specific elementary schools.[2]

- Elementary schools implementing features of RWCT[3] in their curriculum.

- Elementary schools implementing features of Dalton planeducation in their curriculum.[4]

The research sample was designed on the base of a multi-stage random choice of learners attending a 5th grade in the school years of 2015/2016[5].

Instrument: A modified tool for finding the level of reading comprehension of scientific contents and metacognitive monitoring

For the purposes of our research we have used a relaxed problem in Progress in International Reading Literacy Study (PIRLS) – „The Mystery of the Giant Fang" including professional terminology from the fields of history and archeology (Janotová & Šafránková, 2013). The problem deals with the changes in ways of people´s

---

[1] The fundamental and unifying curricula document which is the result of reformatory efforst in Czech education (the first drafts of verification were started in 2004 and were followed by issuing the document on September 1 2007, with its binding principle to be accepted by all 1st and 6th graders). Even though the school autonomy was strengthened and it was highly recommended to implement teaching methods other than transmissive and the potential cooperation with others was vastly accentuated, these ideas have remained to be mere proclamations rather than their potential having been fulfilled in real school life.

[2] By conventional schools we mean non-specialized institutions with no reflection of a particular educational curriculum (e.g. Step by Step, Montessori, Waldorf, RWCT, Dalton plan etc.). These may from here on be referred to a "N-S ES).

[3] Reading and Writing for Critical Thinking.

[4] From here onwards „ES Dalton plan" or „ES Dalton". The Czech Republic does not provide a wholly fulfilling approach to Dalton plan so that it would entirely copy the original teaching concept of its founder Helen Parkhurst, but it is combined with more conventional organization of subjects with classes based on self-dependence, responsibility and a free choice of the activity chronology at a ratio of 3:1 to the advantage of a conventional subject organization.

[5] The data has been excerpted from a list of schools of the Ministry of Education, Youth and Sports: http://rejskol.msmt.cz/ (publicized January 15, 2016)

thinking over historical finds and therefore it is ideal for the subject of Science. From a perspective of the character of the problem used in this paper we have chosen a text of which the idea was to read it and gain information and put it to practice (corresponding with reading informative texts; compared to reading literary texts to gain literary experience in the context of pleasures and interests). The lack of knowledge of a particular term (in our case e.g.: Iguanodon, Jura, Cetraceous Period etc.) may be an obstacle for an otherwise efficient reader. The obstacle may well be overcome by introducing metacognitive strategies (self-questioning, word meaning deduction from its context and its combination with various passages, etc.). The problem illustrated above has been modified in parts for the purposes of the study. Three questions have been excluded (1x to deduce conclusions and 2x to fill in forms); we excluded them because of the fact that a problem with binary scores we are unable to find dimensions of metacognitive monitoring. In general, the modified tool for finding the level of reading comprehension of scientific contents have comprised of 11 tasks out of which 4 are of cognitively less challenging character and 7 of cognitively more challenging character. To gather data leading to marking the dimension of metacognitive monitoring, a 10cm-long segment (11 point scale) has been added so that the learner had a chance to express their level of confidence judgement to their correct response (0 % - 100 %).

RESULTS OF RESEARCH

The statistics with both observed dimensions (the level of reading comprehension and metacognitive monitoring) comprises of logical order of steps: (1) descriptive statistics shows level of data positions, in its inductive part as follows (2) normality tests through adequate statistical method, (3) visualisation of data through quartile graphs, (4) finding differences between observed types of educational programmes through adequate statistical method and (5) determining particular fields where significant differences occurred.

Descriptive statistics

Out of the entire number of respondents (N=1103) 97 % filled in the information on their gender: (N=1075). Out of the count 52. 2 % were boys and (N=561) 47. 8 % were girls (N=514).

Tab. 1: Allocation of boys and girls according to the type of their educational programme.

	N-S ES	ES Dalton	ES RWCT	Total
Boys	360 (53.1 %)	143 (52.4 %)	58 (46.8 %)	561 (52.2 %)
Girls	318 (46.9 %)	130 (47.6 %)	66 (53.2 %)	514 (47.8 %)
Total	678	273	124	1075

The actual description has been divided according to the field researched (reading comprehension and metacognitive monitoring). The reliability of the reading comprehension of scientific contents level test equalled (Kuder-Richardson coefficient - 206).

Tab. 2: Basic descriptive data for reading comprehension of scientific contents level test and metacognitive monitoring.

	Understanding text			Metacognitive monitoring		
	N-S ES	ES Dalton	ES RWCT	N-S ES	ES Dalton	ES RWCT
N	699	279	125	699	279	125
Ø	7.75	8.43	8.78	0.21	0.19	0.17
Me	8.00	9.00	9.00	0.18	0.17	0.16
Mod	10.00	10.00	10.00	0.18	0.09	0.00
SD	2.51	2.18	2.18	0.13	0.12	0.11

---

[6] The KR20 formula has been used as it can be interpreted as a specific type of Cronbach´s alpha and after it has been used we may gather identical results in binary data.

Inductive statistics to reading comprehension results

According to (Levine & Hullett, 2002; Tolson, 1980), the starting point for the choice of the significance level are conclusions based on applied statistics saying that while verifying a new pedagogical method as well as others a higher risk level may be opted for (e.g. α = .10). The character of statistics testing shows the significance level (p-level value) to be strongly dependent on the size of the set.

With regard to the scope of the set and the impact of the research we shall consider the significance level to be 0.01.

Since none of the samples analysed provided the criteria of data normality, it would be more than difficult to verify practical significance next to statistical significance (Soukup, 2016). Testing data normality was through Shapiro-Wilk normality test (Shapiro & Wilk, 1965), where null hypothesis was tested against to prove that the data evaluated has normal distribution. The p-level values were as follows:

Tab. 3: P-level values found with the results in reading comprehension at individual types of educational programmes.

Field analysed	p-level value
Standard ES	**p <.0001**
ES Dalton	**p <.0001**
ES RWCT	**p <.0001**

P-level value was p < .01 in all cases researched and that is why it is possible to reject null hypothesis at normal data distribution and accept alternative hypothesis that the data has different than normal distribution.

Since there was need to compare more than two sets and we cannot assume a normal data distribution, Kruskal-Wallis test was used to prove differences in individual groups (Kruskal & Wallis, 1952), and was followed by a post hoc analysis (multiple observations). Dunn method (Dunn, 1964) was used with post hoc analysis since we do not operate with balanced classification.

The following alternative hypothesis has been formulated:

H1   High scores of the test finding the level of reading comprehension differ in dependence of the type of education programme.

For the alternative hypothesis, we stated null hypothesis, which was tested on a one- percent significance level:

H1$_0$ High scores of the test finding the level of reading comprehension are not dependent on the type of the education programme (medians are equal).

The p - level value for Kruskal-Wallis test is p < .0001. We therefore reject null hypothesis on the equality of medians and confirm the alternative hypothesis on the differences in learners´ performance in a test finding the level of reading comprehension in dependence of the type of education programme. The following post hoc analysis determines where there is statistically significant difference between observed groups.

Tab. 4: Post hoc analysis values found in the area of reading comprehension. Numbers in bold refer to education programmes with statistically significant differences in the test finding the level of reading comprehension.

r/s	|tr-ts| (critical value)		
	N-S ES	ES Dalton	ES RWCT
N-S ES	---------	**62.418 (30.353)**	**147.074 (37.164)**
ES Dalton	**62.418 (30.353)**	---------	**84.656 (43.907)**
ES RWCT	**147-074 (37.164)**	**84.656 (43.907)**	---------

Statistically significant differences on a one percent significance level showed in all areas researched. The situation in detail is to be seen in the quartile graph.
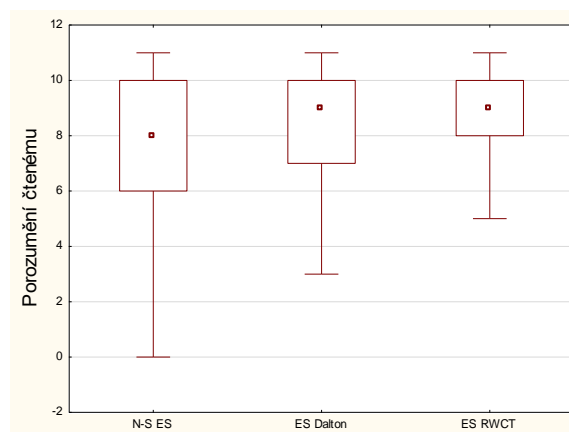
Fig. 1: Quartile graph to the results in reading comprehension with individual types of educational programmes.

Inductive statistics to the results of metacognitive monitoring

The normality of data was tested through Shapiro-Wilk normality test (Shapiro & Wilk, 1965) where we test the normal distribution of data with the data analysed against the null hypothesis. With regard to p-level values shown in the table below (Tab. 5.) we are to reject the null hypothesis of a normal data distribution on a one percent significance level.

Table 5. P-level values found with results of metacognitive monitoring level in the specific domain of reading at individual types of educational programmes.

Formula	Absolute accuracy		
Schools	N-S ES	ES Dalton	ES RWCT
p-level	**.0000**	**.0000**	**.0008**

With regard to other than normal data distribution we used Kruskal-Wallis test to verify hypothesis H2.

H2   The absolute accuracy values gathered from learners differ in dependence on the type of educational programme.

$H2_0$ The absolute accuracy values gathered from learners are independent of individual educational programme (medians are equal).

The *p* - level value at Kruskal-Wallis test was *p* = .009. On a one percent significance level we reject the null hypothesis and confirm alternative hypothesis stating the existence of statistically significant differences at the level of metacognitive monitoring in dependence of the type of educational programme. The post hoc analysis to follow states where the statistically significant differences are to be found. The following table is a post doc analysis based on Dunn method (Dunn, 1964). The numbers in bold are to determine with which educational programmes statistically significant differences exist at values of metacognitive monitoring reached.

Tab. 6: Values found with post hoc analysis of metacognitive monitoring

r/s	|tr-ts| (critical value)		
	N-S ES	ES Dalton	ES RWCT
N-S ES	-----------	**42.616 (32.908)**	**83.949 (45.128)**
ES Dalton	**42.616 (32.908)**	-----------	41.333 (50.016)
ES RWCT	**83.949 (45.128)**	41.333 (50.016)	-----------

A statistically significant difference on one percent significance level has been proved in all areas researched, except for the areas of ES Dalton and ES RWCT.

The situation can be studied in detail on the base of quartile graphs. The individual quartile graphs visualise the results first (left) for N-S ES learner, further for ES Dalton learners and consequently for ES RWCT learners.
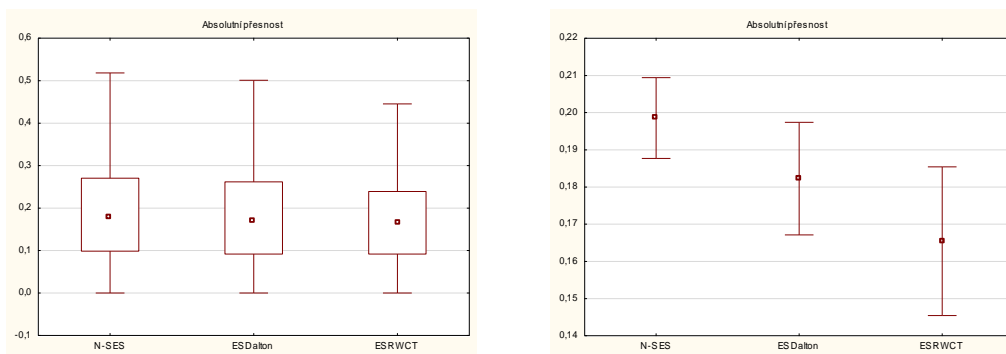
Fig. 2: Quartile graphs for results of the metacognitive monitoring level in a specific domain of reading with individual types of educational programmes.

With regard to the data interpretation it is vital for us to see if the values of metacognitive monitoring provide a statistically significant difference to a null value out of the reason that if the values reached for the field of absolute accuracy **do not differ** with statistic significance, so we may well state that a given group of learners is quite precise in their estimations on items answered correctly and incorrectly. The issue has been covered by hypothesis H3.

H3 The values gathered on metacognitive monitoring with learners will differ from the null value.

We tested the null hypothesis, that is, in generall, in the form: $H_0 : \tilde{x} = x_0$, compared to the alternativ one: $H_0 : \tilde{x} \neq x_0$. The testing was conducted using the sign test as well as Wilcoxon test of mean value for single selection, which is stronger (Hendl, 2012). Observed p-levels are in the following table.

The values in bold show if values reached with learners at a particular dimension of metacognitive monitoring is significantly different from null value.

Tab. 7: *P* - level values for Wilcoxon paired test for dimension of metacognitive monitoring.

Areas researched	P-level values for Wilcoxon paired test
N-S ES	***p* <.01**
ES Dalton	***p* <.01**
ES RWCT	***p* <.01**

The null hypothesis is to be rejected in all researched areas and we confirm the practical hypothesis of statistically significant differences of reached metacognitive monitoring values from the value of null.

Discussion

Significant differences on the level of understanding scientific contents have occurred with all three groups of schools. ES RWCT scored best (Me 9, Ø 8,78), followed by the ES Dalton plan learners (Me 9, Ø 8,43). The poorest results were with learners at standard N-S ES (Me 8, Ø 7,75). The results of our research have proved significant differences (p <.01) on the level of metacognitive monitoring in dependence of the type of educational programme. Both ES RWCT Me 0,15; Ø 0,17) together with ES Dalton plan (Me 0,16; Ø 0,19) learners have reached significantly better values (values closer to null) compared to learners at standard N-S ES (Me 0,17; Ø 0,21); with no noteworthy differences between learners at ES Dalton plan and ES RWCT. The absolute accuracy index in the field of research in pedagogical psychology means the level of how accurate our assumption of our own performance is compared to the real performance. The research of calibration scopes very widely and some of the results have an essential find in common (Kruger & Dunning, 1999): the level of accuracy of one´s own assumption correlates to the students´ success in the subject. We have verified the same through our research. Both ES Dalton plan and ES RWCT learners in comparison to standard N-S ES learners scored significantly better in reading comprehension test and at the same time they were more accurate in their assumptions. Independent of the genre, it has been proved that the readability of the text has an influence on the accuracy of the assumption. Garofalo and Lester (1985) drew attention to an individual´s subjective reception of the text´s attraction: preferred stories are easier to read than those unpopular with the

individual. Weaver and Bryant (1995) have found interaction effect between the genre and confidence judgement for various types of information (topic-based information with narrative texts are more revealing and practical texts are easier to guess for detailed information), nevertheless they made an essential conclusion: with no regard to the genre (just as well as it is in the case where an unbiased judge is estimated to be either easy or demanding), all texts lead to a more accurate estimation of their own performance in reading if they are within the optimum (adequate) scope of difficulty. In this respect the authors have concluded that such interaction effect has become non-existent through subjectively perceived difficulty. The confidence judgment for correct test answers referring to the read text is determined (see Conslusion 1) by the entire text difficulty (if it is to be found on the subjectively adequate level of difficulty; or, if on a general level the task/text is perceived favourably, and (see Conclusion 2) the character of the test answer (e.g. searching for an explicitly stated piece of information x making conclusions) (Weaver & Bryant, 1995). According to Prins et al. (Prins, Veenman, & Elshout 2006) we need to stress that the benefits from applying metacognitive processes are gained by the learner at the very moment when the text/task is found in a subjectively perceived (adequate) scope of difficulty. On the contrary, a high rate of knowledge may lead to the reduction of subjectively perceived difficulty of task and therefore there is no need to implement metacognitive approaches. The wider the domain-specific knowledge, the less needed it is to implement a strategy to enable for the task to be solved (Waters & Waters, 2010). They, who have previous knowledge, profit from the integration of new data into the existing concept. The presumption being that the knowledge "at hand" (terms) show content links to the knowledge which is to be gained.

If our long-term memory should file knowledge items with hardly any link to the previous knowledge, it is important to construct a new knowledge base first and through the means of higher cognitive processes. The higher level of understanding text of scientific content in ES RWCT and ES Dalton plan learners (contrary to N-S ES learners) would support the idea of a higher rate of previous knowledge relevant to the content with ES RWCT and ES Dalton plan learners, and at the same time a question would occur: How is it possible that ES RWCT and ES Dalton plan learners have disposed of significantly better level of absolute accuracy, so at implementing metacognitive monitoring, which is beneficial to the learner when the task is within a subjectively adequate range of difficulty (if an individual has a required knowledge base, needn´t they implement metacognitive processes? More poignantly we could solve the problem by drafting, verifying reliability and validity of and setting a pre-test finding the level of previous knowledge linked to a domain – topic (background knowledge). This factor deserves closer attention as domain-specific knowledge is according to some authors repeatedly shown as the strongest predicator of school performance (Nietfeld & Schraw, 2002; Watkins, Lei, & Canivez, 2007). More reasoning (see Conclusion 3) may be found in the effect caused by the learner´s expectation on the character of the test (test expectancy). The expectation has been influenced by previous experience and significantly influences matacognitive monitoring and correctness of answering the items. It has been proved through experiments (Thiede, Griffin, & Wiley, 2011) that in the very moment when the learner expects the items in the didactic test to be constructed primarily on a lower level cognitive level of thinking (memory test; it is a type of test designed so that it evaluates memory recollection of details specifically mentioned in the text), they show a higher rate of confidence at these items as well as they score better. On the contrary, where the expectations are of higher thinking operations (inference test; it is a type of test designed to evaluate the ability to derive conclusions of implicit information and textual ideas), they show higher rate of confidence as well as they are more successful at these items. Thiede, Griffin and Wiley, for an experiment, divided university students into two groups. They were both to read a series of three texts. After reading each text they were to express the confidence judgement how many out of the total number of 5 questions would be answered correctly in the following test. The procedure was repeated three times. The first group was instructed that the test given is to evaluate their ability to remember particular information included in the texts. The second group, also before the introduction of the test, was told that the tests are to evaluate their ability to link different parts of the text and make conclusions out of them (meaning the answer was not to be found in recalling explicit information from the text but on the grounds of an implicit level). In the second follow-up the students had 6 texts to read with the difference that the tests followed the last text´s evaluation. The order of texts was made random for each tested student. Every one of them was to answer in both the memory tests and inference tests where one half of the students were given expected tests (depending on which type they filled in the first round), and the other type of tests came after the first. The other half of the students were in a reversed order, meaning they were first given tests they had not expected (a different type of test compared to the one in the first round). The efficiency in the test and the accuracy of monitoring (in terms of total evaluation score in texts) was higher if students were given the expected type of test.

The significant relevance of high level of metacognitive monitoring is assigned to the method of teaching at ES RWCT and ES Dalton plan schools (Conclusion 4). ES RWCT schools apply E-U-R model of which its second phase (understanding) explicitly incites the application of such teaching methods and forms of teaching that involve active and conscious work with text (learning material) so that the learner realises what they know and what they do not know; what the intentions of the author are and what they found interesting, if the text offers contradictory information, how is textual information to be linked to previous knowledge etc. (Grecmanová & Urbanovská, 2007; Hausenblas & Košťálová, 2006). The explanatory cause of high level of metacognitive monitoring in ES Dalton plan learners is a find executed by Stipek, Feiler, Daniels and Milburn (1995), who found out that in an open environment, where the teacher provides their learners with more freedom a higher level of metacognition is found (learners use learning strategies at a higher rate and unless they meet with an obstacle, they are able to find a way to overcome it).

A specific cause for a higher calibration level with ES RWCT and ES Dalton plan learners could be the conditions of setting educational context. Roebers, Moga and Schneider (2001) in their study granted learners (six-, seven- and eight-year-olds) (a) plus 1 point for a correct answer, (b) minus 1 point for an incorrect answer and (c) 0 point for an "I don´t know" answer. (Points gained could later be exchanged for a material gift.) Roebers (2002) suggests that in experiments where learners are given the option of answering "I don´t know" (and in cases of certain insecurity to skip the question to heighten the overall accuracy of metacognitive monitoring), the learners still prefer to answer, which may be explained by the fact that the respondents simply try to justify their answer and that is why they link their confidence judgements towards higher confidence. Individuals commonly search the confirmation of what they think and they know that by seeking reasons why their answer should be true rather than they would seek reasons for why their answer should not be true (Koriat, Lichtenstein, & Fischhoff, 1980). In the moment when Koriat et al. (2001) replicated the research with the difference that the children had been instructed that should there be one incorrect answer they lose all points gained, the calibration was far more accurate and the learners filled in significantly fewer items (compared to the control group). The results of researches show that the level of metacognitive monitoring is higher in correlation to stricter rules applied. What are the actual conditions and criteria for learners´ evaluation (esp. in grading tests) at institutions participating in the research? There is strong need for further research in the field.

Conclusions and implications

ES Dalton plan and ES RWCT learners were more accurate in their estimation and that could have been caused due to (Conclusion 1) the fact that The Mystery of the Giant Fang was highly readable in the sense of individual paragraphs to be understandable and clear to the learners – and therefore were within the subjectively adequate range of difficulty. The second interpretation level (Conclusion 2) relates to the question character. Even though the tests finding the level of reading comprehension were enriched by more demanding items in comparison to the less challenging ones at a rate nearly 2:1 (as recommended for a design of didactic tests e.g. Škoda, Doulík, & Hajerová-Müllerová, 2006), it was not a problem for either ES RWCT or ES Dalton plan learners to score better than N-S ES. It is then possible that it was the character of items that had caused the higher scoring at the reading comprehension test as well as the higher level of calibration. The test items were subjectively perceived as adequately demanding. If it really was the case, it would be a proof of notable discrepancies in what is considered an adequate test question by ES RWCT and ES Dalton plan learners against N-S ES learners. The instant application level finds it more relevant to learn what agents have been at work with the occurrence of the explanatory cause as shown by our paper that ES RWCT and ES Dalton plan learners could subjectively perceive paragraphs in the text and items in the worksheet as adequately difficult (What are the causes of causes then?) What causes could it have that caused ES Dalton plan and ES RWCT learners consider the worksheet items of adequate difficulty? This question is a rather complex one and would definitely deserve a sole analysis going beyond the frame of this paper.

Concusion 3 refers to the phenomena of text expectancy. Significantly higher accuracy of ES RWCT and ES Dalton plan learners can give hints that these learners expect tasks of more challenging nature. Are the learners used to similar types of didactic tests? Is it possible to think that the actual learning activities in these types of educational programmes are focused on higher thinking operations? If this assumption was correct, we may acknowledge it. If we accept the premise, then learners at N-S ES may more often face questions of cognitively less demanding nature on the base of memorizing (in the context of the tools used in our paper, e.g.: Who found a petrified fang in England?, which may be alarming.

Sternberg (2002) says that metacognitive monitoring lies in the process of observing one´s own thoughts, where in principle we realise what we know and what we do not know (but what we also want to know and

what we perceive as contradictory information in relation to the previous knowledge or other information from the text etc.). The depth (quality) together with temporal side (quantity) of the process of understanding is quite individual. Every learner proceeds in the process of understanding at their own pace. Hádková and Jindráček (2012) show that intra-communication is a highly important factor and the teacher should provide the learner with enough time for their inner speech (see Vygotsky´s concept of metacognition as "inner speech"). The executed research supports the premise (Conclusion 4) that metacognitive monitoring is only beneficial in cases where the learner is given a chance to use their knowledge (and meta knowledge) at their disposal and control their learning according to their own needs (individual tempo). The above mentioned shows that the pedagogue should link teaching situations inciting the learner to deep understanding and at the same time there will be enough space for their own processing of the learning contents. Based on these findings we see the transmission of individual understanding in teaching contents through monologue methods into the "learners´ heads" at the tiniest inhibition of metacognitive development in the learners. ES RWCT along with ES Dalton plan learners are provided with a particular space to their own understanding through specific teaching method at ES RWCT (e.g. I.N.S.E.R.T. – Interactive Notating System for Effective Reading and Thinking), or setting all classes taught where the learner chooses freely the order of fulfilling the activities and they are solely responsible for the results of their work (ES Dalton plan). Such didactic approaches add to the development of a learner´s metacognitive potential. As Bransford et al. (Bransford, Brown, & Cocking, 2000) suggest, the development of metacognition im learners should be the standard practice of every teacher.

In conclusion it is essential to add that the absolute accuracy index values in all learners at all types of educational programmes significantly differed from null (absolute accuracy values did not reach such point where the level would not be notable from the exquisite fidelity in its absolute sense of the word – null value), which provides space for all teachers whose learners participated in the research to improve their accuracy (Conclusion 5). With accuracy in our performance judgement we improve the possibility to control our own new-knowledge/skill learning process and use the strategies designated for reaching our goals. In other words, with absolute accuracy the potential of our self-regulation rises (Greene & Azevedo, 2009).

Acknowledgements

References

[1] Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-Level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders, 25*(1), 65-83.

[2] Bjorklund, D. F., & Coyle, T. R. (1995). Utilization deficiencies in the development of memory strategies. In F. E. Weinert & W. Schneider (Eds.), *Memory development and competencies: Issues in growth and development* (pp. 161-180). Mahvah, NJ: Erlbaum.

[3] Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

[4] Braselton, S., & Decker, B. C. (1994). Using graphic organizers to improve the reading of mathematics. *The Reading Teacher, 48*, 276-281.

[5] Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: Perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology, 90,* 60–77.

[6] Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, *111*, 524–542.

[7] Dunn, O. J. (1964). Multiple contrast using rank sums. *Technometrics, 5*, 241-252.

[8] Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage Publications.

[9] Gambrell, L. B. (2011). Seven rules of engagement: What's most important to know about motivation to read. *The Reading Teacher, 65*(3), 172-178.

[10] Garofalo, J., & Lester Jr, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for research in mathematics education, 16*(3), 163–176.

[11] Grecmanová, H., & Urbanovská, E. (2007). *Aktivizační metody ve výuce, prostředek ŠVP*. Olomouc: Hanex.

[12] Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology, 34*(1), 18–29.

[13] Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429-457). New York: Psychology Press.

[14] Hádková, M., & Jindráček, V. (2012). *Princip názornosti ve výuce českého jazyka a literatury*. Ústí nad Labem: PF.

[15] Hausenblas, O., & Košťálová, H. (2006). Co je E-U-R: Podrobněji k fázi uvědomění si významu informací. *Kritické listy, 23*, 57-59 [online]. Retrieved from: http://rejskol.msmt.cz/. [cit. 08-08-2016].

[16] Hendl, J. (2012). *Přehled statistických metod: analýza a metaanalýza dat*. Praha: Portál.

[17] Janotová, Z., & Šafránková, K. (2013). *Čteme nejen v hodinách českého jazyka*. Praha: ČŠI.

[18] Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 163–182.

[19] Klooster, D. (2000). Co je kritické myšlení? *Kritické listy*, *1-2* [on-line]. Retrieved from: http://www.kritickemysleni.cz/materialy.php?co=kriticke_listy&co2=02/cojeKM. [cit. 12-09-2016].

[20] Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2) [on-line]. Retrieved from: http://www.communicationcache.com/uploads/1/0/8/8/10887248/reasons_for_confidence.pdf. [cit. 04-11-2016].

[21] Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134.

[22] Kruskal, W. H., & Wallis, A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association, 47*(260), 583-621.

[23] Kurtz, B. E., & Borkowski, J. G. (1984). Children´s metacognition: Exploring relations among knowldedge, process, and motivational variables. *Journal of Experimental Child Psychology, 37*, 335-354.

[24] Lamanauskas, V. (2012). Development of scientific research aktivity as the basic component of science education. *Journal of Baltic Science Education, 11*(3), 200-202.

[25] Lamanauskas, V. (2013). Natural science education importance in adolescence. *Journal of Baltic Education, 12*(4), 396-398.

[26] Leonesio, J. R., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16*, 464-470.

[27] Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research, 28*(4), 612-625.

[28] Listiana, L., Susilo, H., Suwono, H., & Suarsini, E. (2016). Empowering students´ metacognitive skills through new teaching strategy (group investigation integrated with think talk write) in biology classroom. *Journal of Baltic Science Education, 15*(3), 393-400.

[29] Maki, R. H., Foley, J. M., Kojer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16*, 609-616.

[30] Maki, R. H., & McGuire, M. J. (2002). Metacognition for text: Findings and implications for education. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied Metacognition* (pp. 39-67). Cambridge: University Press.

[31] Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology, 97*, 723–731.

[32] Narens, L., Nelson, T. O., & Scheck, P. (2008). Memory monitoring and the delayed-JOL Effect. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 137-155). New York: Psychology Press.

[33] Nelson, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not the absolute performance on an individual item: comments on Schraw (1995). *Applied Cognitive Psychology, 10*, 257–260.

[34] Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). Metacognitive monitoring accuracy andstudent performance in the classroom. *Journal of Experimental Education, 74*, 7-28.

[35] Nietfeld, J. L., & Schraw, G. (2002). The effect of knowledge and strategy training on monitoring accuracy. *The Journal of Educational Research, 95*(3), 131–142.

[36] Pressley, M., & Ghatala, E. S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist, 25*, 19-33.

[37] Pressley, M., Harris, K. R., & Marks, M. B. (1992). But good strategy instructors are constructivists. *Educational Psychology Review, 4*, 3-21.

[38] Prins, F. J., Veenman, M. V. J., & Elshout, J. J. (2006). The impact of interllectual ability and metacognition on learning: New support for the threshold of problematicity theory. *Learning and Instruction, 16*, 374-387.

[39] Roebers, C. M. (2002). Confidence judgements in children´s and adults event recall and suggestibillity. *Developmental Psychology, 38*, 1052-1067.

[40] Roebers, C. M., Moga, N., & Schneider, W. (2001). The role of accuracy motivation on children's andadults' event recall. *Journal of Experimental Child Psychology, 78*, 313–329.

[41] Roeschl-Heils, A., Schneider, W., & Kraayenoord, Ch. E. (2003). Reading, metacognition and motivation: A follow-up study of German students in Grades 7 and 8. *European Journal of Psychology of Education, 28*(1), 75-86.

[42] Sadi, Ö, & Uyar, M. (2013). The relationship between self-efficacy, self-regulated learning strategies and achievement: a path model. *Journal of Baltic Science Education, 12*(1), 21-33.

[43] Sen, S. (2016). The relationship between secondary school students' self-regulated learning skills and chemistry achievement. *Journal of Baltic Science Education, 15*(3), 312-324.

[44] Serafini, F. (2013a). Close readings and children's literature. *The Reading Teacher, 67*(4), 299-300.

[45] Serafini, F. (2013b). Supporting Boys as Readers. *The Reading Teacher, 67*(1), 40-42.

[46] Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples*). Biometrics, 52*(3-4), 591-601.

[47] Schleifer, D. (2009). Metacognition and performace in the accouting classroom. *Issues in accounting education, 24*(3), 339 – 367.

[48] Schneider, W., & Artelt, C. (2010). Metacognition and mathematics education. ZDM: The International *Journal on Mathematics Education, 42*, 149–161.

[49] Schneider, W., & Bjorklund, D. F. (1998). Memory. In W. Damon, D. Kuhn, & R. S. Siegler (Eds.), Handbook of child psychology: *Cognition, perception, and language (Vol. 2,* pp. 467–521). New York: Wiley.

[50] Schneider, W., & Pressley, M. (1997). *Memory Development Between Two and Twenty*. Psychology Press.

[51] Schneider, W., Schlagmüller, M., & Visé, M. (1998). The impact of metamemory and domain-specific knowledge on memory performance. *European Journal of Psychology of Education, 13*, 91–103.

[52] Schraw, G. (1998). Promoting general metacognitive awereness. *Instructional science, 26*(1-2), 113-125.

[53] Soukup, P. (2013). „Věcná významnost výsledků a její možnosti měření." *Data a výzkum – SDA Info, 7*(2), 125-148.

[54] Sternberg, R. J. (2002). *Kognitivní psychologie*. Praha: Portál.

[55] Stipek, D., Feiler, R., Daniels, D., & Milburn, S. (1995). Effects of different instructional approaches on young children's achievement and motivation. *Child Development*, *66*(1), 209-223.

[56] Škoda, J., Doulík, P., & Hajerová-Müllerová, L. (2006). *Zásady správné tvorby, použití a hodnocení didaktických testů v přípravě budoucích učitelů* [on-line]. Retrieved from: http://cvicebnice.ujep.cz/cvicebnice/FRVS1973F5d/. [cit. 07-10-2016].

[57] Thiede, K. W., Griffin, T. D., & Wiley, J. (2011). Test expectancy affects metacomprehension accuracy. *British Journal of Educational Psychology*, 81, 264–273.

[58] Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. S. (2009). Metacognitive monitoring during and after reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition andself-regulated learning* (pp. 85–106). New York: Routledge.

[59] Thiede, K. W., Redford, J. S., Wiley, J., & Griffin, T. D. (2012). Elementary School Experience With Comprehension Testing May Inluence Metacomprehension Accurancy Among Seventh and Eighth Graders. *Journal of Educational Psychology*, *104*(3), 554-564.

[60] Tolson, H. (1980). An adjuct to statistical significance ω2. *Research Quarterly for Excercise and Sport*, *51*(3), 580-584.

[61] Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences*, *15*, 159-176.

[62] Vilenius-Tuohimaa, P., Aunola, K., & Nurmi, J. (2008). The association between mathematical word problems and reading coprehension. *Educational Psychology*, *28*(4), 409-426.

[63] Wang, M., Haertel, G., & Walberg, J. H. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Psychology*, *84*, 30-43.

[64] Waters, H. S. (2000). Memory strategy development: Do we need yet another deficiency? *Child development*, *71*(4), 1004–1012.

[65] Waters, H. S., & Schneider, W. (2009). *Metacognition, Strategy Use, and Instruction*. Guilford Press.

[66] Waters H. S., & Waters, T. E. (2009). Bird experts: A study of child and adult knowledge utilization. In H. S. Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 113-134). New York: Guifold Press.

[67] Watkins, M. W., Lei, P. W., & Canivez, G. L. (2007). Psychometric intelligence andachievement: A cross-lagged panel analysis. *Intelligence*, 35, 59–68.

[68] Weaver, C. A. I., & Bryant, D. S. (1995). Monitoring of comprehension: The role of text difficulty in metamemory for narrative and expository text. *Memory & Cognition*, *23*, 12-22.

[69] Winne, P. H. (1995). Inherent details on self-regulated learning. *Educational Psychologist*, *30*(4), 173-187.

[70] Winne, P. H., & Hadwin, A. (1998). Studying as Self-Regulated Learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277-304). Mahwah: Erlbaum.

[71] Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*(1), 3-17.