

Developmental differences in dynamic indicators of three variously simple cognitive sub-systems functioning at girls and boys aged 8-17 years

Mislav Stjepan Žebec^{1*} & Katja Kaurić²

¹Institute for Anthropological Research, Zagreb

²Rimac Technology

* Corresponding author: mislav.stjepan.zebec@inantro.hr

Received February 10th, 2025

Accepted for publication June 26th, 2025

Online First July 7th, 2025

Keywords: reaction time (RT), performance dynamics indicators (PDI), task complexity, cognitive development, sex differences.

Abstract

Mainstream post-Piaget cognitive development researchers mostly ignored performance dynamics behind the total paper-and-pencil tests score, or average reaction time (RT) on computerized tests. This research focuses on several indicators of performance dynamics while solving three simple computerized cognitive-motor tests of various complexity. To get developmental picture of the related cognitive subsystems dynamics, the tests were solved by girls (N=228) and boys (N=235) aged 8-17 years.

Participants were students of a primary and a secondary school from Zagreb (Croatia) that individually solved three simple tests of MID KOGTESTER-1 computerized reaction meter, which assessed predominantly perceptual and working memory functioning. Four RT dynamic indicators (minimal, maximal and average time of cognitive task solving, and average time of non-optimal cognitive task solving) were mostly positively correlated, sharing an average variance of 36% - thereby presenting non-redundant measures of cognitive dynamics. On a descriptive level, age-related decrement of all four indicators was non-linear, steadier for girls and suggested girl's superiority in the performance dynamics. The non-trivial and statistically significant results: (1) performance dynamics indicators improved across developmental phases with different intensities, the most in word recognition (WR) task and the least in choice reaction time (CRT) tasks, with the similar pattern for girls and boys; (2) across entire developmental period girls outperformed boys in WR and CRT task, but not in simple reaction time (SRT) task (the weakest advantage appeared in minimal time of cognitive task solving). Dynamics of cognitive subsystem functioning gives more complete picture of related cognitive performance and its development, based on neural structure and its dependence on age and sex.

Introduction

Fundamental cognitive development research, which is focused on the construction or evaluation of theoretical models of cognitive subsystems or human mind development, traditionally use validated instruments/tests of targeted cognitive constructs (Anderson et al, 2001; Demetriou and Kyriakides, 2006; Demetriou et al, 2013; Hicks and Bolen, 1996; Kail, 1997; McArdle et al, 2002). These instruments/tests usually report some total score in the test (whether it is

paper-and-pencil or computerized test) or some average time of test item solving. This kind of report gives quite exact measure of person's average capacity to solve the tasks from the operative domain of the cognitive system under the study. Therefore, by looking these reports across months or ages of human developmental period, we will get an averaged picture of girls' and boys' joined capacity to solve the tasks in domain where the studied cognitive system operates.

However, could we get something more from the set of carefully chosen and validated cognitive tasks that are usually quite numerous (and demanding) with an aim to reliably assess the studied cognitive functions? Can we use the set of answered task outcomes – with and without errors – to get additional information (besides average capacity) on related cognitive system functioning? More precisely, can we assess: (1) how stable related cognitive system was while answering the tasks, or (2) what is its upper limit (i.e. best performance/potential) or (3) lower limit (the worst performance/operative weakness) during functioning?

This information are definitively useful in educational or vocational counseling of children and youth based on cognitive capacity (Gottfredson, 2003; Hodge, 1999; Metz and Jones, 2013; Wai et al, 2018), but can we use them to get more thorough insight in cognitive development dynamics in general? Can they tell when the cognitive system structure starts to change and when the change minimizes with a new harmonized and stable structure? Do stability indicators, or upper and lower performance limits, differ for girls and boys that mature with somewhat different pace (Giedd et al, 2012; Lenroot et al, 2007; Thomas and French, 1985; Žebec et al, 2014)? Do performance indicators differ in simple and complex cognitive activities? Can they tell us how to design some long-term cognitive training or learning process based on a cognitive system under the study?

There are number of arguments from the other domains of cognitive psychology (fundamental or applied) that indirectly suggest positive or extended answers to these refined developmental questions.

Cognitive performance dynamics in non-developmental research

Intraindividual variability in human performance has been studied from perspective of various theoretical approaches and most of them concluded that it is not the consequence of cognitive system noise, but an inherent characteristic of the human that presents cognitive system stability (Boker and Nesselroade, 2002; Rabbitt et al, 2001; Slifkin and Newell, 1998). Stability of cognitive task performance (expressed via RT intraindividual variability) has been widely explored

in processing speed-intelligence research (Deary, Der and Ford, 2001; Jensen, 2006; Neubauer et al, 1997). Theoretical models of an individual's RT distribution generated by performance in various cognitive-motor RT tests all include the σ or s parameter of intraindividual variability (Brown and Heathcote, 2005; Ratcliff and Smith, 2004; Ratcliff, Van Zandt, McKoon, 1999). Finally, the handbooks are written on human behavioral instability that cover various topics of this phenomenon: the role of intraindividual variability in cognitive-motor development, intraindividual variability and mood regulation and self-representation, interindividual variability as a measure of aging process and of vulnerability and resilience (Diehl et al, 2015).

When we move to the lower and upper limits of cognitive (sub)systems engaged in tasks solving, then we find intriguing research on worst and best RT task performance. The lower limit, i.e. worst RT performance in perceptual-motor RT test has been intensively studied under the topic known as *worst performance rule* (WPR). This rule states that during some cognitive-motor RT test solving, the longest RT correlates to intelligence scores more than average or minimal RT (Kranzler, 1992; Larson and Alderton, 1990; Schubert, 2019). This interesting finding actually says that the slowest responses in RT test tell us more about intelligence than the shortest one, or the average of all RTs, and possible explanation researchers looked at working memory (WM) functioning, especially in attention mechanisms of WM central executive (Coyle, 2003; Schmiedek et al, 2007; Unsworth et al, 2010).

The best (or the shortest) response time in cognitive-motor RT tests, although included in previously mentioned WPR research, showed its value predominantly in more applied studies. Because it presents someone's upper limit of RT-performance, it is quite sensitive to environmental conditions of the person engaged in RT test and her/his actual health status. Therefore, minimal RT indicator of performance dynamics various researchers used to analyze the impact of hyperbaric pressure and nitrogen narcosis in shallow air-diving (Petri, 2003), or to analyze the consequences of spontaneous menstrual cycle and of following oral contraceptives on cognitive motor

functioning (Becker et al, 1982). Moreover, minimal RT proved to be a useful indicator of kidney transplantation effects on cognitive and psychomotor functioning (Radić et al, 2011) and for difference analysis in complex psychomotor RT between patients with and without cerebral circulatory disorders signs (Bobić et al, 2002). Besides that, researchers used minimal RT to study RT sex differences at top sprinters (Lipps et al, 2011), but also in more fundamental research of mental processing dynamics (Drenovac, 2001, 2009).

Cognitive performance dynamics in developmental research

The answers on the above-mentioned refined developmental questions - on getting more thorough insight into cognitive development and related consequences from performance dynamics in the set of carefully chosen cognitive tasks - that come directly from cognitive development research, included mostly interindividual variability of performance, but only exceptionally it's lower and upper limits.

Intraindividual variability in cognitive development has been mostly studied within dynamic systems theory (DST) approach to cognitive development, where the authors defined it as a crucial mechanism of development (Smith and Thelen, 2003; Thelen and Smith, 1998; van Geert and van Dijk, 2002). Within this approach, the authors studied intraindividual variability in various developmental areas: in infant behavior (de Weerth et al, 1999), in manual reaching behavior during middle childhood (Golenia et al, 2017), in language development (van Dijk and van Geert, 2007), and many others.

Closely to DST, K. Fisher pointed to important role of intraindividual variability in his dynamic skill theory (Fischer and Yan, 2002; Yan and Fischer, 2002). Nevertheless, the other researchers of cognitive development also recognized the importance of intraindividual variability as a presumption of development and as a measure of cognitive functioning stability (Demetriou et al, 2013; Siegler, 1994).

Finally, intraindividual variability change across the lifespan was the research topic of authors interested in

sex differences of this phenomenon (Deary and Der, 2005; Dykiert, 2012).

However, despite valuable findings of these developmental research of cognitive performance dynamics in answering the refined questions of the cognitive development, they did not include some useful dynamic indicators and methodological specificities of experimental cognitive research of reaction time (RT) in cognitive tasks. For example, these developmental researches mostly did not include upper and lower limits of cognitive performance that define the range of person's cognitive capacity (one rare exception was Žebec et al study, 2014). In addition, they did not differentiate cognitive from motor development in RT cognitive-motor tasks performance and although detangling cognitive from the motor component of RT response is an issue from embodied cognition perspective (Anderson, 2003; Wilson, 2002), there are models on how to do it, at least partially (e.g. Hick paradigm, described in Jensen, 2006, or in Neubauer, 1990). This differentiation is important because there are sources of specific motor variability that are a consequence of specific anthropometric features (e.g. hand dimensions, muscles development), minor somatic health confinements (e.g., ophthalmological, previous hand injuries), and previous experience in manual activities (e.g. related to computer games, specific sports). This motor variability confounds cognitive development assessment conducted by RT measures, designed in a way that demands manual actions. Finally, previous cognitive development research based on performance dynamics in cognitive-motor tests did not differentiate correct answers (to the test tasks) performed after errors, from correct answers performed after previous correct answer. This differentiation is important because errors in the previous RT task usually modify answering of the actual RT task while, depending on the subject's personality (e.g. emotional stability), they can prolong or even shorten RT answer of the actual task (Burns, 1971; Drenovac, 2009). In that case, we will not get a clear RT measure of targeted cognitive process.

In order to analyze cognitive development with as many indicators of cognitive-motor performance dynamics as we can derive it from an individual RT

distribution, by including motor and cognitive development differentiation and measures of targeted cognitive process decontaminated from errors impacts, we had to design developmental research with specific RT equipment. Moreover, to avoid confounding effects of girls' and boys' different developmental pace we included enough female and male participants and analyzed their performance in cognitive tasks separately.

Measures of cognitive development are numerous, but thorough insight in cognitive development asks for sensitive, real-time indicators that correspond to neurobiological bases of cognitive functions and, at the same time are noninvasive and applicable to larger groups of participants. Therefore, we used RT on cognitive-motor tasks that reveal the functioning of basic cognitive processes like perception and working memory (which contains attentional processes).

For describing performance dynamics of related cognitive sub-systems, we have chosen four indicators of performance dynamics in cognitive RT tasks that have been proved as valuable in the literature. Those were best (minimal RT), worst (maximal RT), and average time of cognitive task solving (mean RT), but also average time of non-optimal cognitive task solving (i.e. average non-optimal performance instability measured by mean deviation from best RT).

All previously discussed arguments directed this study toward empirically addressing several key issues. To check whether:

1. The four indicators of the cognitive component of simple cognitive-motor task solving separately contribute to the dynamics description of related cognitive sub-system functioning (i.e. whether they justify their inclusion in the description of activated cognitive sub-systems),
2. The four cognitive indicators of cognitive-motor task-solving dynamics improve across ages of the observed developmental period, and how possible improvement looks like,
3. Possible change of cognitive indicators of cognitive-motor task performance dynamics across

developmental phases depends on (i) the dynamics indicator's type, (ii) cognitive task complexity, and (iii) person's sex (female/male),

4. There are sex differences in cognitive indicators of task performance dynamics and how they depend on (i) the dynamics indicator's type and (ii) cognitive task complexity.

Answering the above-mentioned problems should enable us to conclude whether the analysis of the cognitive component of cognitive-motor task-solving dynamics could give more complete and more correct picture of the cognitive development of perception and working memory systems and of cognitive functioning in general.

Materials and methods

Materials

The study included conducting three cognitive-psychological instruments, but only computerized reaction meter MID-KOGTESTER1 (Žebec, 2005; Žebec, Demetriou, Kotrla Topić, 2015) produced data that were analyzed to answer before mentioned research problems. The rest of two cognitive-psychology scales were used to define study sample.

MID-KOGTESTER1 has been used to evaluate the indicators of task performance dynamics in three variously simple cognitive-motor tests. This instrument is a computer-based battery of eight simple visual-motor tests for the assessment of the control and rate of human information processing, and aspects of working memory. In this study, three tests were used: Simple reaction time (SRT) test, Word recognition go-no go test (WR) and Choice reaction time (CRT) test.

MID-KOGTESTER1 components are:

- (1) A laptop with an installed program designed to generate stimuli and record the subject's responses, i.e. reaction time (RT) in milliseconds. The RT is recorded separately for the cognitive and motor components of the response (according to the Hick paradigm), and errors are recorded. The program differentiates correct answers after error from correct answers after previous

correct answer and calculates various performance dynamics indicators (PDIs).



Figure 1. MID KOGTESTER 1 instrument

(2) A computer monitors on which the stimuli are displayed in various colors. *Note:* The monitor based on LCD technology was not used to control interruptions in the rendering of visual stimuli, caused by periodic status checks by the computer's operating system.

(3) Two panels with response keys. On the upper surface of the first panel, there are five response keys, or target keys, and they are arranged in a semicircle with equal distances from the (semicircle's) center. In the semicircle's center, there is a start key for initiating the stimulus. On the upper surface of the second panel, there are three horizontally arranged keys at equal distances, with the central key having the function of task starting. In answering the tasks of the study tests (SRT, WR, and CRT) only the first response panel has been used.

SRT, WR, and CRT tests consisted of *trial sequence* and *testing sequence* of various numbers of elementary cognitive tasks (ECTs).

The SRT test contains 20 ECTs in a testing sequence, with a previous trial sequence of 10 of them. As a stimulus, six identically colored characters X, in red, white, blue or green, appear on the monitor, in random order and with an unpredictable appearance interval (0.75 - 2.5 seconds). The participant's task is to raise the index finger as quickly as possible from the start key and press the target key, which is located vertically above the start key, as soon as (s)he notices the appearance of Xs of any color.

The WR test contains 30 ECTs in the testing sequence, with a previous trial of 10 ECTs. As a stimulus, a word (written in magenta color) appears on the monitor representing the name of one of the four colors (BLUE, GREEN, WHITE, and RED), or the word COLOR. The word COLOR is the target stimulus to which the participant must respond as soon as possible, and the color names are distractors to which the participant must not respond. The order of appearance of the target stimulus in relation to the distractors is random, and the time between the end of responding and the appearance of a new stimulus varies randomly (0.75 - 2.5 seconds). The participant's task is to raise the index finger from the start key as quickly as possible and press the same target key, which is located vertically above the start key, after the target stimulus appears. If a distractor appears, the participant must not lift her/his finger from the start key, but wait patiently for the distractor to disappear. Although this test primarily measures cognitive inhibition function, it also represents a choice reaction test with two choices - to respond to the target stimulus, or not to do so to the distractor.

The CRT test contains 32 ECTs in a testing sequence, and 12 of them in the previous trial sequence. The stimuli are the names of colors (BLUE, GREEN, WHITE, and RED) written in magenta color and appear in a random order with an unpredictable interval between the end of responding and the appearance of a new stimulus (0.75 - 2.5 seconds). The participant responds to one of the four target keys, according to the rule given during the previous specific instruction. The participant's task is, after perceiving the color name on the monitor, to raise the index finger as quickly as possible from the start key and press the target key in accordance with the specific rule. This is a four-choice RT test since it demands from participant to respond to the appearance of one of four different color names by pressing one of four different keys.

Discriminability of the MID KOGTESTER-1 is very high since the RTs are measured with the precision of one millisecond (SRT, as the simplest test differentiated 97% of the tested participants). Test-retest reliability with one-year time lag ranged from 0.722 (SRT test) to 0.787 (WR test), which is very satisfactory.

After recording all participants' responses in the test, MID KOGTESTER-1 calculates PDIs (*performance dynamics indicators*).

Minimal RT (or best performance) in all three tests is denoted as *ct0min* and is defined as the shortest response time of the cognitive component of correct response after previous correct response on ECT (T^0) of the test SRT (*ct0min_1*), WR (*ct0min_2*) and CRT (*ct0min_3*). It is represented in milliseconds and the calculation formula is presented in Appendix 1.

Maximal RT (or worst performance) in all three tests is denoted as *ct0max* and is defined as the longest response time of the cognitive component of correct response after previous correct response on ECT (T^0) of the test SRT (*ct0max_1*), WR (*ct0max_2*) and CRT (*ct0max_3*). It is also represented in milliseconds and the calculation formula is written in Appendix 1.

Mean RT (or average task-solving performance) in all three tests is denoted as *cat0* and is defined as the average time of the cognitive component of correct response after the previous correct response on ECT (T^0) of the test SRT (*cat0_1*), WR (*cat0_2*) and CRT (*cat0_3*). This indicator is calculated by using T^0 and the number of all correct responses after the previous correct response (N^0) and it is represented in milliseconds. The calculation formula is presented in Appendix 1. Note: *cat0* presents an inverse measure of average task performance (larger *cat0* means worse average task performance).

Mean deviation of response time from the best RT (or average time of non-optimal cognitive task solving, i.e. average non-optimal performance instability) in all three tests is denoted as *catnof0* and is defined as the average time of non-optimized cognitive functioning during correct answering of ECTs of the test SRT (*catnof0_1*), WR (*catnof0_2*) and CRT (*catnof0_3*). It represents the average deviation of the cognitive component of RT in all correct answers (given after the correct answer), from the cognitive component of the best answer. It is represented in milliseconds and the calculation formula is written in Appendix 1.

The other two instruments were cognitive-psychology scale and checklist that were used to select

appropriate sample of the target population: right-handed pupils of the age range 8-17 years, without health confinements relevant for visual-motor answering on elementary cognitive tasks.

Hand dominance scale (Tadinac, 1993), consisting of 12 questions considering which hand participant uses in 12 common life situations (e.g. *Which hand do you write with?*, or *Which hand do you hold the scissors with when cutting?*), was conducted to determine students' hand dominance, since only right-handed students participated in the study (85-90% of the population). Participants responded by selecting one of five possible answers ($-2 = \text{always with left}$, $-1 = \text{usually with left}$, $0 = \text{with one and the other hand}$, $1 = \text{usually with right}$, $2 = \text{always with right}$) and they were classified as right-handed if the total score (on all 12 questions) was higher than 8. The questionnaire was administered in groups, in each class separately. The author of the scale (Tadinac, 1993) did not report psychometric characteristics of the scale (reliability, validity), but insight in the related results within our study suggests that there were no statistical preconditions for doing that. Namely, standard psychometric calculation of reliability and validity coefficient is based on Pearson correlation among response values of the instrument's items (questions/statements) and the most of prerequisite assumptions for Pearson correlation calculation were violated in data measured by *Hand dominance scale administration* (highly asymmetric and bimodal distribution, variability reduced to only 5 values, sub-interval measurement scale).

Students' health status scale (Žebec, 2005), consisting of 12 questions, was conducted to determine students' health status focused on health characteristics relevant to cognitive-motor functioning (e.g. *Does your child distinguish colors well?*, or *Does your child have persistent or frequent difficulties with attention?*). The scale completed the parents of student-participants of the study, one or both of whom were present at the parent-teacher conference organized before other cognitive functions measurements (primarily necessary to determine the final sample of participants). Parents responded by answering "yes" or "no" and if their answers confirmed more than one relevant health problem, the child was not included in the research

sample (one confirmed health problem was taken into account when conducting the MID-KOGTESTER 1 measurement). Psychometric characteristics of the scale (reliability, validity) were not calculated because the scale was conceptualized as heterogeneous checklist with rigorous exclusion criteria, not as standard questionnaire.

Research design

Measurements were part of the larger developmental study (Žebec et al, 2015) and were organized by using two research designs:

(1) *cross-sectional correlational design* ensured data collection needed for answering the first study problem (dependence among four indicators of cognitive task performance and inferring on their separate contribution to related cognitive functioning description)

(2) complex quasi-experimental mixed design with between-groups component (defined by independent variables of age, developmental phase, and sex) and within-subjects component (defined by independent variables of PDI's type and task complexity), enabled data collection needed for answering all other research problems (main and interaction effects of age/developmental phase, sex, task complexity and PDI's type on indicator's magnitude). Additionally, the observed, or dependent variables (DVs) were values of four PDIs (*best, worst and average performance, and non-optimal performance instability*), while independent variables (IVs) were previously mentioned in quasi-experimental mixed design.

Participants

The research participants were students of one elementary and one high school in the City of Zagreb, and as such they represent a convenient sample from the population of healthy children and youth aged 8 to 17, in urban areas of the Republic of Croatia. Based on the health status of the participants recorded by the scale, those whose health problems may be relevant to the subject of measurement or are of such a nature that it is not possible to speak with certainty about normal cognitive development, are excluded. A total of 463

right-handed participants took part in the study (228 were girls) and a more detailed age and sex structure is shown in the following table.

Table 1. Age-sex structure of study participants' sample

		sex		Total
		female	male	
age	8	34	46	80
	9	28	33	61
	10	17	6	23
	11	25	30	55
	12	19	16	35
	13	28	29	57
	14	8	12	20
	15	27	23	50
	16	21	19	40
	17	21	21	42
Total		228	235	463

Note: The age-sex groups associated with the ages of 10, 12, and 14 are noticeably less numerous than other age groups for two reasons: (1) the research was organized so that only 6 school ages were included - 2nd, 3rd, 5th, and 7th grade of elementary school and 1st and 3rd grade of high school; (2) in each of the 6 examined school ages, one age was dominant, but - due to the specific relationship between the date of birth and the date of measurement - part of the students of the given school age moved to "neighboring" age groups. In this sense, numerical uniformity of age and sex groups was not expected *a priori*.

Institute of the Improvement of Education at The Ministry of Education and Sports of Republic of Croatia has ethically approved the study, and the parents of the students, study participants, gave informed consent.

Procedure

In the first part of the research (as preparation for ECT performance measurement via MID KOGTESTER-1), *Hand dominance* and *Health status scales* were applied. Due to the organization of the stimuli on the MID KOGTESTER-1, which is adapted to right-handed people, participants with a dominant left hand were excluded from the study. Additionally, those with health dysfunctions that could be relevant for the measurement validity (for example, color blindness or hand motor disorders) or for the aim of measurement, i.e. normal cognitive development (for example, brain damage or color blindness) were also excluded from the study sample.

In the main part of the research, an individual examination was conducted for the duration of one school hour. It was done in school classrooms, specially organized to ensure the conditions of undisturbed individual examination. Before starting the MID KOGTESTER-1 tests, the participants were given a general instruction about the instrument, which explained the basic concept of the measurement, the content of the stimuli that appear on the screen, the way to answer on the answer boards, and the different types of errors when answering and the accompanying warning sounds. Additionally, before each ECT test (SRT, WR, CRT), they received specific instructions for solving it, in which the participant gets to know in detail the relevant stimuli and the specific way of responding to each of them. Moreover, they were advised to sit comfortably on the chair in front of the instrument in a position that would enable them to clearly perceive the stimuli and answer the tasks accurately and quickly, and to ignore potential distractors. Before each test, a trial sequence of the corresponding ECT was conducted in order for the participants to familiarize themselves with the stimuli and practice the way of answering, and to eliminate the possibility of misunderstanding the specific instructions.

During solving the MID KOGTESTER-1 test the participant dictates the pace of assigning tasks, because the display of the stimulus item starts only after the participant places the index finger on the start key. Initially, a fixation point appears on the monitor for a

duration that randomly varies from 0.75 to 2.5 seconds. It has the function of a warning signal since it announces the stimulus appearance, and the participant is obliged to look at the point. After the stimulus appears on the monitor, the participant should raise the index finger from the start key as quickly as possible, press the corresponding target key, and then return the index finger to the start key. At the same time, the computer measures the time from the stimulus rendering until the moment the index finger is lifted from the start key (*cognitive processing time*) and from the moment the index finger is lifted from the start key until the target key is pressed (*movement time*). In addition, various possible errors are registered and followed by a sound warning, after which the subject returns the index finger to the start key and waits for a further stimulus.

Before the test, the importance of accuracy first, and then speed, is emphasized to the participants. The tests do not always appear in the same order, so STR can be the 1st or 2nd test, WR the same, and CRT the 3rd, 4th, or 5th test in order in the entire battery of eight tests, which varies from participant to participant (control of the exercise in within-subjects research design). In addition, the four possible layouts of target keys for the color are changing across the participants, thus controlling the strategy transfer among them and reducing the possible effect on test performance.

Results

Data were previously screened for outliers and then statistically processed by the IBM SPSS statistical package. All dependent variables (DVs) were measured as ratio scale data and therefore we used predominantly parametric inferential statistics (except in cases when other presumptions were significantly violated).

Statistical analyses were organized in relation to research problems but in the developmental aspect it differentiated two types of analysis: (1) *Age-related analysis* of the cognitive indicators of tasks performance dynamics (*Research problem 2*), (2) *Developmental phase's related analysis* of the same indicators (*Research problems 1, 3, 4 and 5*). This differentiation has been partially generated by violation

of factorial ANOVA assumptions related to equal age-sex sample sizes and homogeneity of variances, if we have used age as independent variable (IV) in answering *Research problems 3, 4 and 5*. Nevertheless, the differentiation of developmental statistical analysis primarily shaped the theoretical framework of cognitive development, which states that developmental changes mainly occur between successive developmental phases (Demetriou et al., 2018). On the other hand, within developmental phases, there should be no significant change, i.e. cognitive functions and their relations should be mostly stable.

Research problem 1 analysis and related results

In order to check if the four indicators of the cognitive component of simple cognitive-motor task solving *separately contribute* to dynamics description of related cognitive sub-system functioning, we applied correlational analysis. Namely, if these indicators represent mostly separate aspects of performance dynamics, then they should represent mostly independent dimensions of variable-vector space for dynamics description. In other words, *ct0min*, *ct0max*, *cat0*, and *catnof0* should not be highly correlated, i.e. they should not share more than 50% of the variance.

To calculate correlations among four indicators on homogenous participants' samples (to avoid the effect of developmental variables that modify the correlations as a hidden third variable) we decided to perform calculations on participants' subsamples defined by cognitive developmental phases. Namely, DVs are cognitive variables and within these phases cognitive and relevant variables should be mostly stable. By using Demetriou et al (2018) theoretical model of cognitive development, we identified four developmental phases (Dev.Phase) in our participants' sample: (i) 6 – 8 years (*emerging rule-based representations*), $n_{DP1}=80$; (ii) 9 – 11 years (*integration of rules into rule-based systems*), $n_{DP2}=139$; (iii) 12 – 13 years (*emerging principle-based representations*), $n_{DP3}=92$; (iv) 14 – 17 years (*integrated principles*), $n_{DP4}=152$.

On these subsamples, we calculated Pearson's correlations on all six pairs of PDIs at all three tests (SRT,

WR, CRT). Applying calculation on various tests would raise the generalizability of the research problem findings. The results of this calculation are given in Table 2.

Table 2 reveals several important findings:

- (1) Around 95% (of 72) of correlations were statistically significant and all *insignificant* ones were correlations between *ct0min* and *catnof0*.
- (2) The average of all correlations was 0.599 meaning that the average percentage of shared variance (or common factors of the variables) among PDIs equals 35.9%.
- (3) The weakest correlation stands between *ct0min* and *catnof0* (average *ct0min-catnof0* correlation across tests and Dev. Phases was 0.219).
- (4) The strongest correlation stands between *cat0* and the other three indicators, depending on the test: (a) in SRT test $r(cat0-ct0min)= 0.807$, (b) in WR test $r(cat0-ct0max)= 0.704$ and (c) in CRT test $r(cat0-catnof0)= 0.842$.
- (5) In most of the tests (i.e. SRT and WR) the strongest correlation appears in 2nd Dev. Phase, while in most of the tests (i.e. SRT and CRT) the weakest correlation appears in 4th Dev. Phase.

Research problem 2 analysis and related results

Since previous correlational analysis suggested that all observed performance dynamic indicators, except *average performance*, mostly reflect activation of specific combination of neurocognitive resources, we found it interesting to check whether these resources show specific developmental trajectories in the age range 8 – 17.

In order to do that, we rendered age-related trajectories of four indicators (*cat0*, *catnof0*, *ct0min*, *ct0max*) in all three tests (SRT, WR, CRT), separately for girls and boys. These trajectories are shown in the next 12 figures following Table 2.

Table 2. Pearson inter-correlations of four indicators of cognitive performance dynamics (*ct0min*, *ct0max*, *cat0*, *catnof0*) in three simple cognitive-motor tasks of various complexity (SRT, WR, CRT) across four theoretically based developmental phases (Dev. Phase).

test type	correlational pairs	Pearson correlation (r)			
		1. Dev. Phase	2. Dev. Phase	3. Dev. Phase	4. Dev. Phase
SRT	ct0min - ct0max	0.517**	0.582**	0.589**	0.474**
	ct0min - cat0	0.790**	0.798**	0.786**	0.853**
	ct0min – catnof0	0.413**	0.417**	0.213*	0.246**
	ct0max – cat0	0.661**	0.782**	0.858**	0.707**
	ct0max – catnof0	0.694**	0.751**	0.709**	0.706**
	cat0 – catnof0	0.834**	0.817**	0.640**	0.612**
WR	ct0min - ct0max	0.238*	0.391**	0.343**	0.386**
	ct0min - cat0	0.642**	0.700 **	0.594**	0.772**
	ct0min – catnof0	-0.033	0.067	-0.223*	-0.126
	ct0max – cat0	0.631**	0.769**	0.71**	0.705**
	ct0max – catnof0	0.696**	0.69**	0.614**	0.609**
	cat0 – catnof0	0.532**	0.453**	0.458**	0.418**
CRT	ct0min - ct0max	0.381**	0.613**	0.757**	0.361**
	ct0min - cat0	0.675**	0.797**	0.853**	0.648**
	ct0min – catnof0	0.391**	0.559**	0.647**	0.062
	ct0max – cat0	0.695**	0.849**	0.876**	0.829**
	ct0max – catnof0	0.804**	0.851**	0.810**	0.846**
	cat0 – catnof0	0.814**	0.883**	0.941**	0.731**

Legend: * denotes statistical significance at $p \leq 0.05$, while ** denotes statistical significance at $p \leq 0.01$.

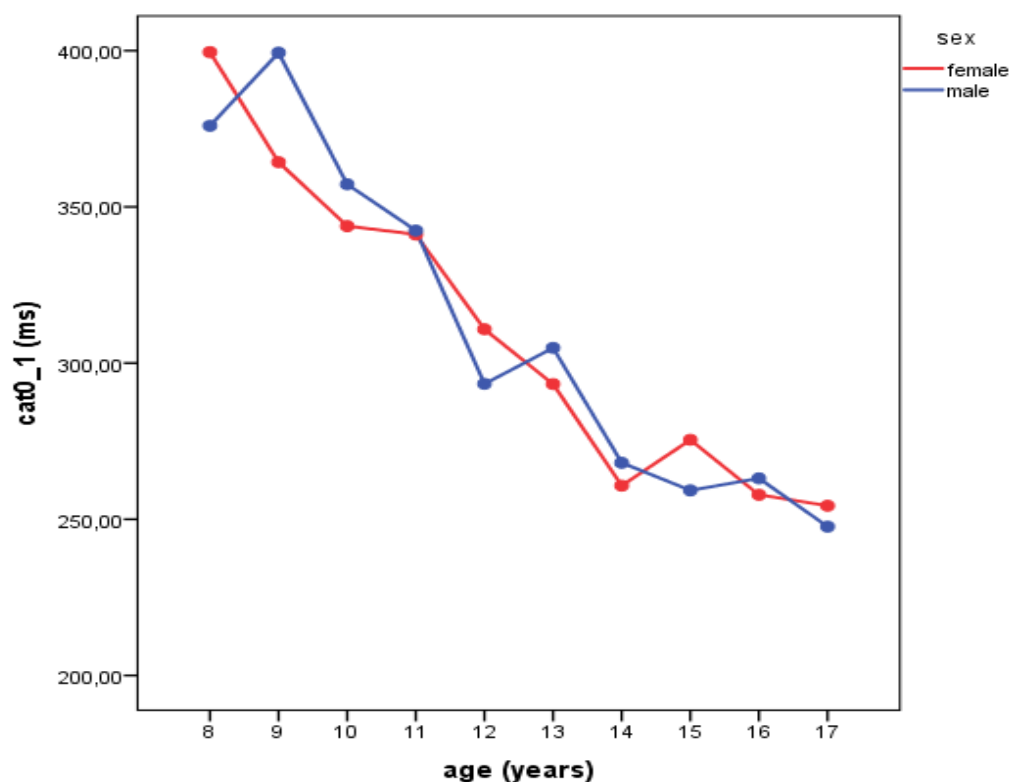


Figure 2a. Age-related differences of average cognitive response time in correct answered tasks of *Simple reaction time test* (cat0_1) for female and male students

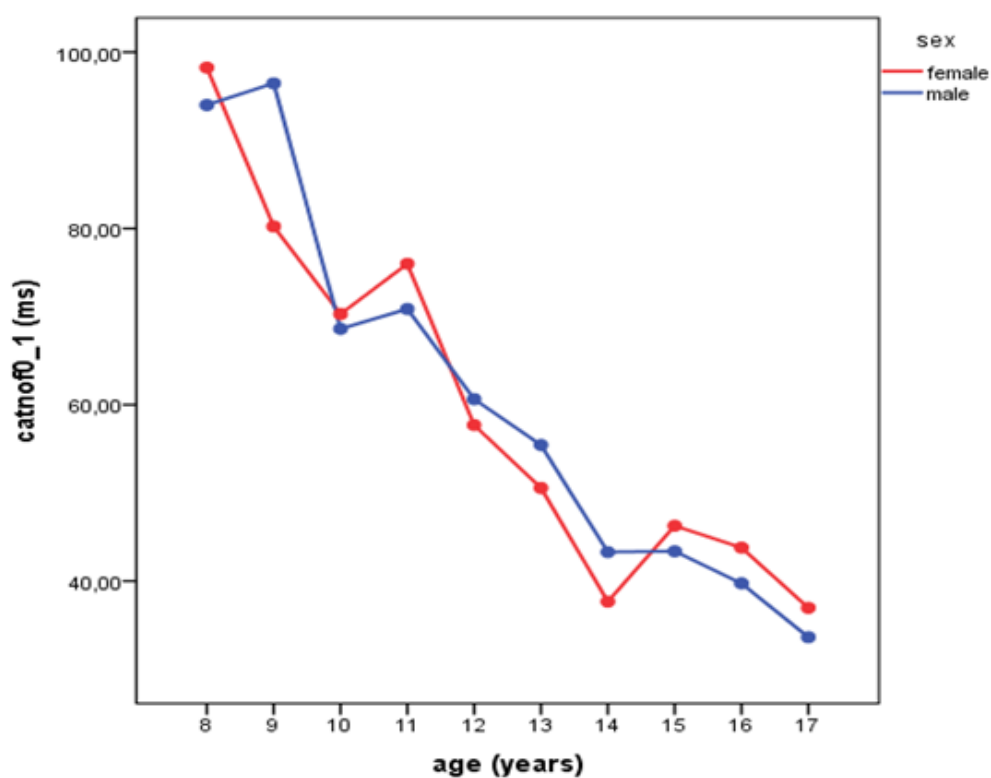


Figure 2b. Age-related differences of average non-optimized cognitive response time in correct answered tasks of *Simple reaction time test* (catnof0_1) for female and male students

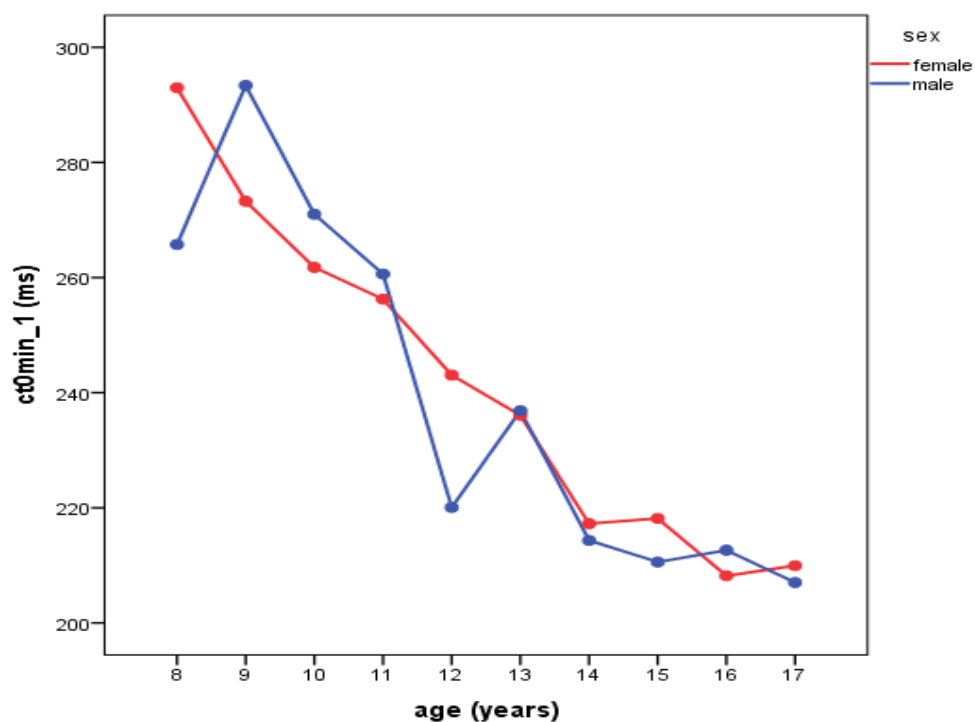


Figure 2c. Age-related differences of minimal cognitive response time in correct answered tasks of Simple reaction time test (ct0min_1) for female and male students

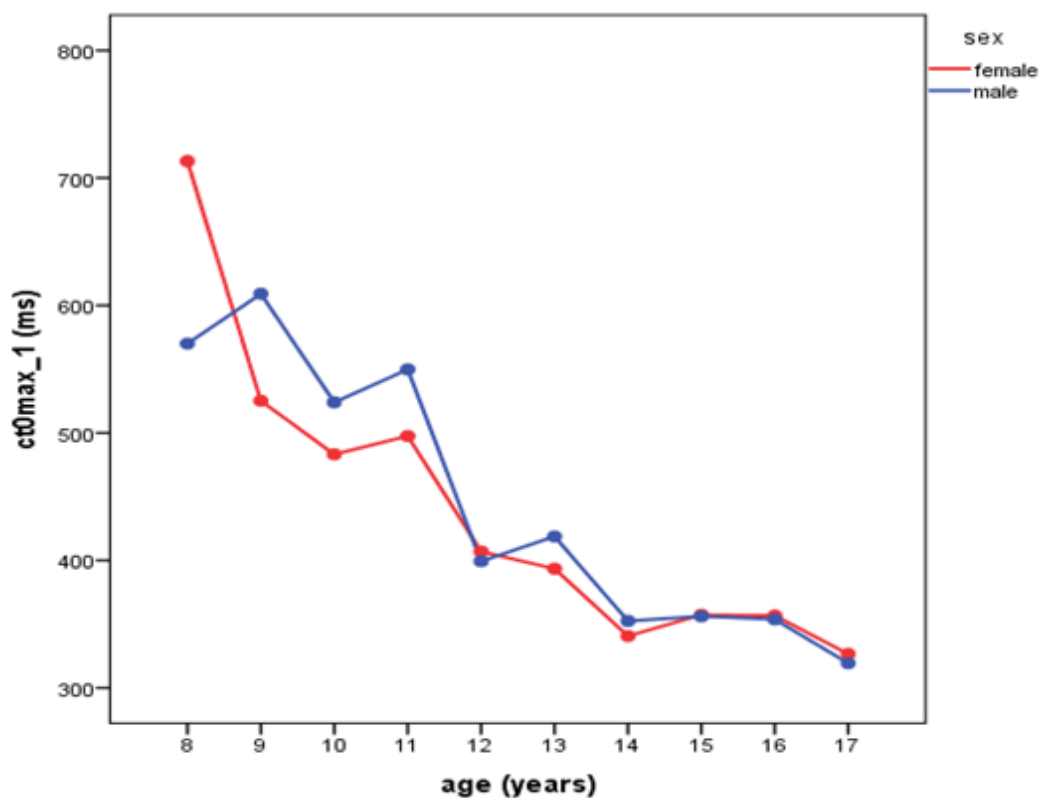


Figure 2d. Age-related differences of maximal cognitive response time in correct answered tasks of Simple reaction time test (ct0max_1) for female and male students

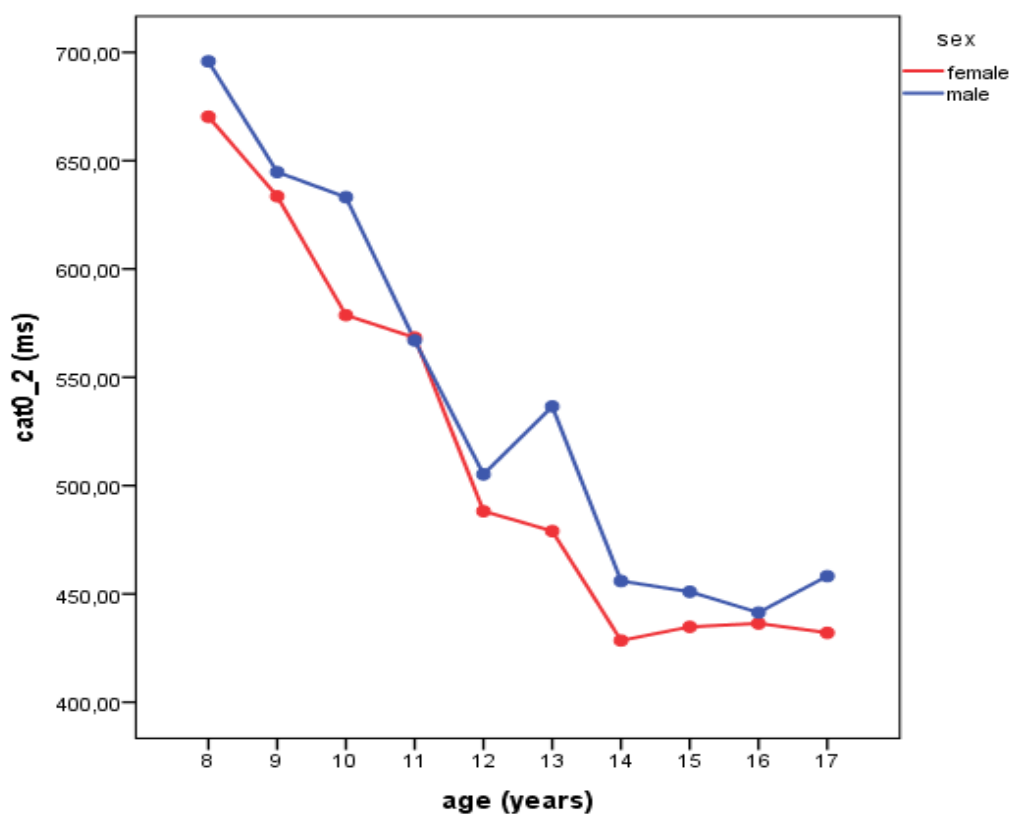


Figure 3a. Age-related differences of average cognitive response time in correct answered tasks of *Word recognition go-no go test* (cat0_2) for female and male students.

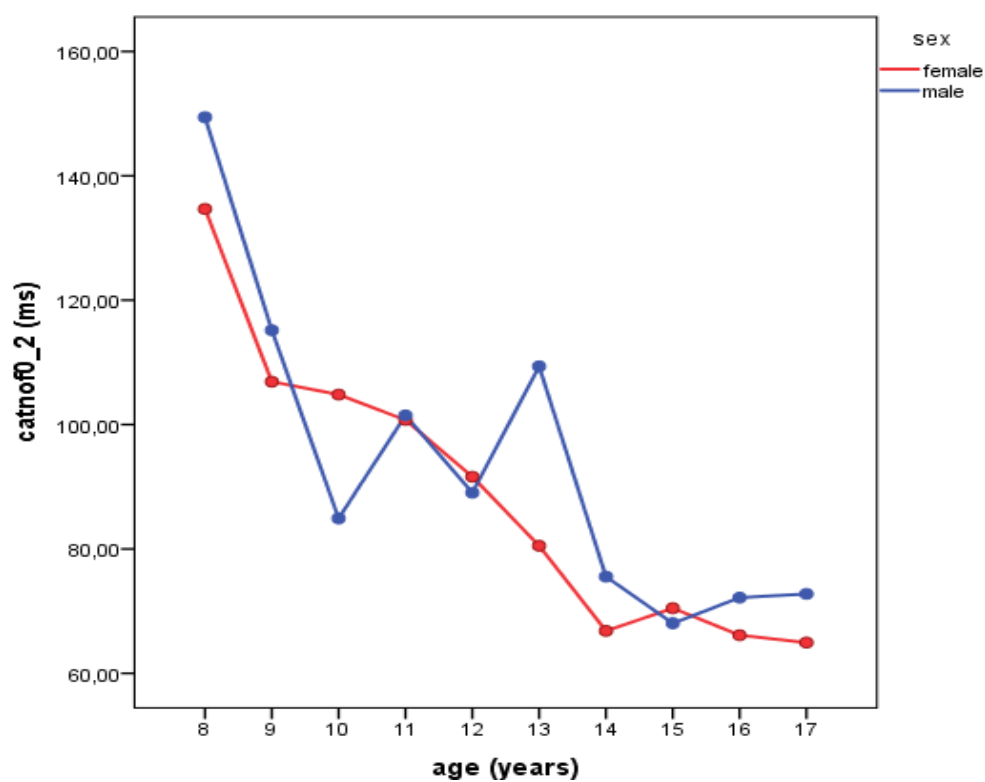


Figure 3b. Age-related differences of average non-optimized cognitive response time in correct answered tasks of *Word recognition go-no go test* (catnof0_2) for female and male students.

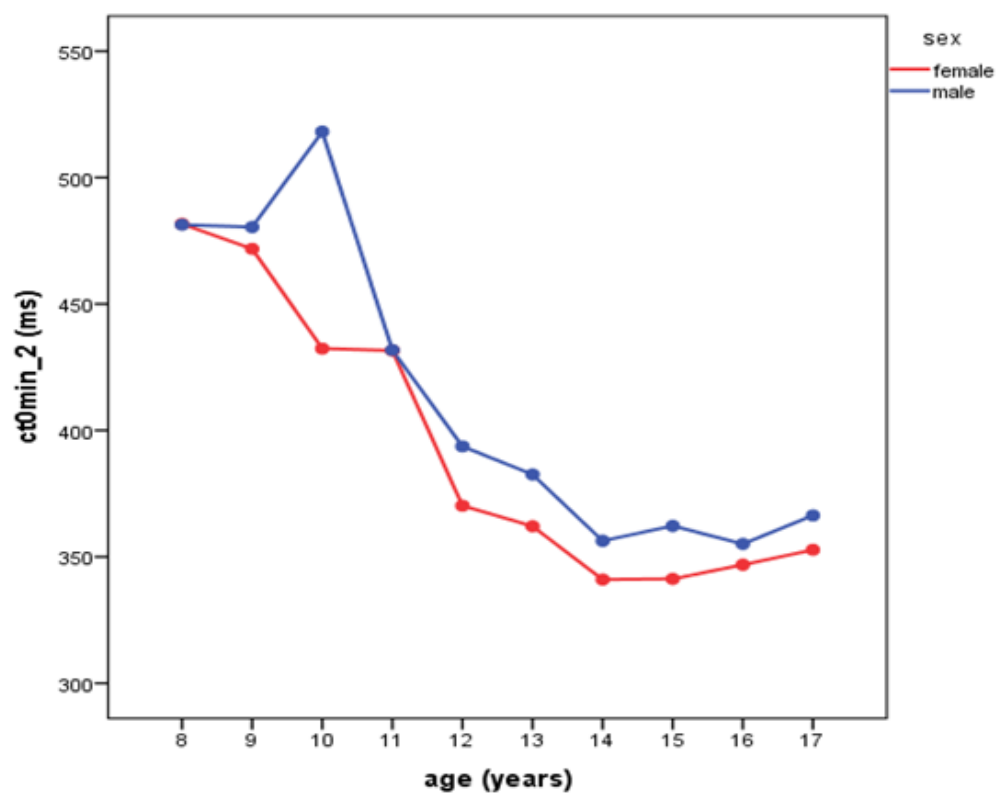


Figure 3c. Age-related differences of minimal cognitive response time in correct answered tasks of *Word recognition go-no go test* (ct0min_2) for female and male students.

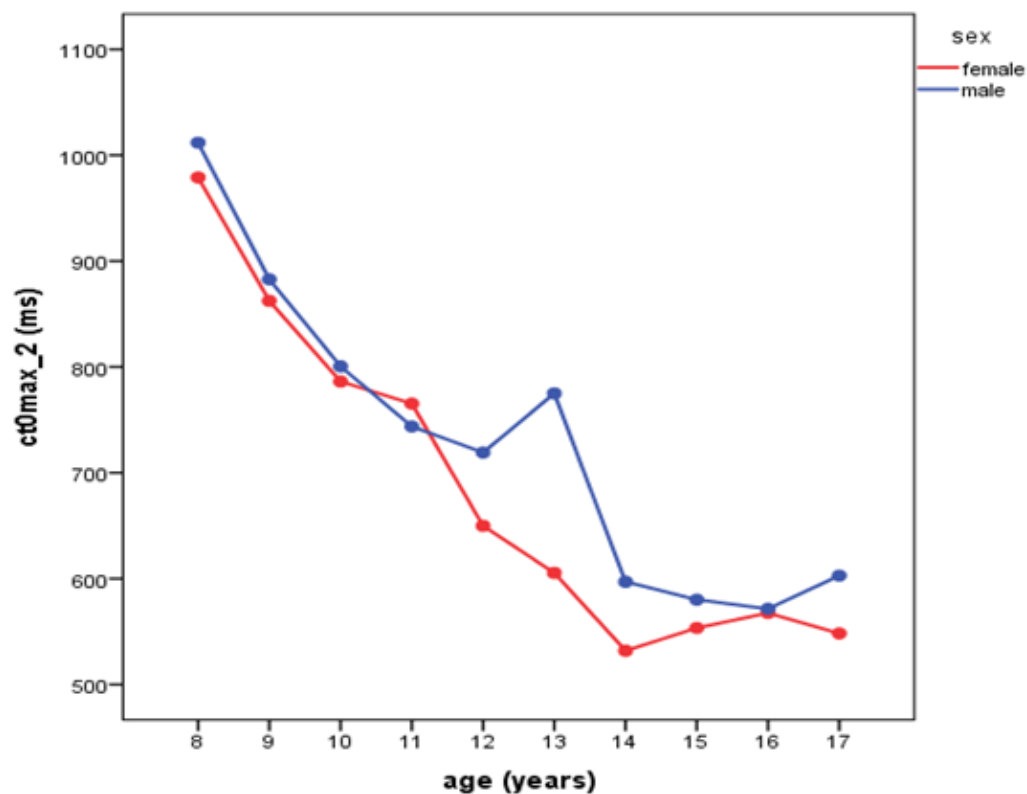


Figure 3d. Age-related differences of maximal cognitive response time in correct answered tasks of *Word recognition go-no go test* (ct0max_2) for female and male students.

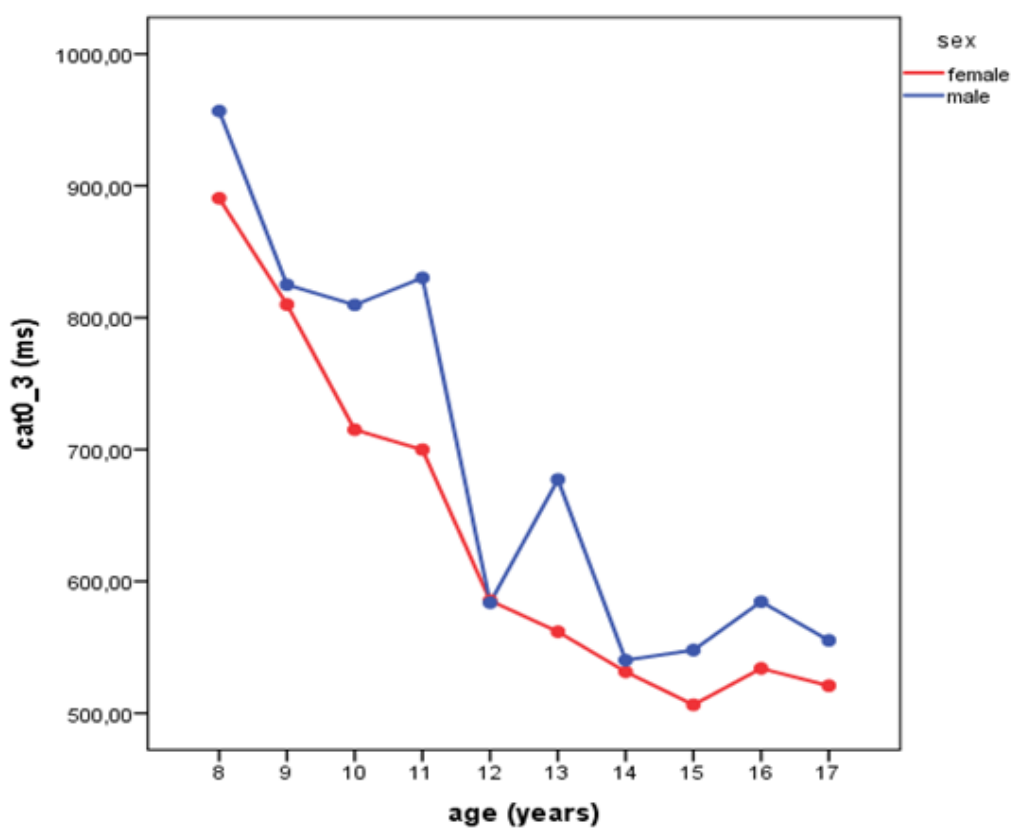


Figure 4a. Age-related differences of average cognitive response time in correct answered tasks of *Choice reaction time to the color name test* (cat0_3) for female and male students.

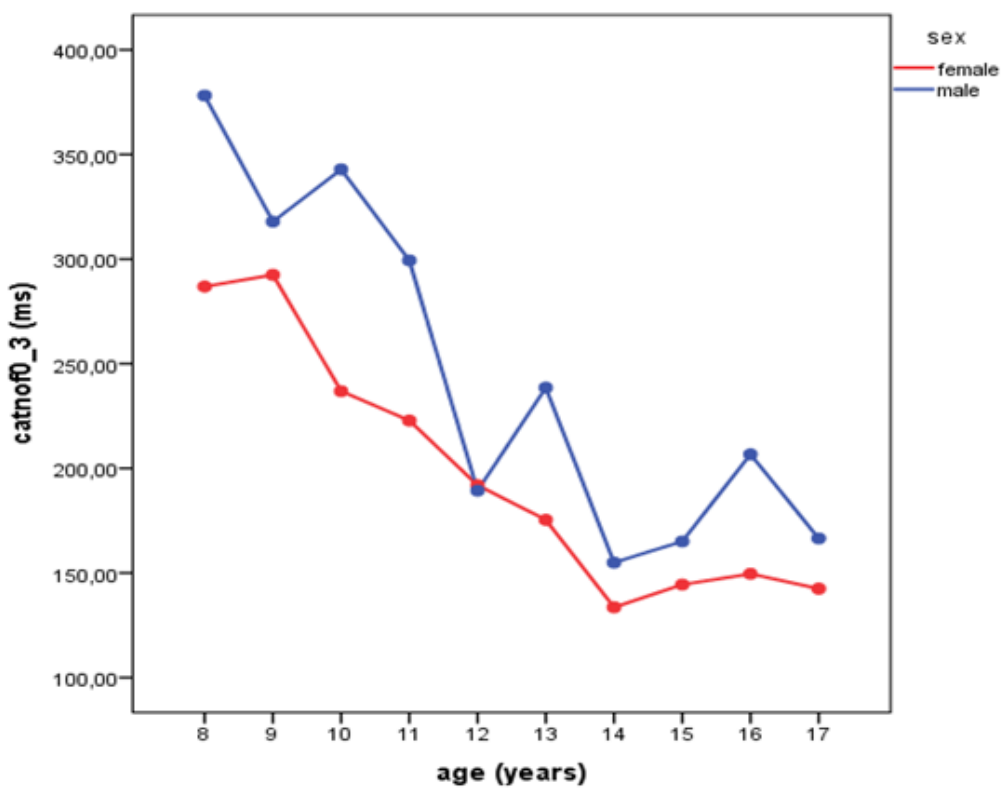


Figure 4b. Age-related differences of average non-optimized cognitive response time in correct answered tasks of *Choice reaction time to the color name test* (catno0_3) for female and male students.

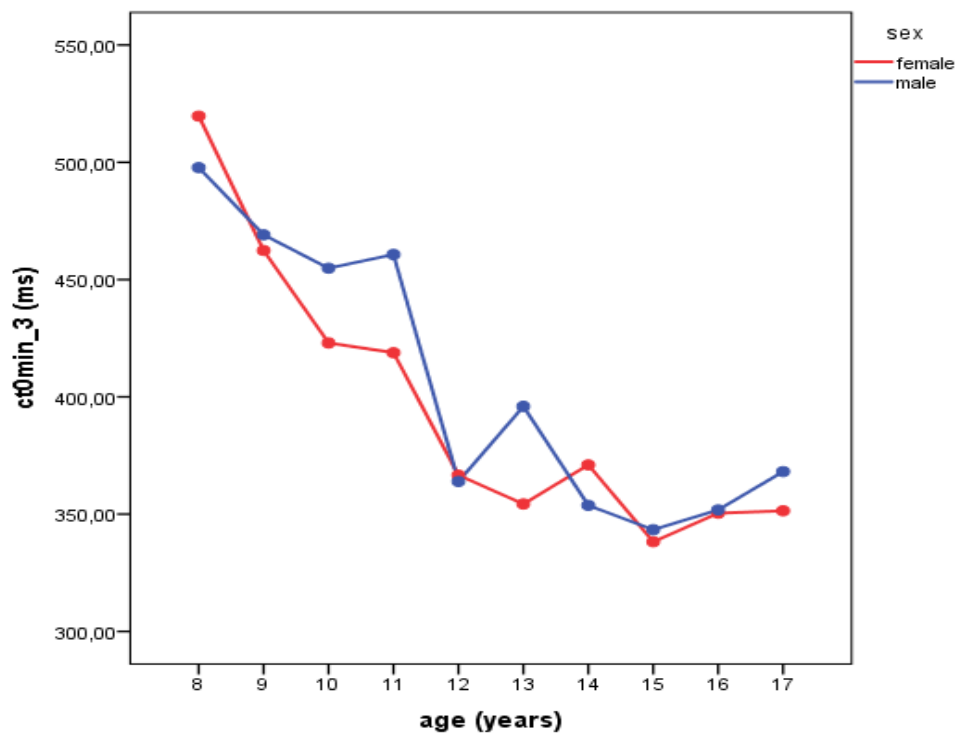


Figure 4c. Age-related differences of minimal cognitive response time in correct answered tasks of *Choice reaction time to the color name test* (ct0min_3) for female and male students

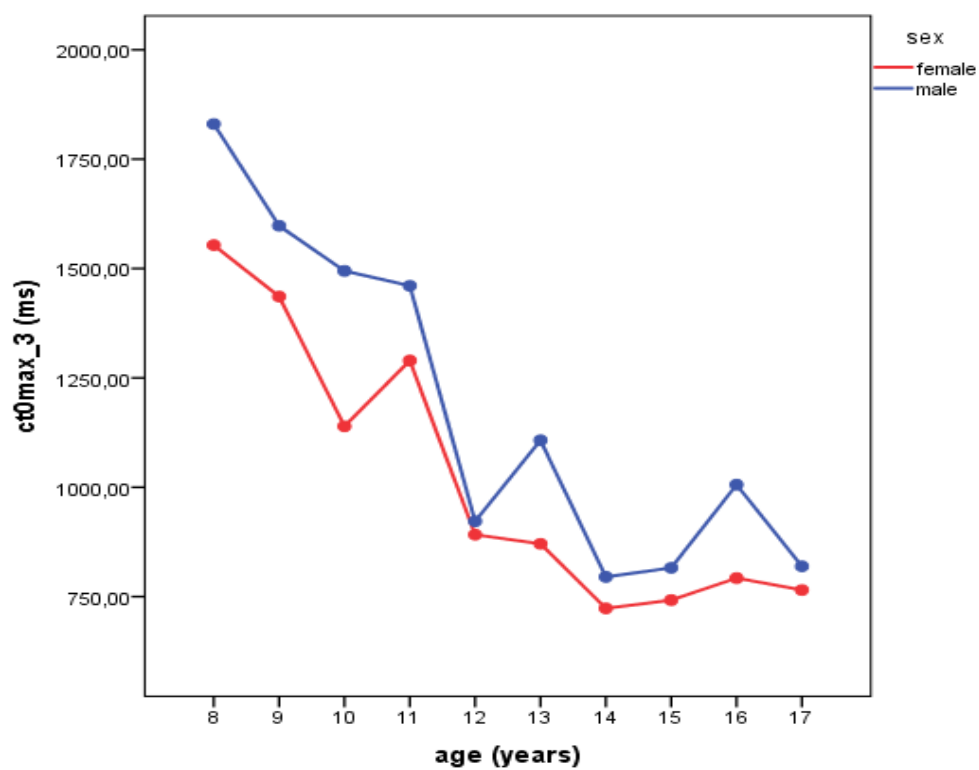


Figure 4d. Age-related differences of maximal cognitive response time in correct answered tasks of *Choice reaction time to the color name test* (ct0max_3) for female and male students.

By analyzing *Figures 2a – 2d*, *Figures 3a – 3d*, and *Figures 4a – 4d*, related to sex-specific age-related differences of PDIs in *SRT*, *WR* and *CRT* test, we derived several findings:

(1) Values of all four PDIs decrease in the observed developmental period; considering that smaller PDI value means better performance, all four PDIs improve in the observed period,

(2) The smallest relative decrement from age 8 to age 17 (i.e. absolute indicator's decrement divided by the average value of the indicator) is found at *ct0min* indicator,

(3) The largest relative decrement from age 8 to age 17 is found at *catnof0* indicator in *SRT* and *WR* tests and is found at *ct0max* indicator in *CRT* test,

(4) Sex differences in PDI's values in the observed developmental period increase from *SRT* task to *CRT* task and are present at ages of 9-10, 12-13 and 17 years for most of PDIs; More precisely:

- in *SRT* test the observed sex differences are not clearly expressed, although there is indication that girls show better performance until the age 12 or 13 years,

- in *WR* test girls perform better at age 10 (in *cat0* and *ct0min*) and at ages 12-14 (in all PDIs)

- in *CRT* test girls perform better at ages 10-11 and at ages 13-16, in the most of PDIs

(5) The age-related means variation across observed successive ages of most of the PDIs has been more pronounced in boys, but this tendency is more clear in *WR* and *CRT* tests, then in *SRT* test (where girls' means varied more for *cat0* and *catnof0*, while boys' means varied more for *ct0min* and *ct0max*),

(6) Age-related changes in PDIs are non-linear since their mean decrement from 8 to 12 years of age is several times larger than their decrement from 12

to 17 years (depending on the indicator); thereby the smallest nonlinearity has been found mostly at *catnof0* (in *WR* and *CRT* tests), while the largest one has been found at *ct0min*. Additionally, nonlinearity grows with test complexity. More precisely, in:

- *SRT* test (age related decrement)₈₋₁₂ is 1.5 – 2.8 times larger than (age related decrement)₁₂₋₁₇

- *WR* test (age related decrement)₈₋₁₂ is 2.5 – 4.8 times larger than (age related decrement)₁₂₋₁₇

- *CRT* test (age related decrement)₈₋₁₂ is 4 – 24 times larger than (age related decrement)₁₂₋₁₇

Research problem 3 analysis and related results

To statistically test specificities of four PDIs' developmental trends (related to developmental change intensity, to cognitive-motor task complexity, and to children's/youth's sex) revealed by previous descriptive statistical analysis, we conducted ANOVA statistical procedures. Thereby we analyzed these trends across cognitive Dev. Phases (not across developmental ages) to satisfy ANOVA prerequisites and theoretical guidelines of cognitive development.

(i) Possible specificities of various PDIs development across cognitive Dev. Phases

To test whether four PDIs (*best*, *worst* and *average performance* and *non-optimal performance instability*) change with different intensity across four cognitive Dev. Phases, we performed analysis that included relation to participant's sex and relation to task complexity (independent variables that possibly determine differential PDIs decrement across Dev. Phases). Therefore, we performed two models of three-way ANOVA: **(1)** (*Dev. Phase*) X (*PDI's type*) X (*participant's sex*) and **(2)** (*Dev. Phase*) X (*PDI's type*) X (*task complexity*). First ANOVA model is conducted separately for three variously complex cognitive-motor tasks

(SRT, WR, CRT), while second ANOVA model included task complexity as one of three independent variables of the model. The target outcomes of these four ANOVA calculations (three calculations of the 1st three-way ANOVA model,

and one calculation of the 2nd three-way ANOVA model) were interaction effects (*Dev Phase*) X (*PDI's type*) on PDI's values, and they are presented in *Table 3*.

Table 3. Outcomes of two models of three-way repeated measures ANOVA calculation related to interaction effect of *PDI's type* and *Dev. Phase* on PDIs' values (with first model conducted for all three cognitive-motor tasks - SRT, WR, CRT).

three-way ANOVA model	Test complexity	df ₁	df ₂	Mean Square	F	Sig.	Partial Eta Squared	Observed Power
Outcome of interaction effect of <i>PDI's type</i> and <i>developmental phase</i> on PDIs' values								
(1) <i>DevPhase</i> X <i>PDIType</i> X <i>Sex</i>	SRT	3.61	548.20	657010.109	58.051	<0.001	0.277	1.000
	WR	4.85	735.88	874601.700	87.062	<0.001	0.365	1.000
	CRT	3.36	508.07	7091459.432	37.901	<0.001	0.200	1.000
(2) <i>DevPhase</i> X <i>PDIType</i> X <i>Task complexity</i>	SRT							
	WR	3.48	531.29	3133686.701	23.567	<0.001	0.134	1.000
	CRT							

Note: interaction effects were calculated within two models of three-way repeated-measures ANOVA by using Greenhouse-Geisser correction of degrees of freedom (sphericity assumption violated).

Results of both models of three-way ANOVAs clearly show that four observed interaction effects (*PDI's type*) X (*Dev. Phase*) on PDI's values are statistically significant, which means that various PDIs values decrease across *Dev. Phases* with different intensities (or slopes). *Partial Eta Squared* values in ANOVA model (1) suggest that the strongest (*PDI's type*) X (*Dev. Phase*) interaction was observed at WR test and the smallest interaction is observed at CRT test. These differences in the observed interaction effects among variously complex tasks proved to be significant in ANOVA model (2) ($F = 16.810$, $df_1 = 4.03$, $df_2 = 614.71$, $p < 0.001$).

To elaborate differences in decrement intensity of various PDI's across four *Dev. Phases* we performed 12 *one-way ANOVAs*, with *Dev. Phase* being the only independent variable, for all PDIs in

three variously simple tasks (SRT, WR, CRT). *Partial Eta Squared* values (that give us information on the effect size that *Dev. Phase* exerts on PDI's values), presented in following *Table 3.1*, suggest that *Dev. Phase* produced the greatest decrement effect on *cat0* and the smallest decrement effect on *catnof0* (with *ct0max* being mostly on the third and *ct0min* mostly on second place). This finding is not aligned with some conclusions of the 2nd *Research problem analysis* (that stated the greatest *relative* developmental decrement for *catnof0*, and the smallest for *ct0min*) and this discrepancy stems from two sources. First, *one-way ANOVAs* are performed with *Dev. Phase* as independent variable, and *relative* decrement in 2nd *Research problem analysis* included *age* as independent variable. Second, *Partial Eta Squared* calculation is based on different elements than *relative* decrement calculation.

Table 3.1. One-way ANOVA statistical procedures outcomes conducted to assess PDIs' decrement (*ct0min*, *ct0max*, *cat0*, *catnof0*) across cognitive Dev. Phases in variously simple cognitive-motor tasks (SRT, WR, CRT).

Task's type	Performance dynamics indicator	df ₁	df ₂	Mean Square	F	Sig.	Partial Eta Squared	Observed Power
SRT	<i>ct0min_1</i>	3	459	117038.227	79.194	<0.001	0.341	1.00
	<i>ct0max_1</i>	3	459	1825302.006	75.545	<0.001	0.331	1.00
	<i>cat0_1</i>	3	459	382218.300	115.782	<0.001	0.431	1.00
	<i>catnof0_1</i>	3	459	67324.607	72.991	<0.001	0.323	1.00
WR	<i>ct0min_2</i>	3	459	430340.406	113.975	<0.001	0.427	1.00
	<i>ct0max_2</i>	3	459	3561253.32	146.699	<0.001	0.489	1.00
	<i>cat0_2</i>	3	459	1265585.14	209.876	<0.001	0.578	1.00
	<i>catnof0_2</i>	3	459	98751.700	82.950	<0.001	0.352	1.00
CRT	<i>ct0min_3</i>	3	458	547249.454	63.642	<0.001	0.294	1.00
	<i>ct0max_3</i>	3	458	18657973.3	49.842	<0.001	0.246	1.00
	<i>cat0_3</i>	3	458	3316738.35	73.468	<0.001	0.325	1.00
	<i>catnof0_3</i>	3	459	723555.193	41.694	<0.001	0.214	1.00

By considering that all PDIs are inverse measures of performance success (the longer decision times reflect worse performance of activated cognitive-motor system), the above mentioned ANOVA results point to developmental improvement of performance dynamics of the cognitive systems activated during solving three simple cognitive-motor tests (*SRT*, *WR*, *CRT*). The largest developmental improvement showed *average task performance* of the cognitive system, while the smallest improvement showed cognitive system *non-optimal functional instability*. These differences in developmental improvement among PDIs are the most expressive in *WR* cognitive-motor system and the least expressive in *CRT* cognitive-motor system.

At the end of this analysis of different intensities in four PDIs developmental improvement, it should be noted that differentiated improvement has the same form for girls and boys only in *CRT* task ($F=0.530$, $df_1=3.36$, $df_2=508.07$, $p>0.05$). In *SRT* task ($F=8.030$, $df_1=3.62$, $df_2=548.20$, $p<0.001$) and *WR* task ($F=2.875$, $df_1=4.85$, $df_2=735.88$, $p<0.05$) differentiated improvement of four PDIs did not have the same form for both sexes. Namely, in *SRT* task performance, previously proved differentiated PDIs' developmental decrement, at girls is continuous/smooth from the first developmental phase, while at boys this differentiated PDI's decrement starts to be continuous/smooth from the second developmental phase. On the other hand, in *WR* task performance, previously proved differentiated PDIs' developmental decrement, at

girls contains steeper developmental decrement of *ct0max* and *cat0* indicators, than at boys.

- (i) *Possible dependence of PDIs development on cognitive-motor task complexity*

To test whether developmental change of any PDI varies across cognitive-motor task complexity, we performed two-way repeated measures ANOVA for every PDI, with *Dev. phase* and *Task complexity* as independent variables and we focused to (*Dev. phase*) X (*Task complexity*) interaction effect.

Table 4. Outcome of two-way repeated measures ANOVA related to interactional effect of *task complexity* and *developmental phase* on PDIs' values, for all four PDIs (*ct0min*, *ct0max*, *cat0*, *catnof0*).

Indicator's type	df ₁	df ₂	Mean Square	F	Sig.	Partial Eta Squared	Observed Power
ct0min	4.90	748.73	61942.851	15.406	<0.001	0.092	1.000
ct0max	3.50	534.84	4383895.606	19.765	<0.001	0.115	1.000
cat0	3.51	543.46	633183.923	25.664	<0.001	0.144	1.000
catnof0	3.53	540.75	189161.491	18.670	<0.001	0.109	1.000

Note: interaction effects were calculated within a two-way repeated-measures ANOVA (*developmental phase* X *task's complexity*) by using Greenhouse-Geisser correction of degrees of freedom (sphericity assumption violated).

Table 4 results point to significant (*Dev. phase*) X (*Task complexity*) interactional effect on PDI's value at all PDI's types, which means that any observed PDI improves developmentally in a different way for differently complex cognitive-motor tasks. This finding is confirmed within two models of three-way ANOVA. (1) (*Dev. phase*) X (*Task complexity*) X (*Participant's sex*) ANOVA model showed that interaction effect of *Dev phase* and *Task complexity* on PDI's values is significant for all four PDIs ($F_{ct0min}= 15.01$, $p<0.001$; $F_{ct0max}= 18.69$, $p<0.001$; $F_{cat0}= 24.55$, $p<0.001$; $F_{catnof0}= 17.76$, $p<0.001$). (2) (*Dev. phase*) X (*PDI's type*) X (*Task complexity*) ANOVA model showed significance of the same interaction effect ($F= 23.57$, $df_1= 3.48$, $df_2= 531.29$, $p<0.001$).

The Table 3.1 results tell us how the observed PDIs decrease across Dev. Phases differently at SRT, WR and CRT cognitive-motor task. Namely, *Partial Eta Squared* values of all four PDIs vary across task

types with the same pattern: developmental change of all PDIs is the most present/obvious in WR task performance, and the least present/obvious in CRT task performance.

- (ii) *Possible dependence of cognitive PDIs' developmental change on participant's sex*

Since the conclusions of the 2nd Research problem included possible sex differences in PDIs' values and in PDIs' development on the descriptive level, it is expected to statistically test (on inferential level) these differences in cognitive performance dynamics. To do that, we first changed the analytic scale from *ages* to *Dev. Phases*, and then calculated (*Dev. phase*) X (*Participant's sex*) interactions within three-way ANOVA model (*Dev. phase*) X (*Participant's sex*) X (*Task complexity*) for every PDI (*ct0min*, *ct0max*, *cat0*, *catnof0*). After that, we elaborated the obtained results within additional

two-way ANOVA calculations of the same (*Dev. phase*) X (*Participant's sex*) interaction effects for every PDI, in every cognitive-motor task (*SRT*, *WR*, *CRT*).

Table 5. Outcomes of (*Dev. phase*) X (*Participant's sex*) X (*Task complexity*) ANOVA for all PDIs (*ct0min*, *ct0max*, *cat0*, *catnof0*) related to (*Dev. phase*) X (*Participant's sex*) interaction.

PDI type	df ₁	df ₂	Mean Square	F	Sig.	Partial Eta Squared	Observed Power
ct0min	3	454	15655.837	2.172	0.091	0.014	0.552
ct0max	3	454	75988.207	0.473	0.701	0.003	0.146
cat0	3	454	10175.976	0.411	0.745	0.003	0.132
catnof0	3	455	7289.194	1.006	0.390	0.007	0.274

Note: In calculating two-way interaction effect (*Dev. phase*) X (*Participant's sex*) on PDI-dependent variable, this variable is formed as the average of PDI_{SRT}, PDI_{WR} and PDI_{CRT} values.

Table 5 content states that (*Dev. phase*) X (*Participant's sex*) interaction is not significant at any PDI, which means that all PDIs improve across developmental phases in the same way for girls and boys. Since in this ANOVA model PDI-dependent variable (in the observed two-way interaction) is the average of three variously complex components (PDI_{SRT}, PDI_{WR} and PDI_{CRT} values), the obtained interaction effects do not say anything about the observed interaction effect on the level of particular task (*SRT*, *WR*, *CRT*). Therefore, we calculated 12 two-way ANOVAs (for every PDI, in every task) and presented the related results in Table 5.1 in condensed form.

Table 5.1. Outcomes of two-way ANOVA calculation of interactions (*Dev. phase*) X (*Participant's sex*) for every PDI, in every cognitive-motor task (*SRT*, *WR*, *CRT*).

		Task type		
		<i>SRT</i>	<i>WR</i>	<i>CRT</i>
PDI	<i>ct0min</i>	F= 4.79 df ₁ =3 ; df ₂ =455 p<0.01	F= 0.46 df ₁ =3 ; df ₂ =455 p>0.05	F= 1.38 df ₁ =3 ; df ₂ =454 p>0.05
	<i>ct0max</i>	F= 8.36 df ₁ =3 ; df ₂ =455 p<0.001	F= 3.32 df ₁ =3 ; df ₂ =455 p<0.05	F= 0.43 df ₁ =3 ; df ₂ =454 p>0.05
	<i>cat0</i>	F= 2.59 df ₁ =3 ; df ₂ =455 p=0.052	F= 0.77 df ₁ =3 ; df ₂ =455 p>0.05	F= 0.26 df ₁ =3 ; df ₂ =454 p>0.05
	<i>catnof0</i>	F= 0.82 df ₁ =3 ; df ₂ =455 p>0.05	F= 1.26 df ₁ =3 ; df ₂ =455 p>0.05	F= 1.06 df ₁ =3 ; df ₂ =455 p>0.05

The findings of Table 5.1 show that developmental improvement of PDIs still differs for girls and boys at *ct0min* and *ct0max* indicators during *SRT* task performance and at *ct0max* indicator during *WR* task performance. Conducting independent t-test at every *Dev. Phase*, to test the possible sex

differences in $ct0min$ and $ct0max$ at SRT and WR task, gave next findings. Sex differences were noted at first developmental phase in $ct0min_{SRT}$ and $ct0max_{SRT}$ ($t= 2.714$, $df= 78$, $p= 0.008$ and $t= 2.216$, $df= 47.364$, $p= 0.031$, respectively) and at second developmental phase only in $ct0max_{SRT}$ ($t= -2.500$, $df=137$, $p= 0.014$). Thereby, boys outperformed girls in $ct0min_{SRT}$ and $ct0max_{SRT}$ at phase 1, and girls were superior in $ct0max_{SRT}$ at phase 2. Furthermore, testing sex differences in $ct0max$ of WR task (at every developmental phase), showed that girls' and boys' worst performance was the same during the first two Dev. Phases. However, girls outperformed boys during the last two observed phases ($t_{phase1}= -0.662$, $df=77.97$, $p=0.510$; $t_{phase2}= -0.193$, $df=137$, $p=0.847$; $t_{phase3}= -5.419$, $df=76.84$, $p<0.001$; $t_{phase4}= -2.676$, $df=149.95$, $p<0.01$).

Research problem 4 analysis and related results

Although the analysis of the 3rd Research problem proved that, there are sex differences in some PDIs' values at some developmental phases, the systematic analysis of sex differences in performance dynamics still have to be done. To do that, we conducted three-way repeated measures ANOVA (*participant's sex*) X (*PDI's type*) X (*task complexity*) and focused to sex differences on the observed developmental period as a whole, and their possible dependence on PDI's type and on task complexity. To elaborate results of this analysis, we conducted additional t-tests with *participant's sex* as independent variables.

The outcomes of the above mentioned three-way ANOVA, relevant for sex differences analysis, are given in *Table 6*.

Table 6. Outcomes of three-way repeated measures ANOVA (*participant's sex*) X (*PDI's type*) X (*task complexity*), relevant for sex differences analysis of performance dynamics in three variously simple cognitive-motor tasks.

Source of variation	df1	df2	Mean Square	F	p	Partial Eta Squared	Observed Power
participant's sex	1	460	2313856.18	10.632	0.001	0.023	0.902
(participant's sex) X (PDI type)	1.13	519.33	1191061.84	9.671	0.001	0.021	0.901
(participant's sex) X (task complexity)	1.15	531.01	1559605.26	10.384	0.001	0.022	0.924

Note: Testing prerequisite assumptions for repeated measures ANOVA showed that sphericity assumption has been violated, so we applied related Greenhouse-Geisser correction.

(i) Sex differences in overall performance dynamics

Main effect of participant's sex on PDIs' values, from *Table 6*, is significant and points to sex differences in performance dynamics while solving three variously simple cognitive-motor tasks. This effect is present in overall observed developmental period, but its size is small (Partial Eta Squared value shows that only 2.3% of total PDIs' variability is explained by sex differences).

To describe direction and structure of these sex differences in overall performance dynamics during completely observed developmental period, according to dimensions of PDI type and task complexity, we conducted 12 independent sample t-test (4 PDIs X 3 tasks). The results of these t-tests were presented in *Table 6.1* in the condensed form.

Table 6.1. Means (M) and standard deviations (SD) of PDIs' values in girls and boys, calculated on overall developmental period and distributed according PDI's type and task complexity.

		Task complexity					
		SRT		WR		CRT	
PDI	ct0min	M _f =246.1	SD _f =45.75	M _f =401.5	SD _f =77.69	M _f =403.5	SD _f =101.68
		M _m =244.9	SD _m =48.67	M _m =419.1	SD _m =83.14	M _m =420.4	SD _m =117.13
		M_f-M_m = 1.2		M_f-M_m = -17.6*		M_f-M_m = -16.9	
	ct0max	M _f =460.9	SD _f =204.29	M _f =709.9	SD _f =208.59	M _f =1071.4	SD _f =613.54
		M _m =469.6	SD _m =173.99	M _m =769.3	SD _m =222.08	M _m =1281.7	SD _m =765.83
		M_f-M_m = -8.7		M_f-M_m = -59.4**		M_f-M_m = -210.3**	
	cat0	M _f =318.2	SD _f =72.76	M _f =529.2	SD _f =112.15	M _f =655.7	SD _f =234.98
		M _m =322.6	SD _m =78.94	M _m =559.9	SD _m =124.04	M _m =731.2	SD _m =273.39
		M_f-M_m = -4.40		M_f-M_m = -30.7**		M_f-M_m = -75.5**	
	catnof0	M _f =184.7	SD _f =149.88	M _f =92.4	SD _f =37.55	M _f =207.0	SD _f =123.39
		M _m =190.7	SD _m =168.99	M _m =102.9	SD _m =46.66	M _m =263.6	SD _m =164.20
		M_f-M_m = -6		M_f-M_m = -10.5**		M_f-M_m = -56.6**	

Notes: The difference in PDIs' mean values between girls and boys (M_f-M_m) has been tested with independent samples' t-test; * = p<0.05, ** = p<0.01.

Table 6.1 findings show that girls' PDIs' values are smaller than the ones from boys in 91% of the cases (differences M_f-M_m are negative), which means that girls' performance dynamics at the observed developmental period is superior to the boys' performance dynamics – although this difference is small.

(ii) Sex differences' dependence on PDI's type

Since interaction effect (*Participant's sex*) X (*PDI type*), presented in the Table 6, is statistically significant, it can be concluded that small girls' superiority in overall performance dynamics depends on PDI's type. To interpret this dependence, the magnitudes of statistically significant differences M_f-M_m of Table 6.1 are used. Namely, from these magnitudes it is visible that the biggest girls' advantage appears at ct0max PDI, while the smallest girls' advantage appears at ct0min PDI.

(iii) Sex differences' dependence on task complexity

Finally, to infer if small, but significant girls' superiority in overall performance dynamics (and at overall observed developmental range) depends on task complexity, we refer to (*Participant's sex*) X (*Task complexity*) interaction effect. Findings on this effect from Table 6 tell us that the observed sex differences significantly depend on cognitive-motor task complexity. The form of this dependence we derive from Table 6.1 results: since differences M_f-M_m predominantly continuously increase from SRT-task to CRT-task, we conclude that small girls' superiority increases with cognitive-motor task complexity, in the observed developmental period.

Discussion

Concerning the 1st Research problem, there are several arguments that suggest that four PDIs of the cognitive component of simple cognitive-motor task solving separately contribute to description of related cognitive sub-system functioning. First, mutual PDIs' correlations, although positive and significant in the most

correlation pairs (across all three tasks and all four Dev. Phases), in average do not explain more than 36% of common variance (median value of all correlations suggests 45% of common variance). That means that, with exception of *cat0* correlations with other PDIs (whose average value of common variance is 52.6%), the other three PDIs mostly didn't share mutually more than 50% of the variance and therefore, mostly presented *specific and separate performance dynamics descriptors*. Due to the reliability of the measurement (which ranges from 0.72 to 0.79%), the question arises whether the percentage of shared variance among PDI constructs is greater than 36%, but there are also developmental reasons why this increase is not realistic to expect. Specifically, the correlations in *Table 2* were calculated within each developmental phase, and from the graphical representations of the 2nd *Research problem*, it is clear that within these phases all PDIs change (decrease) with age. Therefore, the *age variable* also acts in the background of all correlations in *Table 2*, due to which the correlations are partly increased. Therefore, the sum of the antagonistic effects of *measurement reliability* and *age* on the correlations of *Table 2* suggests that their estimate of the relationship among the four PDIs stays valid.

Average performance efficiency (i.e. *cat0*) logically showed higher correlations with other PDIs, while it incorporates other PDIs by its definition (i.e. calculation formula). Furthermore, significant positive mutual PDIs' correlations at almost all indicators (across the three variously simple tests and across four Dev. Phases) is also logical since they are calculated from the same individual's RT distribution. However, these four PDIs do not use the same part of individual's RT distribution in their calculation (the exception is *cat0* that uses the whole RT distribution) and therefore mutual correlations among PDIs are of confined magnitude. That suggest that these PDIs measure

mostly different constructs of cognitive process activated during task responding. Additionally, *best performance (ct0min)* and *average non-optimal performance instability (catnof0)* calculation formula explains the lowest correlations between these two indicators: instability is average deviation of all RTs in individual's RT distribution from the best performance (and *X* cannot be highly correlated with *deviation from X*).

The last finding of the 1st problem analysis shows that PDIs are somewhat largely correlated in first two Dev. Phases (average $r = 0.615$) than in the last two (average $r = 0.584$), which can be explained by more intensive age-related change of all indicators in the first two Dev. Phases (see Figures 2a – 4d). Namely, within these two phases *age*, as a third variable accompanied to all PDIs, increases their correlations more than in the last two Dev. Phases.

All these 1st problem findings suggest that *best*, *worst* and *average non-optimal instability* indicators reflect specific features of someone's cognitive (sub)system activated during repetitive cognitive-motor task solving (except the average performance indicator, which is cumulative indicator that incorporates the rest of three). This conclusion finds support in WPR studies that explain the nature of worst RT in terms of WM and attention mechanisms lapses and in terms of individual differences in neural oscillations (Coyle, 2003; Schmiedek et al, 2007; Unsworth et al, 2010). The evidence also come from neurocognitive studies that explain person's intra-individual variability/instability while performing equivalent RT tasks by neuronal activity source (Booth et al, 2019; Paraskevopoulou et al, 2021; Reed, 1998). Moreover, some researchers proposed cognitive-motor abilities typology based on *best RT performance* (dimension of speed) and *average non-optimal instability performance* (dimension of stability), besides some other PDIs (Drenovac, 2009). Further evidence for specific contribution of

best, worst and average non-optimal instability indicators to performance dynamics description, comes from the similar research, but with different methodology (Žebec et al, 2014). This research also found that mutual correlations between *best, worst, average performance* and *performance instability* are medium in average, which explains less than 50% of the common factors. Moreover, the same research shows different effects of age and sex on different PDIs, which suggests that various PDIs are variously determined by biological variables. Finally, various effects of Dev. Phase and participant's sex on various PDIs are also proven in this research, within 3rd and 4th Problem's analysis.

Concerning the 2nd Research problem, which deals with age-related change of four PDIs during 8-17 years of age, three findings deserve attention.

First, in all three cognitive-motor tasks all four PDIs improved across the observed ten years long developmental period, and this improvement was non-linear: in the period 8 to 12 years it was several times larger than in the period 12-17 years of age. This finding is expected and in accordance with well-known developmental studies of processing rate development (Cerella & Hale, 1994; Kail & Salthouse, 1994), and inhibition mechanisms development – related primarily to WR-task (Band et al, 2000; Fortenbaugh et al, 2015; Ridderinkhof & van der Molen, 1997). On the other hand, this finding suggest common developmental mechanism for all four PDIs, mostly related to neurological maturation in central part (myelination, synaptic pruning or tuning) and peripheral part (myelination) of nervous system (Hale, Fry & Jessie, 1993; Luna et al, 2004; Travis, 1998, Petanjek et al, 2023), although activated in different developmental periods for different parts of central nervous system.

Second, the intensity of PDIs' relative age-related improvement is different for different PDIs. At one

hand, that means that biological variable of age affects differently the combination of neurological resources activated while producing the best or the worst response, or neural bases responsible for performance stability (Booth et al, 2007; Paraskevopoulou et al, 2021; Reed, 1998). On the other hand, age might also affect the strategies of repetitive fast responding via experience with that kind of tasks (which cumulates during childhood and adolescence), and strategies variously affect different PDIs. Namely, there is more "space" for strategy effect in non-optimal instability improvement (produced by responding in all tasks and caused also by distractors and personal variables activated during distractors' inhibition), than in best performance improvement (which appears in only one task, when distractors are absent or minimally present and related personal variables are not needed to activate). The question is whether these differences in relative intensity of age-related improvement (based on two extreme values of PDIs - at the age of 8 years, and at the age of 17 years) are more relevant than ANOVA outcomes of developmental change intensity (which consider not only variability across ages but also within the ages). This will be discussed at 3rd Research problem.

The third important finding at 2nd Research problem analysis reveals possible differences between girls and boys in PDI development. Figures 2a – 2d until 4a - 4d show that red developmental line (girls) at more ages lies below the blue line (boys), which means that in all four PDIs girls tend to attain shorter decision times (DTs) while solving the task and they less deviate from the best DT than boys. This tendency becomes more obvious as cognitive-motor task becomes more complex. Since this tendency is on descriptive level, we will not comment it more, unless we discuss ANOVA tests within the 4th

Research problem, but on the other analytic scale (Dev. Phases).

One more aspect of possible sex-specific age-related changes of four PDIs provokes curiosity. Namely, the average of PDI's mean changes between all successive ages show that boys' means oscillate more than the girls' means. It might be the consequence of sex differences in sample size at particular age, but quick look at the *Table 1* (age-sex structure of the study's sample) reveals that differences in boys and girls age-related samples' size are not significant. Furthermore, difference in the size of age-related oscillations of PDI's means might be the consequence of sex differences in age-related group variability, so we checked these differences (Levene's test). The conducted variance comparisons revealed that sex differences of results' variances in age-related groups, mostly cannot explain greater boys' oscillations in PDIs' developmental improvement (variances were the same at 82.5 – 95 % comparisons). When we consider that these developmental improvement oscillations in boys appeared at all four PDIs, in all three cognitive-motor tasks, without finding right methodological explanation, we suppose that some developmental mechanism should be responsible for that. The only comparable research available to us shows similar outcome only at analogous performance instability indicator (Žebec et al, 2014).

The analysis of the study results under the scope of the 3rd Research problem was to show whether the developmental improvement in cognitive-motor tasks' performance dynamics is determined by the PDI's type (*best, worst, average performance, average non-optimal instability*), by the complexity of the cognitive-motor task (*SRT, WR, CRT*) and by the sex of the task performer (*girls, boys*).

First, two repeated-measures three-way ANOVA models ([*Dev. Phase*] X [*PDI Type*] X [Participant's

sex] and [*Dev. Phase*] X [*PDI Type*] X [*Task complexity*]) showed that the developmental improvement was of different intensity for different PDIs. Specifically, across *Dev. Phases*, the average task performance (*cat0*) improves the most, and the average non-optimal instability (*catnof0*) improves the least. Such a finding is not surprising because the *cat0* includes the remaining three PDIs (*ct0min, ct0max* and *catnof0*), and given that these three PDIs improve through *Dev. Phases*, then in the development of the average performance those three improvements are cumulated. Similarly, the weakest developmental improvement of the *catnof0* is also logical. Namely, this instability indicator is largely determined by non-systematic variations in the individual's attention and the non-systematic appearance of distractors, and any independent variable - including the *Dev. Phase* - hardly affects non-systematic variations. Another explanation of the lowest *catnof0* developmental decrement is psychometric. Non-optimal instability (*catnof0*) is calculated as the sum of the differences between two significantly correlated quantities ($DT_i - DT_{min}$, $i = 1, 2, 3, \dots, 30$ or 32 number of task items; DT = decision time), because they are part of the same person's response time to equivalent cognitive-motor tasks. Since the difference of two correlated variables presents composite variable with reduced reliability, systematic effect of any other variable (e.g. *Dev. Phase*) on low-reliable (composite) variable, *a priori* cannot be large.

Various intensive developmental improvement of different PDIs with an analogous outcome was observed in the study by Žebec et al. (2014), who used a different reaction meter: the most intensive developmental improvement was noted in *average task performance*, and the least intensive improvement was noted in *performance instability*. Considering that the same finding has been obtained with two different instruments,

with differently sized samples, and on a different developmental scale, it can be reliably concluded that the differentiated developmental improvement of different PDIs is quite firmly proven. Additionally, the study of Dykiert et al. (2012) also observed developmental changes in *average task performance* and in *task performance instability*. With somewhat different data analysis they came to findings that the developmental changes of these two indicators are best fitted by *different* mathematical functions (quadratic and cubic), and that the absolute developmental change is more pronounced in the case of *average performance* than in *performance instability*.

The additional finding obtained under the scope of 3rd Research problem (in our research), which says that the differentiated improvement of different PDIs is more pronounced in the WR-task than in the SRT and CRT tasks, is difficult to comment on because this effect was not found in the literature available to us. We faced the similar situation with the finding that the pattern of differential developmental improvement for different PDIs is different for girls and boys in the SRT and WR cognitive-motor task, but not in the CRT task.

The next specificity of PDIs' developmental changes (found under the scope of the 3rd Research problem) is their non-equally intensity in three variously complex tasks (*SRT*, *WR*, *CRT*). Namely, ANOVA showed that each PDI manifests the most pronounced developmental change in the WR task, and the least pronounced change in the CRT task. Such consistency of the task complexity effect on different parts of the intra-individual DT distribution (related to a series of equivalent tasks) indicates that task complexity operates through a certain general mechanism.

An explanation based on a general mechanism involves two inferential premises. The first premise indicates that DT in *elementary cognitive tasks*

(ECTs, which contain one to two information-processing stages) decreases nonlinearly with the *ECT-rate* across Dev. Phases. The second premise indicates that total DT in solving more *complex cognitive tasks* (CCTs, which contain more than two information-processing stages) is equal to the sum of DT_is in all *k* information processing stages (*i*= 1, 2, 3, ..., *k*). The conclusion of the two premises: the sum of DT_is of all CCT-solving phases will be smaller at an older Dev. Phase than at an earlier Dev. Phase. As a result, differences (DT_{CCT} - DT_{ECT}) in earlier Dev. Phases will be nonlinearly greater than differences (DT_{CCT} - DT_{ECT}) in later Dev. Phases. In other words, DT_{CCT} will decrease across Dev. Phases non-linearly with *CCT-Rate* being greater than *ECT-Rate*.

The hypothesis on general mechanism of the information processing rate development in the human cognitive system was considered by several relevant authors in the field (Cerella & Hale, 1994; Kail & Salthouse, 1994; Miller & Vernon, 1997), although their explanation of this mechanism was partially differently. That hypothesis had its opponents (Cowan et al, 1998; Ridderinkhof and van der Molen, 1997; Madden, Pierce, & Allen, 1993) and apparently, its applicability to development of different measures of information processing rate is not comprehensive. However, it is possible to find childhood research studies that confirmed different rates of RT developmental change in solving variously complex cognitive tasks (Dykiert et al, 2012; Kiselev, Espy, Sheffield, 2009). Additionally, there are more works that recorded the same phenomenon in the aging period - of course, with the opposite direction of change (Salvia et al, 2016; Gorus, De Raedt and Mets, 2006; Tun & Lachman, 2008). Nevertheless, it looks like general mechanism explanation does not fit to our data. Namely, 3rd Research problem's ANOVA showed that the *Partial Eta Squared* values are smaller for the

developmental changes of all 4 PDIs in the CRT task than in the SRT task. On the other hand, it is clear that the CRT task is more complex than the SRT task. Therefore, the obtained result of our research probably reflects not only the above-mentioned general mechanism, but also a significantly greater inter-individual variability of the CRT task, than of the SRT task, within Dev. Phases (and this variability reduces the *Partial Eta Squared value*).

The last type of PDIs' developmental changes considered under the scope of *3rd Research problem* was possible sex determination of the developmental improvement in PDIs across the four observed Dev. Phases. The results of *Table 5* and *Table 5.1* showed layered picture of participant's sex effect on PDIs' developmental improvement. When we consider this effect regardless to task complexity (*Table 5*), then the four observed PDIs' developmental changes follow the same pattern for girls and boys (regardless of whether there is a difference between girls and boys at the level of the entire sample). However, when we analyze this sex effect by considering task complexity - i.e. separately for SRT, WR and CRT task (*Table 5.1*) - PDIs' developmental changes partially show different pattern for girls and boys. Specifically, sex differences in *developmental changes* manage to be manifested only in some tasks (SRT, WR) and only in some PDIs (*ct0min*, *ct0max*), and when they are manifested, they show that the initial developmental advantage of boys disappears in later Dev. phases, or even girls get better. We evaluated these task-specific findings within the context of possible methodological factors' effects (equality of girls' and boys' sample size, equality of girls' and boys' PDIs values' variability, extreme values' effects) and didn't find any methodological source of observed sex differences in PDIs' *developmental changes*.

Therefore, the cause of the rare sex differences in *developmental changes* of the best (*ct0min*) and

the worst task performance (*ct0max*) should be sought in some developmental mechanism, and this will be done within discussion of the next (and last) research problem, which is widely related to sex differences in the studied tasks' performance.

Within the scope of the 4th Research problem of this study, we should analyze overall girls' vs boys' differences concerning the cognitive part of performance dynamics while solving the set of equivalent cognitive-motor tasks.

Although the analysis of sex differences in *developmental changes course* revealed that boys in **SRT task** attained higher *best* (*ct0min*) and *worst* (*ct0max*) cognitive performance during the first Dev. Phase, this advantage in the *best performance* disappears during later Dev. Phase. Moreover, in the *worst SRT-performance* girls even over performed boys during the second Dev. Phase (i.e. they showed statistically lower *ct0max*). However, when we observe sex differences in the SRT task cognitive performance over the entire observed period as a whole, *Table 6.1* clearly suggests that the cognitive dynamics of the SRT task performance is in every respect the same in girls and boys. More precisely, boys show statistically insignificant tendency to advantage with the *best cognitive performance*, and girls show same level tendency to advantage with the *worst* and *average cognitive performance*, and with *average non-optimal instability*.

It seems demanding to comment on the obtained findings in relation to previous research because there is very little available research on sex differences in SRT tasks considering cognitive/decision time (DT) in the developmental period of 8-17 years, by using four PDIs.

Regarding the *average cognitive performance* (*cat0*) and *non-optimal cognitive performance instability* (*catnof0*) of the SRT task, in which our study did not find sex differences, only two

comparable studies are available. The study by Eckert and Eichorn (1977) found that boys showed a faster *average cognitive SRT performance* when releasing a key after perceiving a simple visual signal, than girls did. Additionally, the same study did not test sex differences in *performance instability*, although it contained it. On the other hand, the study by Lynn & Ja-Song (1993) did not find differences between 9-years-old girls and boys concerning the *average performance* and *performance instability* of the cognitive part of the SRT task. Nevertheless, there are additional developmental studies of sex differences in *average SRT performance* and *performance instability* that do not separate DT from total RT, which by using total RT confirm a better *average performance* of boys (Dykiert et al, 2012; Ghisletta et al, 2017). However, the same studies show that there are no sex differences in *intra-individual variability* (i.e. *instability*) of SRT performance in the observed development period. Finally, the absence of sex differences in the best (*ct0min*) and worst (*ct0max*) SRT task cognitive performance from our research we could not compare with any other developmental research on SRT task performance because the aforementioned research did not include those two PDIs.

The picture of sex differences in the performance of the cognitive component of the **CRT task** is somewhat simpler. Namely, the analysis of sex differences in *developmental changes course* in the cognitive performance of the CRT-task found that girls and boys show the same pattern of developmental changes across all four PDI. On the other hand, *Table 6.1* shows that girls, at the level of the entire observed development period, showed better cognitive CRT performance across all four PDI, whereby the advantage in the *best performance* indicator (*ct0min*) did not approach statistical significance. Thus, the pattern of developmental changes of the four cognitive PDIs

in girls was systematically shifted to lower values compared to boys, and it can be concluded that the girls' dynamics of the CRT task's cognitive performance in all segments, except for *the best performance*, was superior to the boys' dynamics.

The comparison of this study's findings with the previous researches' findings is again made difficult by the specificity of our research PDIs (only cognitive component of the response was considered, and *the best* and *worst performance* of the CRT task was not used in previous developmental researches). More specifically, the available literature shows that the *average cognitive performance* and the *average instability of the cognitive performance* in CRT task was analyzed only in Lynn and Ya-Song (1993) study. This study reported that 9 years old girls and boys did not differ in terms of the *average performance* and *instability* of the cognitive part of the CRT task. Nevertheless, three additional developmental studies analyzed sex differences in *average performance* and *average instability* of CRT task performance via total RT (rather than just the cognitive part, i.e., DT). Dykiert et al (2012) showed no sex differences in the two PDIs during childhood and adolescence, while Noble, Baker and Jones (1964) showed (1) girls' superiority in the *average CRT tasks performance* from 10 to 13 years of age, and (2) boys' superiority in the rest of the developmental period. The *average instability* of CRT task performance was not investigated in this study. Finally, Žebec et al. (2014) showed that, in the entire developmental period between 8 and 18 years of age, girls show a smaller *average instability* of CRT task performance. However, in terms of *average performance* the picture was more complex: sex differences are not significant throughout the developmental period, while locally girls show better *average performance* at 12, 13 and 16 years of age, and boys at 18 years of age.

The absence of sex differences in *the best CRT task performance*, and the advantage of girls in *the worst CRT task performance* in our study, can be compared with the findings of the study by Žebec et al. (2014). Namely, in that study, boys proved to be more successful in CRT-task *the best performance*, while in regard to *the worst performance*, the picture is more complex: there were no sex differences in the whole observed developmental period, but locally, girls showed better *worst performance* at 12 and 13 years of age, and boys at 10 years of age.

Finally, the analysis of sex differences in *developmental changes course* in the cognitive performance of the **WR-task** revealed the only advantage of girls in *the worst cognitive performance* during the 3rd and 4th Dev. phases (in the previous two phases they were equal to boys). However, when main effect of sex on four PDIs is analyzed at the level of the entire observed development period, then Table 6.1 shows that, across all four PDIs, the dynamics of girls' cognitive performance is superior (although the smallest advantage was shown for *the best cognitive performance*).

Again, it is difficult to comment on these results in relation to previous studies presented in the literature, because of at least two reasons. More precisely, (1) findings on non-average RT measures (*best performance*, *worst performance* and *instability of performance*) are not available to us, and (2) for the *average chronometric performance* of children and adolescents in Go-No Go tasks, only studies using total RT are available, without insight into DT (that we used in our research). The use of total RT includes motor time (MT) in addition to DT, and MT has been consistently shown to be shorter in males of all ages. Therefore, in Go-No Go type RT tasks (that use RT, not DT) boys are given a certain advantage in testing sex differences. In the context of these methodological differences between our

and previous research, it should be noted that girls' superiority in the cognitive performance dynamics in WR task, found in our research, is not recorded in the *average RT* of previous research. In particular, research by Clark et al (2006), Gur et al (2012) and Pascualvaca et al (1997) show that boys' *average RT* is shorter than that of girls in the used tasks of the Go-No Go paradigm (in certain periods of the developmental range of 8-17 years). On the other hand, the developmental research of Roalf et al (2014) establishes that there is no sex difference in the *average RT*. However, in all these studies girls showed significantly fewer performance errors.

When the particular images of sex differences in all four PDIs, considered through the SRT, WR and CRT tasks in developmental period 8-17 years, are put together into one whole, we come to general picture of the sex differences in the totality of cognitive performance dynamics (i.e. linear combinations of all four PDIs: *ct0min*, *ct0max*, *cat0* and *catnof0*). This picture indicates that the overall dynamics of *cognitive performance* of three variously simple cognitive-motor tasks, in the developmental period from 8 to 17 years of age, is mostly superior in girls. However, this girls' advantage is least pronounced in the *best cognitive performance* (*ct0min*), and most pronounced in the *worst performance* (*ct0max*), although their advantages in *cat0* and *catnof0* are not significantly less. Finally, this girls' advantage in the dynamics of the cognitive performance systematically increases from the SRT-task to the CRT-task. In other words, the advantage increases with an increase in the working memory load (although the difference between the advantage of girls in the WR-task and the CRT-task is not large), but also with an increase in the proportion of verbal material in the tasks, based on which one should react.

It is possible to offer several explanations for this girls' superiority (in the period of 8-17 years of age)

in the dynamics of the cognitive component of the WR and CRT tasks performance (while there was no difference in the performance dynamics of the SRT task).

The first explanation is the often-mentioned earlier maturation of cognitive functions in girls (Gur et al, 2012; Lynn, Backhoff, Contreras-Niño, 2004; Lynn, Alik, Must, 2000; Pascualvaca, 1997; Roalf et al, 2014; Waber, 1976; Žebec et al, 2014), related to sexual dimorphism in brain development (De Bellis et al., 2001; Lenroot et al., 2007). This maturational difference, especially after girls enter puberty, leads to higher cognitive level of girls than boys of the same chronological age.

Another explanation is related to the verbal character of the WR and CRT tasks, in which the superiority of girls was manifested in the largest scope of performance dynamics (as opposed to the SRT-task, in which the stimulus did not have a verbal character - a series of 6 X's, which are not used in the Croatian language). Given that the females' superiority in perception and manipulation of verbal content is one of the longest-standing findings on sex differences in cognition (Camarata & Woodcock, 2006; Kimura, 1999; Rajchert, Žultak, Smulczyk, 2014; Robinson & Lubinski, 2011; Roivainen, 2011; Waber, 1976), it is expected that the girls will show an advantage in RT test with the verbal stimuli.

The third explanation relates to the fact that in the RT tasks used, the analysis was carried out on the decision time (DT), in which the motor component of the response is negligible, thus neutralizing the systematically proven advantage of boys (Clark et al, 2006; Era et al, 2010; Gur et al, 2012; Lynn & Ja-Song, 1993; Roalf et al, 2014).

Finally, it is possible that the girls' superiority in the cognitive performance dynamics of the WR and CRT tasks was also contributed to by a different task responding strategy, which is discussed in

neuroscientific research (Clements-Stephens, Rimrodt and Cutting, 2009) or, in research on greater commitment of the males to greater speed at the expense of accuracy (Ibbotson & Roque-Gutierrez, 2023). Nevertheless, this explanation is difficult to advocate because neuroscientific measurements were not used in the research, and sex differences in the speed-accuracy trade off approach were not recorded.

Conclusions

All previously discussed results of the dynamics of cognitive performance of three variously simple cognitive-motor tasks (which gradually include perception, inhibitory mechanisms of attention and simple working memory processes) show that non-average PDIs justify their introduction into developmental research. The evidence for this come from (1) medium mutual correlations of the four PDIs (whose shared variance indicate PDIs are more unrelated than related), (2) from the fact that PDIs are calculated from different parts of intra-individual RT distribution, related to different neuropsychological processes, and (3) from different effect of sex, developmental phase and task complexity on different PDIs.

The description of *age-related changes* of four PDIs through all 10 years of the observed developmental period (8-17 years) confirms the standard finding on non-linear information processing rate improvement (measured by simple cognitive-motor tasks RT) and therefore the validity of the study. The description suggests superiority of girls over the entire developmental period – increasing from the simplest (SRT) to the most complex (CRT) task – both in terms of shorter decision time, and in less variable developmental improvement compared to boys. These findings are different and original contribution to information processing rate development

research, and based on methodological improvement in “cognitive speed” measurement.

Consideration of the four PDIs’ changes *across four developmental phases* of cognitive development (theory of A. Demetriou et al., 2018) showed set of interesting findings. First, these phase-changes are of different intensity at different PDIs, independently of other variables’ effects (the most intensive at *average cognitive performance* of the task, and the least at *average non-optimal performance instability*). The explanation of this differentiated PDIs’ developmental course at least partly comes from the definition and neuropsychological basis of the four PDIs and the related psychometric consequences. Furthermore, the developmental changes of the four PDIs differed for different cognitive-motor tasks: all four PDIs showed the strongest determination by developmental change in WR-task (cognitive speed + inhibition), and the weakest determination by developmental change in CRT-task (cognitive speed + working memory functioning). Possible explanation of this finding is based on general mechanism of development of all cognitive process phases during the task solving, and on the different inter-individual variability of the three cognitive-motor tasks within the observed developmental phases. Finally, PDIs’ developmental changes showed layered relation to the participants’ sex. When these changes are considered independent of the task type, they show the same pattern for girls and boys (although this pattern was shifted towards lower decision times in girls, observed over the entire developmental period). However, when PDIs’ developmental changes are observed separately for each of the three cognitive-motor tasks, at *average performance* and *non-optimal instability* they kept the same pattern in girls and boys. Nevertheless, PDIs’ developmental changes’ pattern was somewhat different in girls and boys at the *best* and *the worst performance* in SRT-task

and WR-task. This sex differentiation of the *best* and *worst performance* development can possibly be explained by sexual dimorphism in the maturation, which is especially evident when entering puberty (3rd developmental phase).

The sex differences analysis of cognitive performance dynamics in three cognitive-motor tasks, *over the entire observed developmental period*, showed that boys and girls do not differ in the performance dynamics of the SRT-task (cognitive speed with non-verbal stimuli). However, in CRT and WR tasks, girls were superior in terms of all four PDIs, although the least in terms of the *best performance*. This girls’ cognitive performance dynamics’ superiority in more complex cognitive-motor tasks is explained primarily by girls’ earlier maturation, superiority in cognitive processing speed of verbal material (represented in the WR and CRT-task) and by the minimal presence of motor component of the response (in which boys usually show an advantage).

The limitations of the study are primarily related to (1) unequal sample sizes from all age groups of the observed developmental range, (2) sub-optimal reliability of decision time measurements, and (3) the use of verbal material in more complex tests. Therefore, the recommendations for future research would include avoiding these limitations, but also some extensions, in order to clarify results interpretation. In that sense, it would be useful to include motor component of response, within developmental phases’ changes and inter-individual variability development in analysis. Moreover, fruitful application of this chronometric aspect of cognitive functioning would be relating its findings with more complex cognitive functions measurement (reasoning, decision making and problem solving).

In conclusion, the findings of the conducted study clearly suggest the use of *non-average indicators* of the cognitive performance dynamics in simple tasks (calculated by decision times, relieved of the motor component of the response). These indicators give layered, broader and partially different picture of ontogenetic development of basic cognitive functions, than the one obtained by analyzing only the *average performance*, measured by total reaction time (RT). This picture becomes even more provocative when the effect of sex and various task complexity are introduced into

analysis, since some stereotypes on male's reaction times superiority are questioned.

Acknowledgments

We thank to the Ministry of Education and Sports of Republic of Croatia that supported this research, but also to all the students that diligently participated in the study and all the school staff who patiently adapted to the organizational requirements of the research. Additionally, we thank the assistant researchers who helped in the implementation of the research.

References

- Anderson, M., Reid, C., & Nelson, J. (2001). Developmental changes in inspection time: What a difference a year makes. *Intelligence*, 29(6), 475-486. doi:10.1016/S0160-2896(01)00073-3
- Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial intelligence*, 149(1), 91-130. doi: 10.1016/S0004-3702(03)00054-7
- Band, G. P., van der Molen, M. W., Overtoom, C. C., & Verbaten, M. N. (2000). The ability to activate and inhibit speeded responses: Separate developmental trends. *Journal of experimental child psychology*, 75(4), 263-290. <https://doi.org/10.1006/jecp.1999.2538>
- Becker, D., Creutzfeldt, O. D., Schwibbe, M., & Wuttke, W. (1982). Changes in physiological. EEG and psychological parameters in women during the spontaneous menstrual cycle and following oral contraceptives. *Psychoneuroendocrinology*, 7(1), 75-90. doi: 10.1016/0306-4530(82)90057-9
- Bobić, J., Pavićević, L., & Gomzi, M. (2002). The difference in complex psychomotor reaction time between patients with and without signs of cerebral circulatory disorders. *Collegium antropologicum*, 26(2), 515-520.
- Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research*, 37(1), 127-160. doi: 10.1207/S15327906MBR3701_06
- Booth, T., Dykiert, D., Corley, J., Gow, A. J., Morris, Z., Muñoz Maniega, S., Royle, N. A., Del C Valdés Hernández, M., Starr, J. M., Penke, L., Bastin, M. E., Wardlaw, J. M., & Deary, I. J. (2019). Reaction time variability and brain white matter integrity. *Neuropsychology*, 33(5), 642-657. <https://doi.org/10.1037/neu0000483>
- Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological review*, 112(1), 117. doi: 10.1037/0033-295X.112.1.117
- Burns, J. T. (1971). Error-induced inhibition in a serial reaction time task. *Journal of experimental psychology*, 90(1), 141. doi: 10.1037/h0031335
- Camarata, S., & Woodcock, R. (2006). Sex differences in processing speed: Developmental effects in males and females. *Intelligence*, 34(3), 231-252. <https://doi.org/10.1016/j.intell.2005.12.001>
- Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the life span. *Acta Psychologica*, 86, 109-197. [https://doi.org/10.1016/0001-6918\(94\)90002-7](https://doi.org/10.1016/0001-6918(94)90002-7)
- Clark, C. R., Paul, R. H., Williams, L. M., Arns, M., Fallahpour, K., Handmer, C., Gordon, E., (2006). Standardized assessment of cognitive functioning during development and aging using an automated touchscreen battery, *Archives of Clinical Neuropsychology*, Vol. 21, Issue 5, 449-467, <https://doi.org/10.1016/j.acn.2006.06.005>
- Clements-Stephens, A. M., Rimrodt, S. L., & Cutting, L. E. (2009). Developmental sex differences in basic visuospatial processing: differences in strategy use?. *Neuroscience letters*, 449(3), 155-160. <https://doi.org/10.1016/j.neulet.2008.10.094>
- Coyle, T. R. (2003). A review of the worst performance rule: Evidence, theory, and alternative hypotheses. *Intelligence*, 31(6), 567-587. doi: 10.1016/S0160-2896(03)00054-0
- Cowan, N., Wood, N. L., Wood, P. K., Keller, T. A., Nugent, L. D., & Keller, C. V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General*, 127, 141-160. <https://doi.org/10.1037/0096-3445.127.2.141>
- Deary, I. J., Der, G., & Ford, G. (2001). Reaction times and intelligence differences: A population-based cohort

- study. *Intelligence*. 29(5). 389-399. doi: [https://doi.org/10.1016/S0160-2896\(01\)00062-9](https://doi.org/10.1016/S0160-2896(01)00062-9)
- Deary, I. J., & Der, G. (2005). Reaction time, age, and cognitive ability: longitudinal findings from age 16 to 63 years in representative population samples. *Aging, Neuropsychology and cognition*. 12(2). 187-215. doi: 10.1080/13825580590969235
- De Bellis, M. D., Keshavan, M. S., Beers, S. R., Hall, J., Frustaci, K., Masalehdan, A., Noll, J., Boring, A. M. (2001). Sex Differences in Brain Maturation during Childhood and Adolescence, *Cerebral Cortex*, Vol. 11, Issue 6, 552–557, <https://doi.org/10.1093/cercor/11.6.552>
- Demetriou, A., & Kyriakides, L. (2006). The functional and developmental organization of cognitive developmental sequences. *British Journal of Educational Psychology*. 76(2). 209-242. doi: 10.1348/000709905X43256
- Demetriou, A., Makris, N., Spanoudis, G., Kazi, S., Shayer, M., Kazali, E. (2018). Mapping the Dimensions of General Intelligence: An Integrated Differential-Developmental Theory. *Human Development*. 61 (1): 4–42. DOI: 10.1159/000484450.
- Demetriou, A., Spanoudis, G., & Shayer, M. (2013). Developmental intelligence: From empirical to hidden constructs. *Intelligence*. 41(5). 744-749. doi: 10.1016/j.intell.2013.07.014
- Demetriou, A., Spanoudis, G., Shayer, M., Mouyi, A., Kazi, S., & Platsidou, M. (2013). Cycles in speed-working memory-G relations: Towards a developmental-differential theory of the mind. *Intelligence*. 41(1). 34-50. doi: 10.1016/j.intell.2012.10.010
- De Weerth, C., van Geert, P., & Hoijtink, H. (1999). Intraindividual variability in infant behavior. *Developmental Psychology*. 35(4). 1102.
- Diehl, M., Hooker, K., & Sliwinski, M. J. (Eds.). (2015). *Handbook of intraindividual variability across the life span*. Routledge.
- Drenovac, M. (2001). An analysis of some attributes of the dynamics of mental processing. *Review of psychology*. 8(1-2). 61-68.
- Dykiert, D., Der, G., Starr, J. M., & Deary, I. J. (2012). Sex differences in reaction time mean and intraindividual variability across the life span. *Developmental psychology*. 48(5). 1262. DOI: 10.1037/a0027550
- Eckert, H. M., & Eichorn, D. H. (1977). Developmental Variability in Reaction Time. *Child Development*, 48(2), 452–458. <https://doi.org/10.2307/1128638>.
- Era, P., Sainio, P., Koskinen, S., Ohlgren, J., Härkänen, T., & Aromaa, A. (2011). Psychomotor speed in a random sample of 7979 subjects aged 30 years and over. *Aging clinical and experimental research*, 23, 135-144. <https://doi.org/10.1007/BF03351077>
- Fischer, K. W., & Yan, Z. (2002). Darwin's construction of the theory of evolution: Microdevelopment of explanations of species variation and change. In N. Granott & J. Parziale (Eds.). *Microdevelopment: Transition processes in development and learning*. Cambridge University Press.
- Fortenbaugh, F. C., DeGutis, J., Germine, L., Wilmer, J. B., Grosso, M., Russo, K., & Esterman, M. (2015). Sustained attention across the life span in a sample of 10,000: Dissociating ability and strategy. *Psychological science*, 26(9), 1497-1510. <https://doi.org/10.1177/0956797615594896>
- Ghisletta, P., Renaud, O., Fagot, D., Lecerf, T., & de Ribaupierre, A. (2018). Age and sex differences in intra-individual variability in a simple reaction time task. *International Journal of Behavioral Development*, 42(2), 294-299. <https://doi.org/10.1177/0165025417739179>.
- Giedd, J. N., Raznahan, A., Mills, K. L., & Lenroot, R. K. (2012). Magnetic resonance imaging of male/female differences in human adolescent brain anatomy. *Biology of sex differences*. 3(19). 1-9. doi: 10.1186/2042-6410-3-19
- Golenia, L., Schoemaker, M. M., Otten, E., Mouton, L. J., & Bongers, R. M. (2017). What the dynamic systems approach can offer for understanding development: An example of mid-childhood reaching. *Frontiers in Psychology*. 8. 1774. doi: 10.3389/fpsyg.2017.01774
- Gorus, E., De Raedt, R., & Mets, T. (2006). Diversity, dispersion and inconsistency of reaction time measures: effects of age and task complexity. *Aging clinical and experimental research*, 18(5), 407–417. <https://doi.org/10.1007/BF03324837>
- Gottfredson, L. S. (2003). The challenge and promise of cognitive career assessment. *Journal of Career Assessment*. 11(2). 115-135. doi: 10.1177/1069072703011002001
- Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., Loughhead, J., Connolly, J. J., Qiu, H., Mentch, F. D., Abou-Sleiman, P. M., Hakonarson, H., & Gur, R. E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology*, 26(2), 251–265. <https://doi.org/10.1037/a0026712>.
- Hale, S., Fry, A. F., & Jessie, K. A. (1993). Effects of practice on speed of information processing in children and adults: Age sensitivity and age invariance. *Developmental psychology*, 29(5), 880-892.
- Hicks, P., & Bolen, L. M. (1996). Review of the Woodcock-Johnson psycho-educational battery-revised. *Journal of School Psychology*. 34(1). 93-102.
- Hoge, R. D. (1999). *Assessing adolescents in educational counseling and other settings*. Routledge. doi: 10.4324/9781410601575
- Ibbotson, P., & Roque-Gutierrez, E. (2023). The development of working memory: Sex differences in accuracy and reaction times. *Journal of cognition and development*, 24(4), 581-597. <https://doi.org/10.1080/15248372.2023.2178437>

- Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Elsevier. doi: 10.1111/j.1744-6570.2008.00111_7.x
- Kail, R. (1997). Processing time, imagery and spatial memory. *Journal of Experimental Child Psychology*. 64(1). 67-78. doi: 10.1006/jecp.1996.2337
- Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, 86, 199–225. [https://doi.org/10.1016/0001-6918\(94\)90003-5](https://doi.org/10.1016/0001-6918(94)90003-5)
- Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: A Bradford book, The MIT Press.
- Kiselev, S., Espy, K. A., & Sheffield, T. (2009). Age-related differences in reaction time task performance in young children. *Journal of experimental child psychology*, 102(2), 150–166. <https://doi.org/10.1016/j.jecp.2008.02.002>
- Kranzler, J. H. (1992). A test of Larson and Alderton's (1990) worst performance rule of reaction time variability. *Personality and Individual Differences*. 13(3). 255-261. doi: 10.1016/0191-8869(92)90099-B
- Larson, G. E., & Alderton, D. L. (1990). Reaction time variability and intelligence: A "worst performance" analysis of individual differences. *Intelligence*. 14(3). 309-325. doi: 10.1016/0160-2896(90)90021-K
- Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., ... & Giedd, J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *Neuroimage*. 36(4). 1065-1073. doi: 10.1016/j.neuroimage.2007.03.053
- Lipps, D. B., Galecki, A. T., & Ashton-Miller, J. A. (2011). On the implications of a sex difference in the reaction times of sprinters at the Beijing Olympics. *PloS one*. 6(10). e26141. doi: 10.1371/journal.pone.0026141
- Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child development*, 75(5), 1357-1372. <https://doi.org/10.1111/j.1467-8624.2004.00745.x>
- Lynn, R., Allik, J., Must, O. (2000). Sex differences in brain size, stature and intelligence in children and adolescents: some evidence from Estonia, *Personality and Individual Differences*, Vol. 29, Issue 3, 555-560, [https://doi.org/10.1016/S0191-8869\(99\)00215-9](https://doi.org/10.1016/S0191-8869(99)00215-9)
- Lynn, R., Backhoff, E., Contreras-Niño, L. A. (2004). Sex differences on g, reasoning and visualisation tested by the progressive matrices among 7–10 year olds: some normative data for Mexico, *Personality and Individual Differences*, Vol. 36, Issue 4, 779-787, [https://doi.org/10.1016/S0191-8869\(03\)00132-6](https://doi.org/10.1016/S0191-8869(03)00132-6).
- Lynn, R., & Ja-Song, M. (1993). Sex Differences in Reaction Times, Decision Times, and Movement Times in British and Korean Children. *The Journal of Genetic Psychology*, 154(2), 209–213. <https://doi.org/10.1080/00221325.1993.9914734>
- Madden, D. J., Pierce, T. W., & Allen, P. A. (1993). Age-related slowing and the time course of semantic priming in visual word identification. *Psychology and Aging*, 8, 490–507.
- McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental psychology*. 38(1). 115.
- Metz, A. J., & Jones, J. E. (2013). Ability and aptitude assessment in career counseling. In D. Brown & W. Lent (Eds.). *Career Development and Counseling: Putting Theory and Research to Work* (pp. 449). New Jersey: John Wiley & Sons. Inc.
- Miller, L. T., & Vernon, P. A. (1997). Developmental changes in speed of information processing in young children. *Developmental Psychology*, 33, 549–554.
- Neubauer, A. C. (1990). Speed of information processing in the Hick paradigm and response latencies in a psychometric intelligence test. *Personality and Individual Differences*. 11(2). 147-152. doi: 10.1016/0191-8869(90)90007-E
- Neubauer, A. C., Riemann, R., Mayer, R., & Angleitner, A. (1997). Intelligence and reaction times in the Hick, Sternberg and Posner paradigms. *Personality and Individual Differences*. 22(6). 885-894. doi: 10.1016/S0191-8869(97)00003-2
- Noble, C. E., Baker, B. L., & Jones, T. A. (1964). Age and Sex Parameters in Psychomotor Learning. *Perceptual and Motor Skills*, 19(3), 935-945. <https://doi.org/10.2466/pms.1964.19.3.935>
- Paraskevopoulou, S. E., Coon, W. G., Brunner, P., Miller, K. J., & Schalk, G. (2021). Within-subject reaction time variability: Role of cortical networks and underlying neurophysiological mechanisms. *NeuroImage*. 237. 118127. <https://doi.org/10.1016/j.neuroimage.2021.118127>
- Pascualvaca, D. M., Anthony, B. J., Arnold, L. E., Rebok, G. W., Ahearn, M. B., Kellam, S. G., & Mirsky, A. F. (1997). Attention performance in an epidemiological sample of urban children: The role of gender and verbal intelligence. *Child Neuropsychology*, 3(1), 13–27. <https://doi.org/10.1080/09297049708401365>.
- Petanjek, Z., Banovac, I., Sedmak, D., Hladnik, A. (2023). *Dendritic Spines: Synaptogenesis and Synaptic Pruning for the Developmental Organization of Brain Circuits*. In: A.A. Rasia-Filho, M.E. Calcagnotto, O. von Bohlen und Halbach (eds) *Dendritic Spines. Advances in Neurobiology*, vol 34. (pp 143 – 221). Springer, Cham. https://doi.org/10.1007/978-3-031-36159-3_4
- Petri, N. M. (2003). Change in strategy of solving psychological tests: evidence of nitrogen narcosis in shallow air-diving. *Undersea and Hyperbaric Medicine*. 30(4). 293-304.
- Radić, J., Ljutić, D., Radić, M., Kovačić, V., Dodig-Ćurković, K., & Šain, M. (2011). Kidney transplantation improves cognitive and

- psychomotor functions in adult hemodialysis patients. *American journal of nephrology*. 34(5). 399-406. doi: 10.1159/000330849
- Rabbitt, P. M. A., Osman, P., Moore, B., & Stollery, B. (2019). There are stable individual differences in performance variability, both from moment to moment and from day to day. In *Cognitive Development and the Ageing Process* (pp. 88-119). Routledge.
- Rajchert, J. M., ŻuŹtak, T., & Smulczyk, M. (2014). Predicting reading literacy and its improvement in the Polish national extension of the PISA study: The role of intelligence, trait-and state-anxiety, socio-economic status and school-type. *Learning and Individual Differences*, 33, 1-11. <https://doi.org/10.1016/j.lindif.2014.04.003>
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological review*. 106(2). 261. doi: 10.1037/0033-295X.106.2.261
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological review*. 111(2). 333. doi: 10.1037/0033-295X.111.2.333
- Reed, T. E. (1998). Causes of intraindividual variability in reaction times: a neurophysiologically oriented review and a new suggestion. *Personality and Individual Differences*. 25(5). 991-998. [https://doi.org/10.1016/S0191-8869\(98\)00123-8](https://doi.org/10.1016/S0191-8869(98)00123-8)
- Ridderinkhof, K. R., & van der Molen, M. W. (1997). Mental resources, processing speed, and inhibitory control: A developmental perspective. *Biological Psychology*, 45, 241-261. [https://doi.org/10.1016/S0301-0511\(96\)05230-1](https://doi.org/10.1016/S0301-0511(96)05230-1)
- Roalf, D. R., Gur, R. E., Ruparel, K., Calkins, M. E., Satterthwaite, T. D., Bilker, W. B., Hakonarson, H., Harris, L. J., Gur, R. C. (2014). Within-individual variability in neurocognitive performance: Age- and sex-related differences in children and youths from ages 8 to 21. *Neuropsychology*, Vol 28(4), 506-518. <https://doi.org/10.1037/neu0000067>
- Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, 48(2), 268-302. <https://doi.org/10.3102/0002831210372249>
- Roivainen, E. (2011). Gender differences in processing speed: A review of recent research. *Learning and Individual Differences*, 21(2), 145-149. <https://doi.org/10.1016/j.lindif.2010.11.021>
- Salvia, E., Petit, C., Champely, S., Chomette, R., Di Rienzo, F., Collet, C., (2016). Effects of Age and Task Load on Drivers' Response Accuracy and Reaction Time When Responding to Traffic Lights. *Frontiers in Aging Neuroscience*, Vol.8, 169. <https://doi.org/10.3389/fnagi.2016.00169>
- Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of experimental psychology: General*. 136(3). 414. doi: 10.1037/0096-3445.136.3.414
- Schubert, A. L. (2019). A meta-analysis of the worst performance rule. *Intelligence*. 73. 88-100. doi: 10.1016/j.intell.2019.02.003
- Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current directions in psychological science*. 3(1). 1-5. doi: 10.1111/1467-8721.ep10769817
- Slifkin, A. B., & Newell, K. M. (1998). Is variability in human performance a reflection of system noise?. *Current directions in psychological science*. 7(6). 170-177. doi: 10.1111/1467-8721.ep10836906
- Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in cognitive sciences*. 7(8). 343-348.
- Tadinac, M. (1993). *Ispitivanje lateralizacije funkcija mozgovnih hemisfera tehnikom podijeljenog vidnog polja : doktorska disertacija*. Zagreb: M. Tadinac-Babić
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT press.
- Thomas, J. R., & French, K. E. (1985). Gender differences across age in motor performance: A meta-analysis. *Psychological bulletin*. 98(2). 260. doi: 10.1037/0033-2909.98.2.260
- Travis, F. (1998). Cortical and cognitive development in 4th, 8th and 12th grade students: The contribution of speed of processing and executive functioning to cognitive development. *Biological psychology*, 48(1), 37-56. [https://doi.org/10.1016/S0301-0511\(98\)00005-2](https://doi.org/10.1016/S0301-0511(98)00005-2)
- Tun, P. A., & Lachman, M. E. (2008). Age differences in reaction time and attention in a national telephone sample of adults: education, sex, and task complexity matter. *Developmental psychology*, 44(5), 1421-1429. <https://doi.org/10.1037/a0012845>
- Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. *Intelligence*. 38(1). 111-122. doi: 10.1016/j.intell.2009.08.002
- Van Dijk, M., & Van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of continuity, discontinuity and variability in early language development. *Infant and Child Development: An International Journal of Research and Practice*. 16(1). 7-33. doi: 10.1002/icd.506
- Van Geert, P., & Van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*. 25(4). 340-374. doi: 10.1016/S0163-6383(02)00140-6
- Waber, D. P. (1976). Sex Differences in Cognition: A Function of Maturation Rate? *Science* 192, 572-574, doi:10.1126/science.1257795

- Wai, J., Brown, M. I., & Chabris, C. F. (2018). Using standardized test scores to include general cognitive ability in education research and policy. *Journal of Intelligence*. 6(3). 37. doi: 10.3390/jintelligence6030037
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review*. 9. 625-636. doi: 10.3758/BF03196322
- Yan, Z., & Fischer, K. (2002). Always under construction: Dynamic variations in adult cognitive microdevelopment. *Human development*. 45(3). 141-160. doi: 10.1159/000057070
- Žebec, S. M. (2005). *Dobne promjene čovjekova sustava obrade podataka i njihova uloga u kognitivnom razvoju: neobjavljena doktorska disertacija*. Zagreb: Mislav Stjepan Žebec.
- Žebec, M.S., Demetriou, A. and Kotrla Topić, M. (2015). Changing expressions of general intelligence in development: A 2-wave longitudinal study from 7 to 18 years of age. *Intelligence*. 49. 94-109. <https://doi.org/10.1016/j.intell.2015.01.004>
- Žebec, M. S., Budimir, S., Merkaš, M., Szivovicza, L., & Živičnjak, M. (2014). Sex-specific age-related changes of information processing rate indicators during childhood and adolescence. *Collegium antropologicum*. 38(2). 397-408.

Appendix 1. Calculation formulas for four performance dynamics indicators

A. minimal RT: $ct0min = T^0(\text{rank} = 1) \text{ (ms)}$

B. maximal RT: $ct0max = T^0(\text{rank} = N) \text{ (ms)}$

C. mean RT: $cat0 = \frac{1}{N^0} \cdot \sum_{i=1}^{N^0} T_i^0 \text{ (ms)}$

Whereby T_i^0 represents the correct answer in the i-th task of the test, which followed the correct answer of the previous task.

D. mean deviation of response time from best RT: $catnof0 = \frac{ctnof0}{N^0} \text{ (ms)}$

Whereby $ctnof0$ is calculated by formula: $ctnof0 = \sum_{i=1}^{N^0} [T_i^0 - ctfp]$

And $ctfp$ means cognitive time of functional potential and is calculated via formula:

$$ctfp = \frac{T^0(\text{rank} = 1) + T^0(\text{rank} = 2) + T^0(\text{rank} = 3)}{3}$$

Here $T^0(\text{rank}=1, 2, \text{ or } 3)$ indicates the time of the first, second, and third shortest correct answer time, which followed the correct answer of the previous task.