Intelligence 49 (2015) 94–109

Contents lists available at ScienceDirect

Intelligence

Changing expressions of general intelligence in development: A 2-wave longitudinal study from 7 to 18 years of age

Mislav Stjepan Žebec [a,*], Andreas Demetriou [b,**], Marina Kotrla-Topić [c]

[a] University of Zagreb, Croatia
[b] University of Nicosia, Cyprus
[c] Institute of Social Sciences Ivo Pilar, Zagreb, Croatia

ARTICLE INFO

Article history:
Received 6 October 2014
Received in revised form 12 January 2015
Accepted 12 January 2015
Available online 28 January 2015

Keywords:
Attention control
Intelligence
Intellectual development
Processing speed
Working memory

ABSTRACT

We present a study which investigated the inter-relations between processing speed, attention control, working memory, fluid intelligence, and mathematical reasoning from 7 to 18 years of age. To fulfil this aim, 478 participants drawn from each of the age years 7-17 years at first testing were examined twice, separated by a 12-month interval. Several simple reaction time, divided attention, and selective attention tasks examined processing efficiency. Forward and backward digit span tasks addressed working memory. Raven's standard progressive matrices addressed fluid intelligence and a task battery addressed to mathematical reasoning addressed its investment into a demanding cognitive domain. Relations between processes were explored by several types of structural equation models applied in three age groups: 8-10, 11-13, and 14-18 years. A powerful common general factor underlying all processes at both testing waves in all three age phases was found. The relative weight of these processes in the formation of this grand G differed between phases, with working memory, attention control, and Gf dominating in the three phases, respectively. Cross-lagged modeling revealed three tiers of mental organization (processing, representational, and inferential efficiency) interlinked by a core control program. This core is transcribed into inferential and problem solving ensembles of increasing compositionality at successive developmental phases. Implications for developmental and differential theories of intelligence are discussed.

There is general agreement that human intelligence involves information integration processes activated when dealing with new information or problems. Inductive and deductive inference and problem solving in different domains, such as mathematics, are important mechanisms of information integration. They underlie psychometric fluid intelligence (Gf) (Carroll, 1993; Jensen, 1998; Spearman, 1927) and reasoning studied by developmental (Case, 1985; Halford, Wilson, & Phillips, 1998; Piaget, 1970) and cognitive researchers (e.g., Johnson-Laird, 2013; Rips, 2001). Information processing theories of human intelligence maintain that individual or developmental differences in Gf and reasoning reflect differences in various aspects of processing and representational efficiency, such as processing speed (Jensen, 1998) and working memory capacity (WMC), respectively (Case, 1985; Cowan, Morey, Chen, & Bunting, 2007; Demetriou, Christou, Spanoudis, & Platsidou, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Kyllonen & Christal, 1990). In their developmental version, several theories assume that changes in efficiency reflect changes in executive control. These are assumed to enable individuals to better attend to relevant information and

* Correspondence to: M.S. Zebec, Department of Psychology, Centre for Croatian Studies University of Zagreb, Borongajska cesta 83d, 10 000 Zagreb, Croatia.
** Correspondence to: A. Demetriou, University of Nicosia Research Foundation, University of Nicosia, Makedonitissas Avenue, P.O. Box 24005, 1700 Nicosia, Cyprus.
E-mail addresses: mzebec@hrstud.hr (M.S. Žebec), ademetriou@ucy.ac.cy, demetriou.a@unic.ac.cy (A. Demetriou).

http://dx.doi.org/10.1016/j.intell.2015.01.004
0160-2896/

handle it in working memory and inference (Case, 1985; Diamond, 2013; Zelazo et al., 2003).

However, the exact role of each of these factors is still debated. Some researchers emphasize speed as a purer index of the quality of information processing in the brain (e.g., Coyle, Pillow, Snyder, & Kochunov, 2011; Jensen, 1998, 2006). Others emphasize working memory capacity because it is the workspace of thinking (Kyllonen & Christal, 1990). Others emphasize executive control of attention, which allows self-directed deployment of information selection, processing, decision making, and action. In their view, executive control of attention is common to all, speed, WM, and Gf, explaining their relations (Cowan et al., 2007; Engle et al., 1999; Stankov & Roberts, 1997). Finally, others assume a causal linear relation between them such that changes in speed cause changes (or differences) in control of attention which enhance working memory which, in turn, cause changes (or differences) in Gf (Case, 1985; Coyle et al., 2011; Kail, 1991; Kail & Ferrer, 2007; Nettelbeck & Burns, 2010).

It is commonly accepted that from childhood to early adulthood, speed increases (Kail, 1991, 2007; Kail & Ferrer, 2007), attention becomes more focused, flexible, and selective (Brydges, Fox, Reid, & Anderson, 2014; Diamond, 2013; Pascual-Leone & Johnson, 2011), working memory capacity expands (Case, 1985; Halford et al., 1998; Pascual-Leone, 1970), and Gf can deal with concepts and problems of increasing complexity (e.g., Halford et al., 1998; Piaget, 1970). Demetriou et al. (2013), Demetriou, Spanoudis, Shayer, van der Ven, Brydges, et al. (2014)) showed recently that the relations between these constructs are more complicated than originally assumed. Specifically, they suggested that a common core of processes underlying Gf is always present in intellectual functioning. However, this core is systematically transformed in development. It changes in the kind of representations that it can handle at successive developmenal cycles (e.g., visual → verbal → abstract representations in the three last cycles respectively), the relations that it can build between representations, and the awareness about them. As a result, its relations to measures of processing and representational efficiency, such as speed and WMC, vary in development as a function of its current state.

The developmental version of Gf involves three fundamental processes which are always present in inference and problem solving: Abstraction, alignment, and cognizance (AACog). Abstraction spots or induces similarities between patterns of information, using mechanisms that may vary in development, such as perceptually based induction in infancy and deduction later on. Alignment is a relational mechanism that maps representations onto each other, enabling comparisons driven by current understanding or learning goals. Cognizance is awareness of the objects of cognition, cognitive processes, and cognitive goals. Executive control is a special expression of cognizance in that it reflects the self-regulation possibilities allowed by cognizance. Conceptual development is self-propelled because AACog continuously generates new mental content expressed in representations of increasing inclusiveness and resolution (Demetriou, Spanoudis, & Shayer, 2014).

AACog evolves through four major developmental cycles, with two phases in each. New representations emerge early in each cycle and their alignment dominates later. Each cycle culminates into insight about the cycle's representations and underlying inferential processes that is expressed into executive programs of increasing flexibility. These programs activate transition to the next cycle. In succession, the four cycles operate with episodic representations (birth to 2 years), mental representations (2-6 years), rule-based concepts (6-10 years), and principle-based concepts (11-18 years). Transitions within cycles occur at 4 years, 8 years, and 14 years, when relations between the representational units constructed earlier are worked out (Spanoudis, Demetriou, Kazi, Giorgalla, & Zenonos, 2015). It is notable that, despite differences in descriptions and interpretations, these cycles have been identified by all students of intellectual development (e.g., Case, 1985; Fischer, 1980; Pascual-Leone, 1970; Piaget, 1970). This convergence indicates a strong developmental phenomenon that needs to be understood.

In this paper we focus on the two cycles attained after the age of 6 years as this study is concerned only with them. In the first phase of the rule-based concepts cycle, at 6-8 years, there is a shift from "realistic" representations that are visible to the "mind's eye" to the inferential threads inter-linking them. At the beginning these function as semantic blocks defining generic concepts, such as object classes, number, causal attributions, etc. The integration of various conceptual spaces related to number, such as object arrays, number words, counting, digits, etc., into a common mental number line is a good example of an underlying mental construct in the domain of quantitative reasoning. In the next phase, the rules defining semantic blocks can systematically be aligned with each other, allowing grasping how two or more dimensions intersect with each other defining new forms of objects. Early in the next cycle, at 11-13 years, children grasp relations between rules and encode them as such. Thus, conceptual spaces may be explored as such in reference to one (in the first phase) or more (in the second phase) alternative principles. Analogical and algebraic reasoning in adolescence reflect this possibility. The four levels will be instantiated in Method in reference to the various batteries used.

Demetriou et al. (2013), Demetriou, Spanoudis, & Shayer (2014), Demetriou, Spanoudis, Shayer, van der Ven, et al. (2014) showed that changes in Gf were predicted by speed at the first phase of each cycle (i.e., at 6-8 years and 11-13 years) and by working memory at the second phase (i.e., 4-6 years, 8-10 years, and 13-16 years). They suggested that this pattern reflects differences in the processing requirements of developmental acquisitions. At the beginning of cycles processing speed is a better index because thought in terms of the new mental units is automated and expands fast over different contents. Later in the cycle, when networks of relations between representations are worked out, WMC is a better index because alignment and inter-linking of representations both requires and facilitates WMC. However, speed and WMC index rather than cause transitions in reconceptualization. Executive control and associated awareness of mental processes also change. Spanoudis et al. (2015) found that awareness mediates between processing and representational efficiency and thought, reflecting shifts in the level of executive control that individuals may exercise.

This article aims to further explore the relations between the main constructs of interest from late childhood to late adolescence. Specifically, we focus on the last phase of the cycle of rule-based concepts (i.e., 8-10 years) and the whole cycle of principle-based concepts (11-18 years). Consistent with our aims, we examined children drawn from each of the years 7 through 17 by several tasks addressed to speed of processing, selective and divided attention, working memory, Gf as addressed by Raven's matrices, and mathematical reasoning. All participants were examined twice on all tasks; one year elapsed from first to second testing wave. This design allows testing the following predictions:

1. It is commonplace that all processes develop throughout the years studied here. However, phase boundaries would be discernible in variations in the rate of change captured by longitudinal measures of the various processes. These variations would reflect changes in the relations between processes accompanying reorganizations of processes in AACog. For instance, variations in the rate of change of control process across phases would reflect the different states of automation or alignment associated with successive AACog developmental phases.
2. Despite the variations above, a general factor underlying all constructs across phases and testing waves would have to be found. This would reflect the operation of the AACog processes underlying Gf and mathematical tasks and its interaction with tools of mental efficiency such as attention control and representational capacity or general states of mental efficiency, such as processing speed. In other words, this common factor would reflect both common processes and the developmental bootstrapping orchestrating their change over time.
3. The precise expression of this factor would be phase-specific, to reflect the state of relations between the various processes. That is, the recycling patterns outlined above suggest that in some phases (i.e., 8-10 and 14-17 years) working memory would dominate as an index of G and in other phases (i.e., 11-13 years) processing efficiency would dominate.
4. Executive control becomes increasingly refined with development as it refocuses from processes directly connecting the mind with the world (e.g., perception) to processes internally focused (e.g., backward recall of information). The three measures of executive control used in the present study (i.e., divided attention, selective attention and inhibition, and working memory are systematically arranged along this "touching the world-inwardly directed processes" continuum. Thus, executive control would express itself as divided attention in the first phase, from 8-10 years, to indicate flexibility in attentional control which allowsstimulus selection and stimulus-response alignment. In the following phase, from 11-13 years, it expresses itself as inhibition control as captured by the Stroop task; this reflects mastery of mental processes as such.

1. Method

1.1. Participants

A total of 478 right-handed participants from 7 through 17 years of age (249 male), at first testing, were involved. The exact composition of this sample is shown in Table 1. These participants were all native speakers of Croatian and lived in Zagreb, Croatia's capital. They were students in Croatian public schools and thus SES is about equally represented in each age group. All participants were tested two times, separated by one year. Their age at first testing is shown in Table 1.

Table 1
Composition of sample and mean age (and SD) at first testing wave.

Age group	N	Mean age	SD
7	17	7.84	.11
8	77	8.56	.30
9	60	9.36	.27
10	22	10.79	.14
11	55	11.39	.23
12	36	12.71	.19
13	59	13.30	.20
14	20	14.77	.16
15	50	15.41	.24
16	40	16.63	.20
17	42	17.31	.20

Note 1. The second testing took place 12 months after the first testing.

1.2. Tasks batteries

1.2.1. Processing efficiency: Speed, selective and divided attention

The *MID-KOGTESTER1* was used to test speed of processing, selective and divided attention. This is a computer based test battery that contains eight cognitive tests. The apparatus included a laptop with the necessary software installed, a computer screen for stimulus presentation and two specifically designed, removable response panels (Panel 1 and Panel 2). Panel 1 contains one start key (centrally positioned) and five target keys, equally distanced from the start key (in a semicircle). Panel 2 contains the start key (centrally positioned) and two target keys equally distanced from the start key (left and right of the start key). The participant sits in front of the computer screen and one or two response panels—depending on the test—with basically the same two-step task in every test: (1) Hold the index finger on the start key and carefully look at the fixation mark (violet circle) on the computer screen until the stimulus is presented; (2) quickly retrieve the appropriate response instruction after the stimulus was presented and respond on it by lifting the index finger from the start key and landing it on the target key as fast as possible. It is noted that in all reaction time tasks only decision time was used. This is the time between the stimulus presentation and finger lifting.

Also, all tests involved several practice trials at the beginning. Practice ended when the criterion of 10 or 12 correct answers was met. The main component of each task involved a minimum of 20 correct trials (SRT RH) and some tasks (e.g., CRT-C, CRT-NC, and Stroop) involved up to 32 correct trials. This restriction was deemed necessary to ensure that the main component of the test involved the minimal number of correct trials that would allow an accurate examination of the cognitive processes addressed. Notably, error responses were very limited (mean errors over all tests were 8.7% of the total number of responses).

(1) *Simple reaction time with right (dominant) hand (SRT-RH)*. Participant's index finger rested on the start key with eyes fixated at the fixation mark at the center of the screen. When the stimulus appeared on the fixation mark, she had to lift the finger and press the target key

as fast as possible. To make tasks perceptually similar, color recognition items were made of sets of six same color Xs. This was the average letter length of color names in Croatian, which was the language of the study (e.g., zeleno, meaning green). Set color varied randomly across trials. Participants were asked to recognize Xs' ink color. Ink color varied randomly across trials (they were all red, blue, white or green); the time between pressing the target key and appearance of the new stimulus also varied randomly between 0.75 and 2.5 seconds. The test ended after 10 correct practice and 20 main component trials.

(2) *Simple reaction time with left (non-dominant) hand (SRT-LH).* This test is identical to SRT-RH in all respects but the following: (1) It is performed with the non-dominant hand; (2) the fixation mark is presented on the left quarter of the screen; (3) it ended after 30 correct trials. It involved 30 stimuli (i.e. 6 sets of Xs printed in red or blue or white or green) and 10 exercise trials.

(3) *Word recognition (WR).* Participant's task was to decide either to keep the right index resting on the start key of Panel 1 if a distractor stimulus appears (the name of one of the four colors – RED, BLUE, GREEN, or WHITE was printed in violet), or press the target key (vertically positioned) if the target stimulus appears (the word COLOR printed in violet). This task is similar to the go/no go tasks used in developmental research (van der Ven, Kroesbergen, and Leseman (2012). The test involved 10 correct practice trials (with 6 distractor and 4 target stimuli) and 30 proper trials (with 20 distractor and 10 target stimuli). Presentation order of target and distractor stimuli was random and the time between end of responding and appearance of the new stimulus also varied randomly between 0.75 and 2.5 seconds.

(4) *Choice reaction time to color (CRT-C).* This test was performed with the right hand on the panel 1. Participant's index rested on the start key until the stimulus appeared at the position of the fixation mark, when she had to press the target key corresponding to the color of the target stimulus. There were 4 target stimuli (sets of 6 same color Xs printed in red, blue, white, or green) corresponding to 4 target keys. The sets included four different response instructions involving the stimulus color-target key combination, randomized across participants. The test comprised 12 practice trials and 32 correctly responded items. Time between items also varied randomly between 0.75 and 2.5 seconds.

(5) *Choice reaction time to color name* (CRT-NC); This test was identical to CRT-C in all respects but the stimuli participants reacted to (i.e., the names of the four colors noted above all printed in violet).

(6) *Stroop test.* This test included a set of congruent and a set of incongruent items. There were 32 proper trials. Eight of them were congruent and they were identical to the CRT-NC in all respects but in that each of the four color names was written in its own ink color. There were also 24 incongruent items where each color word was printed in a different ink color. Selection of response keys on Panel 1indicated the color in which the word was written. There were 12 practice trials (4 congruent and 8 non-congruent).

(7) *Object size classification* (OSC). The participant's index finger rested on the start key on panel 2 with eyes fixated at the fixation mark at the center of the screen. When target stimulus appeared on the right of the fixation mark participants had to press the response key indicating if the object on the screen was larger or smaller than the participant; objects were either clearly larger (e.g., elephant, town, church, forest) or clearly smaller than the participant (e.g., ant, egg, shoe, pipe). There were 10 practice items and 30 test items, each including the same number of larger and smaller objects, alternating randomly.

(8) *Divided attention* (DA). This test required simultaneous responding to two different tasks on the two panels, where the stimuli were presented in fast succession (50 to 250 ms). Task 1 was a simple reaction time task from the SRT-LH test. Task 2 is an object size classification task from the OSC test. Participants were asked to respond to Task 1 with the left hand on Panel 1, and on to Task 2 with the right hand on Panel 2. At first, participants looked carefully at the fixation point on the left side of the screen with both index fingers resting on the start keys of the two response panels. When the stimulus for Task 1 appeared at the fixation point, participants had to lift their left index finger from the start key and press the (only) target key on the left response panel, as fast as possible. At the same time, they had to look at the right side of the screen where the Task 2 stimulus (small/large object) was concurrently presented, in order to choose the object category by selecting the correct target key on the right response panel, using their right hand. Participants responded in a fixed order but in very close succession (first Task 1 and then Task 2). There were 10 (correct) practice trials and 30 test trials.

It is noted that SRT (with right and left hand) and WR was located in the 5-choice Panel 1. The four non-target keys of this panel were not masked in these tests but the target key was privileged in relation to the other four keys because it was located right above the resting key, with the others being off side. That was reflected in the very small number of errors observed in SRT-RH and SRT-LH (5.5%) and WR (3.6%).

The tasks above provided the evidence necessary to obtain measures of all three main dimensions of processing efficiency that are of interest here. Specifically, simple reaction times generated measures of processing speed. Tasks causing interference between *two or more* processes (in working memory or reasoning) generated measures of divided attention (DA). In the present context, dual task performance (i.e., participants both performed either on Task 1 or on Task 2 alone and also both of them together) allowed to extract a measure of dual task interference. This is the difference between RT on Task 2 performed together with the Task 1 and RT on Task 2 performed separately. This difference between RT in the dual task condition and to the single task is widely used as an operationalization of divided attention (Pashler, 1998). Selective attention is the selection of a specific stimulus embedded in a stimulus ensemble, such that selection of the target stimulus requires ignoring other stimulus components involved in this ensemble (Pashler, 1998; Sternberg, 1966). Thus, in these tasks there is interference between processing the target stimulus and the inhibition necessary to halt processing of irrelevant stimuli. The Stroop effect is regarded as a powerful measure of selective attention interference (MacLeod, 1991; Pashler, 1998). In the present context, this was the difference between mean reaction time on the non-congruent Stroop tasks and the mean reaction time on the color choice reaction tasks.

Precision was high because RTs were recorded at 1 ms measurement scale in all cases. Presentation order was randomized across participants and items within tasks to control for practice or fatigue. Test-retest reliability (across the two testing waves) was high, varying between .7 (CRT-C) and .85 (DA).

1.2.2. Working memory

The extended version of forward (FDS) and backward digit span (BDS) test included in the Wechsler intelligence scale test for children (WISC) was used to measure components of working memory. In Baddeley's (2012) terms, the first is primarily addressed to the short-term memory component that primarily activates the phonological loop; the second measures working memory because it activates the executive component directly. The FDS test comprised 10 pairs of digit sequences varying from 2 to 11 digits. In each pair there are two equivalent sequences of the randomly chosen one-digit numbers from the 0-9 range. Participants recalled the sequences in presentation order. The BDS test included 8 pairs of digit sequences, varying from 2 to 9 digits. Participants were asked to reproduce the digit sequence backwards starting from the last digit presented. In both tests testing stopped when the trials of a sequence were failed; the previous successful sequence was credited as the person's FDS of BDS.

Test-retest reliability across the two testing waves was satisfactory both for the FDS (.69) and the BDS (.66).

1.2.3. Gf and mathematics

1.2.3.1. Raven's standard progressive matrices. Raven's Standard Progressive Matrices were used to address fluid intelligence. SPM include five sets (or Series) of matrices of increasing complexity. Matrices in Series A require grasping the pattern underlying figures varying along a single dimension. Matrices in Series B require integration of two familiar and obvious dimensions (e.g., shape, size, background, etc.) and the most complex of them require some kind of mental rotation. Matrices in Series C require deciphering a critical dimension through systematic search and transformation of one or more features of the matrices involved. Therefore, matrices in Series C seem to require processes associated with the first phase of the principle-based concepts. Series D and E are similar to Series C in that they also require deciphering dimensions through grasping the thread underlying several transformations of figures. However, each of these transformations occurs on more than one dimension, making them difficult to identify.

In terms of the levels outlined in the introduction, matrices in Series A require abstracting the rule underlying the various figures in a matrix and use it to select the missing matrix among the options. Therefore, this Series addresses inferential abilities associated with the first phase of the rule-based concepts cycle. Matrices in Series B require thought processes associated with the alignment phase of this cycle because two or more rules must be integrated for solution. Matrices in Series C require the ability to formulate a general principle underlying transformations. Series D and E require alignment of several principles. Thus Series C addresses abilities primarily related to the first phase and Series D and E address abilities related to the second phase of the principle-based concepts.

1.2.3.2. Mathematics. A paper-and-pencil battery of mathematical reasoning was administered to the participants. The battery addressed two domains of mathematical reasoning: Arithmetic and algebra. Items in each domain were scaled in four levels of difficulty. In the arithmetic tasks, participants were asked to specify the operations missing from simple arithmetic equations: One (e.g., 5 * 3 = 8), two (e.g., {4 # 2} * 2 = 6), three (e.g., {3 * 2 # 4} @ 5 = 7), and four operations (e.g., {5 @ 2} o 4 = {12 $ 1} * 2) were missing from the items of each level. It is noted that it was made clear to the participants that the meaning of symbols may vary across tasks. This was necessary to ensure that no fixed symbol-operation relation can be transferred across tasks.

The algebraic reasoning tasks required to specify one or more unknowns in an equation. At the first level, the solution could be directly deduced from the elements given (e.g. $a + 5 = 8$, $a = ?$). At the second level, participants must co-ordinate two fully defined symbolic structures to specify the unknown (e.g. $u = f + 3$; $f = 1$; $u = ?$). At the third level, problems required to grasp the relation between two mutually defined structures in order to specify the unknown (e.g., if ($r = s + t$) and ($r + s + t = 30$), specify $r = ?$)). At the fourth level problems required to grasp number as an abstract variable and thus specify general relations rather than specific values (e.g., when is true that $\{L + M + N\} = \{L + P + N\}$?).

In terms of the cycles of development specified in the introduction, the two lower levels of these batteries are primarily related to the two phases of the rule-based concepts cycle, respectively. All first level items required to abstract a general rule running through the various instances. All second level items required to grasp two rules or two instantiations of a rule and mapped them onto each other. Levels three and four of all these batteries addressed the two phases of the principles cycle, respectively. Level three tasks required grasping an underlying general principle relating the various items involved and level four tasks required alignment of two or more principles.

This battery was found to have good psychometric and developmental properties in several studies (Demetriou, Pachaury, Metallidou, & Kazi, 1996). In the present sample *discriminability* (average index of difficulty of 35 tasks is .52 and Ferguson's Δ is .98) and *reliability* were high (Cronbach $\alpha = .92$, split-half = .95).

1.2.4. Handling of outliers

The "Schweinle Method" of data screening (Mickey, Dunn, & Clark, 2004) was adopted for screening outliers within age groups. For all RT tasks, participants with a z-score > 2.57 were considered as outliers (i.e., 1% of the highest and/or the lowest results in a normal distribution). For the working memory and the cognitive tasks, participants with a z-score < -2.57 were considered as outliers, because, in their case, attainement of high scores by chance or error is highly unlikely. Percentage of outliers ranged from 2.33 to 4.02 %, with a means of 3.11%. Their possible influence was controlled by replacing the outlier value with the mean of the respective age group (see Table 1 of supplementaty material).

2. Results

To test the hypotheses of the study two types of analyses were run. First, a series of ANOVAs explored the relations of each process with age. The results of these analyses are primarily related to the first prediction stated in the introduction. Second, a series of structural equation models (SEM) explored the relations between processes in each of the three age phases specified in the introduction. These models are related to predictions 2-4. The results of these analyses are summarized below.

2.1. Relations with age

2.1.1. Processing efficiency

To highlight the relations of speed, divided, and selective attention with age, the two mean simple RT scores (means of the two simple reaction tasks and the word recognition task at each testing wave), the two scores standing for divided attention (means of the two divided attention scores at each testing wave), and the two scores standing for selective attention obtained at each of the two testing waves were subjected to an 11 (age groups) x 2 (waves) x 3 (processes) repeated measures ANOVA. The trends captured by this analysis are shown in Fig. 1 (see complete descriptive statistics in Table 2 of supplementary material). The main effect of age was highly significant and strong, $F_{10,467} = 95.12$, $p < .0001$, $\eta^2 = .67$, indicating that reaction times decreased extensively over all conditions throughout the age span covered by the study. The effect of processing efficiency condition, $F_{2,466} = 1401.89$, $p < .0001$, $\eta^2 = .86$, was also very strong, reflecting the fact that there were large differences between measures (316.37, 1047.20, and 170.00 ms, for simple RT, divided attention, and selective attention, respectively). The effect of testing wave was also highly significant and strong, $F_{1,467} = 57.89$,$p < .0001$, $\eta^2 = .13$, reflecting the fact that performance improved from the first to the second wave. The interactions between these factors were also significant, reflecting the fact that decrease of reaction times with age and testing waves varied as a function of phase and/or condition. Specifically, the age x condition interaction, $F_{20,934} = 22.60$, $p < .0001$, $\eta^2 = .33$, indicated that age differences in reaction times varied with complexity of the measures involved: The more complex the measure (e.g., divided attention vs. simple reaction time) the steeper the difference between measures, especially in the 7 to 10 years phase. The age x testing wave interaction, $F_{10,467} = 7.94$, $p < .0001$, $\eta^2 = .15$, reflected the fact that improvement across waves was larger earlier in age rather than later. Overall, the change from the first to the second testing wave was larger in the 7 to 8 and the 10 to 11 years phase, signifying the phase shift that occurs over these two age periods. Notably, there was practically no change between testing waves after the age of 13 years. The significant condition x testing waves, $F_{2,466} = 36.78$, $p < .0001$, $\eta^2 = .14$, and the age x waves x condition, $F_{20,934} = 6.22$, $p < .0001$, $\eta^2 = .12$, suggested that improvement across testing waves was more pronounced at some age periods rather than other for some measures compared to other measures. Specifically, divided attention improved extensively from 7 to 8 and from 8 to 9 years phase (from circa 1.9 sec to 1.6 sec). There was also a noticeable but smaller improvement from 11 to 12 years. Longitudinal improvement from 9 to 10 or after 12 years was very limited, if existent at all. Selective attention improved more from 9 to 10 and again from 12 through 14 years. Therefore, both divided and selective attention improve extensively in the 7-10 years phase but only selective attention continued to improve in the next phase.

2.1.2. Working memory

To specify the relation of working memory with age, an 11 (age groups) x 2 (FDS vs. BDS) x 2 (testing waves) repeated measures ANOVA was applied on the performance attained on the two working memory tasks across the two testing waves. The results of this analysis are summarized in Fig. 2 (see complete descriptive statistics in Table 3 of supplementary material). The main effect of age was highly significant and very strong, $F_{10,467} = 57.56$, $p < .001$, $\eta^2 = .55$, reflecting the fact that performance improved extensively in both tasks

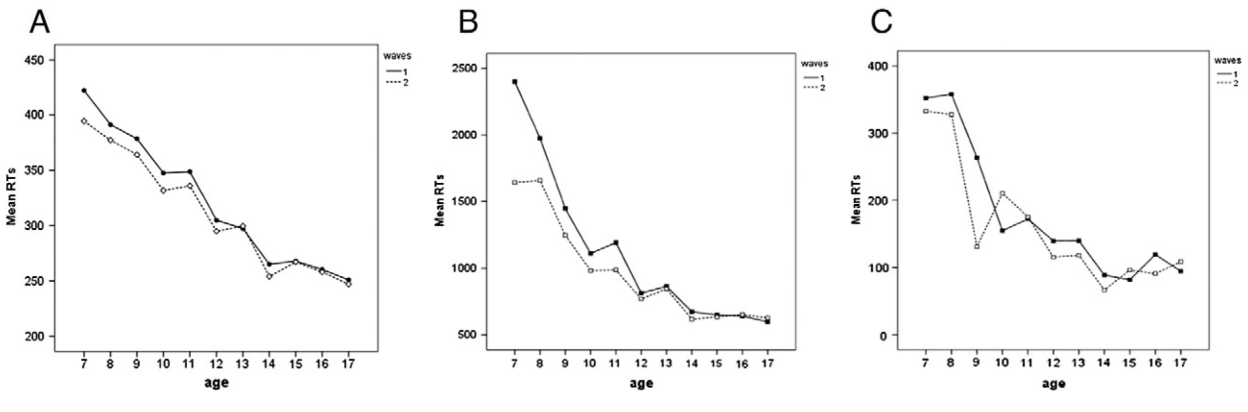

Fig. 1. Speed (panel A), divided attention (panel B), and selective attention ((panel C) as a function of age and testing wave. Note 1. Speed is mean reaction time (in ms) on simple reaction time tasks; divided attention is the difference between mean reaction time to object classification tasks and mean simple reaction time tasks performance simultaneously; selective attention is the difference between mean reaction time to incompatible ink color recognition Stroop tasks and mean simple reaction time to color choice tasks.

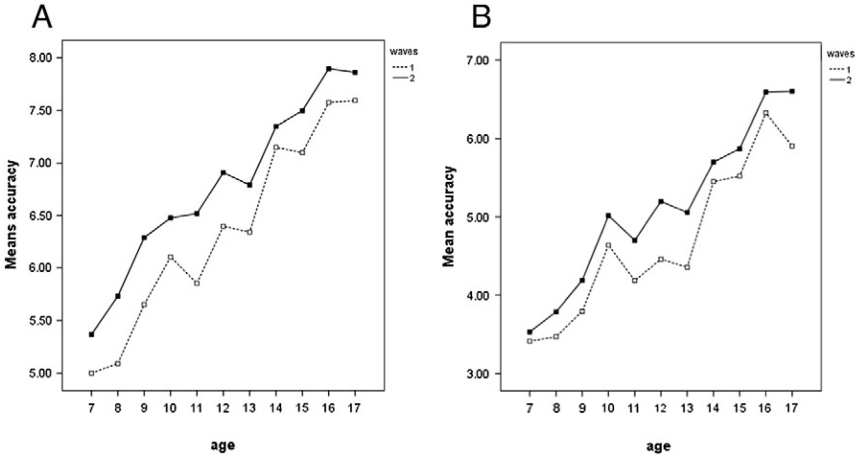

Fig. 2. Forward (panel A) and backward (panel B) digit span as a function of age and testing wave.

throughout the age span from 7 to 18 years studied here. The main effect of task was also highly significant and strong, $F_{1,467} = 1503.29$, $p < .001$, $\eta^2 = .76$, indicating that performance on the FDS task ($M = 6.27$, $SD = 1.35$; $M = 6.75$, $SD = 1.24$) was considerably higher than performance on the BDS task ($M = 4.58$, $SD = 1.42$; $M = 5.03$, $SD = 1.49$), at both the first and the second testing wave, respectively. The main effect of testing wave was also highly significant and moderately strong, $F_{1,467} = 94.28$, $p < .001$, $\eta^2 = .17$, indicating that performance improved significantly from the first to the second testing wave. Of the various interactions, only the age x task interaction was significant, $F_{10,467} = 2.58$, $p < .005$, $\eta^2 = .05$, suggesting that the two tasks were differentially associated with age. This analysis suggested that age differences were larger in the 7 to 10 and the 14-16 years phases rather than in the 11-13 years phase.

2.1.3. Raven matrices and mathematics

To make performance comparable across test batteries and specify attainment with age, we standardized performance on each battery in reference to the four difficulty levels specified above (series E of Raven's SPM was excluded from this standardization to maximize equivalence across batteries). Specifically, the sum score attained on each battery was divided by the number of items addressed to each level in each battery. Thus, scores in each of the three batteries varied from 0 to 4, reflecting dominant level attainment as a function of age and testing wave. These scores were subjected to an 11 (age) x 3 (domains) x 2 (waves) repeated measures ANOVA. The trends captured by this analysis are presented in Fig. 3 (see complete descriptive statistics in Table 4 of supplementary material). It can be seen that the main effect of age was highly significant and strong, $F_{10,\ 467} = 129.46$, $p > .0001$, $\eta^2 > .73$, indicating that performance improved extensively throughout the age span from 7 to 17 years. The effect of testing wave was also highly significant and strong, $F_{1,\ 467} = 233.34$, $p > .0001$, $\eta^2 > .33$, reflecting improvement from first to second testing across all three batteries. The battery effect was also highly significant and strong, $F_{2,\ 466} = 424.17$, $p > .0001$, $\eta^2 > .64$, reflecting the fact that, overall, performance on SPM was higher than performance on arithmetic and this was higher than performance on algebra. However, the significant age x wave, $F_{10,\ 467} = 3.26$, $p > .0001$, $\eta^2 > .06$, age x battery, $F_{20,\ 934} = 424.17$, $p > .0001$, $\eta^2 > .30$, wave x battery, $F_{2,\ 466} = 17.51$, $p > .0001$, $\eta^2 > .07$, and age x wave x battery, $F_{20,\ 934} = 2.02$, $p > .005$, $\eta^2 > .04$, interaction suggested that improvement across age groups and testing waves differed across batteries. Specifically, in SPM level 2 dominated up to the age of 9 years; level 3 was attained at the

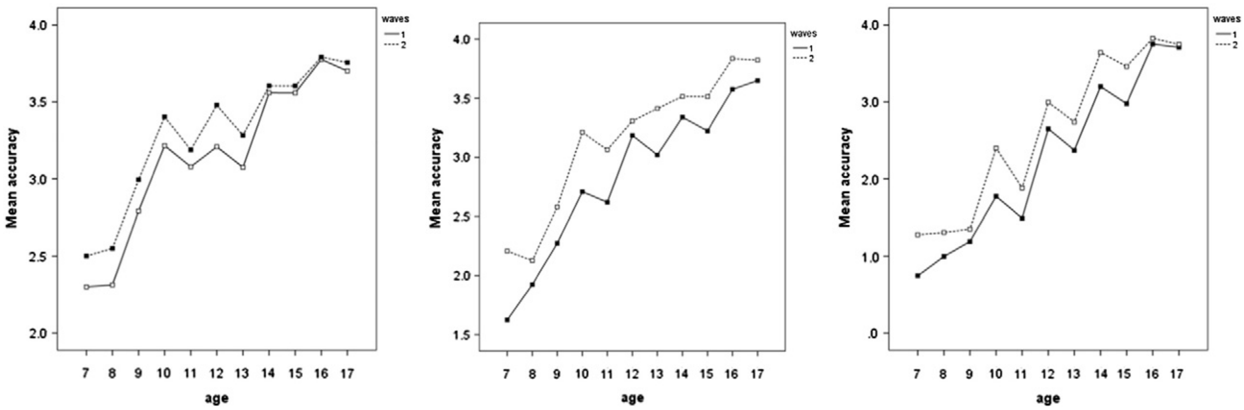

Fig. 3. Mean performance on Raven's SPM, arithmetic operations, and algebraic reasoning. Note 1. Performance was standardized to difficulty level, varying from 0-4.

age of 10 years and stabilized in the 11-14 years phase; level 4 was reached after the age of 14 years, although improvement from first to second testing was minimal in the 14-17 years phase. Performance on the arithmetic test was comparable. Specifically, level 2 dominated from 7 through 10 years. More than 60% of children solved both level 1 and level 2 arithmetic tasks by the age of 9 years but few (less that 40% solved higher level tasks. Level 3 was attained after the age of 10 years (60% of children solved level 3 tasks at the age of 10 years or later but few (less than 30%) solved level 4 tasks at this age. Level 4 tasks were solved by a sufficient number of adolescences only after the age of 13 years. Finally, in algebra, level 1 was attained at the age of 8 years (more than 60% of children solved these tasks), level 2 was reached between 10 and 12 years (at this age 60% or more children solved two of the three level 2 items), level 3 was attained between 12-14 years (only at this age or later 60% of participants solved level 3 tasks) and level 4 was attained after the age of 14 years (only after this age 60% or more of the participants solved the respective tasks).

2.2. Effects of testing experience

In longitudinal studies it may be difficult to dissociate the effects of repeated testing from actual progress reflecting developmental change processes of interest. Anderson, Reid, and Nelson (2001) found that familiarity caused by immediate retesting on inspection time tasks resulted in larger improvement in reaction times than the improvement observed one or two years later in 6-9 years old children. To examine this possibility, the performance attained on the various measures at second testing was contrasted with the performance attained by same age children at first testing. This analysis involved five age groups. That is, the performance attained at second testing by participants who were 8, 10, 12, 14, and 16 years old at first testing (and therefore one year older at second testing) was compared to the performance attained at first testing by 9, 11, 13, 15, and 17 years old participants.

In the case of speeded performance, three mean z scores standing for performance on speed, selective attention, and divided attention were subjected to a 5 (the five age groups) x 2 (two testing times) x 3 (the three measures) repeated measures ANOVA. Nor the main effect of testing time, $F_{1,451} = .00$, $p > .99$, $\eta^2 = .00$, nor any of its interactions with age at testing or process ever approached significance, all Fs < 2.4, all ps > .06, all $\eta^2 < .02$. For working memory, the 2 mean z scores standing for performance on the forward and the backward digit span were subjected to a 5 (age) x 2 (testing time) x 2 (WMC) repeated measures ANOVA. Interestingly, the main effect of testing time was significant, $F_{1,451} = 21.55$, $p < .001$, $\eta^2 = .05$. Of the various interactions only the testing time by age, $F_{4,451} = 2.71$, $p < .03$, $\eta^2 = .02$, was significant. However, this effect was very small (performance on second testing was higher than performance on first testing by only .20 and .28 on forward and backward digit span, respectively). The interaction with age at testing indicated that the improvement from first to second testing was basically limited to 11- and 13-year olds. Finally, the three mean z scores standing for performance Raven matrices, arithmetic, and algebra were subjected to a 5 (age) by 2 (testing time) by three (the three Gf measures) repeated measures ANOVA. The main effect of testing time was not significant, $F_{1,451} = .11$, $p > .74$, $\eta^2 = .00$. However, its interaction with age at testing, $F_{4,451} = 6.35$, $p < .001$, $\eta^2 = .08$, and ability, $F_{2,451} = 4.63$, $p < .01$, $\eta^2 = .02$, was significant. Interestingly, these interactions indicated that performance at second testing on SPM and algebra was slightly better than performance at first testing among only among 11- and 13-year olds.

Therefore, it can be concluded, on the one hand, that the age progress uncovered by the longitudinal part of the study was genuine developmental progress rather than change associated with practice because of repeated testing. On the other hand, however, when present, this effect was very small, concerned with working memory and Gf rather than processing efficiency, and limited to the transition to the principle-based cycle. We will discuss the possible implication of this finding latter on in the discussion.

2.3. Structural relations between processes

To specify the structural relations between processes three types of SEM were run. All of these models involved a factor for each process at each of the two testing waves, as follows: (i) A speed factor (loading on the two simple reaction tasks and the word recognition task), (ii) an attention control factor (loading on the two divided attention and the selective attention difference scores), (iii) a working memory factor (loading on FDS and BDS), (iv) a Raven factor (loading on the five Raven forms), and (v) and mathematical thought factor (loading on arithmetic and algebra). Two alternative versions of a general factor (G_{grand}) may be constructed to test our second prediction. The strongest (and most parsimonious) version would be a single common second-order factor related to the first order factors of both testing waves. This factor would stand for a very powerful common core underlying the various processes both within and across testing waves. This would indicate that this core both participates in the operation of the various processes at successive points in time but also orchestrates their change in time. A weaker version of the general factor would assume that development alters the relations between processes or the role of the common core at different points in time. This version would be consistent with a model where two general factors are assumed, one for each testing wave, with a direct path from the first wave G to the second wave G factor.

These two models were tested in a multiple-groups analysis involving three age groups (i.e., 8-10, N = 159—the 17 7-year-old children were excluded from these analyses because they fall out of the phase boundaries; 11-13, N = 150; and 14-17, N = 152, years participants). Across groups equality constraints were imposed assuming that the task-factor relations were equal across the three age groups. The relations between the first-order factors and the general second-order factor(s) were allowed to vary freely across the three groups, based on the assumption that the strength of the G factor(s) may vary with age.

Although very close, the fit of the model involving one common G, $\chi^2(1029) = 1936.47$, $p = .00$, CFI = .87, RMSEA = .077, model AIC = -121.53), was slightly better than the fit of the model involving one G for each testing wave, $\chi^2(1026) = 1936.48$, $p = .00$, CFI = .87, RMSEA = .077, model AIC = -115.52. For this reason and also in sake of parsimony, the first model was selected for further refinement. Specifically, this model was extended to include age and across waves self-regressions. That is, each first order factor was regressed on age, in addition to its regression on the general factor and the auto-regressions. The fit of this model was significantly better than the first model above and acceptable, $\chi^2(1074) = 1584.43$, $p = .00$, CFI = .93, RMSEA = .056, model AIC = -563.73. This is the model shown in Fig. 4. The values obtained are shown in Table 2 (correlation matrices for these models are shown in Tables 5-8 of supplementary material).

It can be seen (Panel A of Table 2) that the relations between the first-order factors and the general factor varied considerably across age groups. Specifically, in the 8-10 years age group the second-order factor was highly and about equally (> .5) related to working memory, Raven, and mathematics. The relations with speed and attention control, although significant, were moderate (circa -.3). In the 11-13 years age group, the relations were similar to the corresponding relations observed in the 8-10 years age groups, although there was a clear tendency for all relations with WMC, Raven, and mathematics to rise. Finally, in the 14-17 years age group, the second-order general factor was very highly related to the Raven and the mathematics factors (all relations > .7); however, all but one of its relations with the speeded performance dropped below significance.

High autoregressions might reflect one of two different reasons of stability of individual differences in the functioning of a process over time. On the one hand, they might indicate that change depends on a component that is specific to the process itself. Alternatively, it might indicate that the dependence of this process on another one (including G) is stable in time; thus, when the other process changes the process of interest also changes. Measuring both their relations with G and their self-regressions within the same model allows capturing any possible changes in the dependence of the various processes on the common core represented by G. It can be seen (Panel B of Table 2) that autoregressions were significant in all processes but mathematics, suggesting considerable specificity in the change of these processes in time. Interestingly, however, autoregressions of both speed and attention control increased extensively from the 8-10 years phase (.46 and .23 for speed and attention control, respectively) to the phases following (> .7 and > .4, in both phases, respectively), suggesting a tendency for greater autonomy in these processes with age. Autoregressions of WMC and Raven were always significant and relatively high (always > .4). Notably, autoregresssions of mathematics never reached significance, suggesting that change in this domain depends primarily on the general factor.

The relations of the various factors with age (see Panel C of Table 2) are informative by themselves. In the 8-10 years phase, they were all significant and moderate; in the 11-13 phase all but two (attention control at second wave and Raven at both waves) were also significant but generally lower; in the 14-17 years phase the relations of both speed and attention control vanished at both waves. Attention is drawn to the relatively high relations of mathematics with age within all three phases, suggesting that the general factor together with age-depended specific learning experiences are responsible for change in it.

The findings above indicate, in agreement with our third prediction, that G is differentially expressed at different developmental phases. This may indicate that the processes that

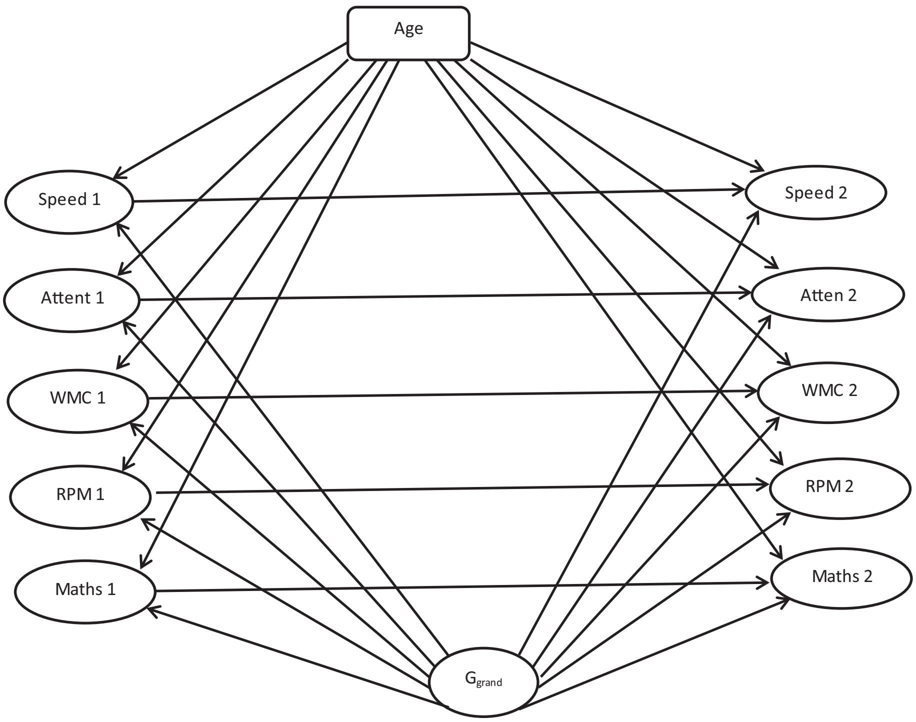

Fig. 4. Prototype model for testing autoregressive relations of the five processes across testing waves and their relations with age and the common general factor (G_{grand}) at both testing waves. Note 1. Results are presented in Table 1.

Table 2
Relations between each process and a factor common (G_{grand}) to the two testing waves (panel A), self-regressions (panel B), and age (panel C) and relations between G at the second testing wave (G_2) and specific processes at the first testing wave (panel D).

	Speed 1	Speed 2	Control 1	Control 2	WM1	WM2	Raven 1	Raven 2	Maths 1	Maths 2
A. G_{grand}										
8-10	-.29	-.28	-.34	-.39	.50	.60	.59	.59	.51	.68
11-13	-.36	-.25	-.31	-.38	.54	.62	.69	.68	.78	.91
14-17	*-.21*	*-.20*	*-.08*	-.35	.27	.37	.75	.81	.74	.85
B. Self-Regressions										
8-10		.46		.23		.51		.48		–
11-13		.73		.44		.44		.56		–
14-17		.74		.41		.56		.35		*.09*
C. Age										
8-10	-.29	-.32	-.40	-.38	.57	.55	.39	.51	.52	.62
11-13	-.44	-.33	-.27	–	.22	.19	–	.13	.47	.41
14-17	–	–	–	–	.24	.26	.31	.32	.62	.46
D. G_2 on processes										
8-10	—		—		1.00		—		—	
11-13	—		-.62		—		—		—	
14-17	—		—		—		1.00		—	

Note 2: Non significant relations are shown in italics.
Note 2: $\chi^2(1074) = 1584.43$, p = .00, CFI = .93, RMSEA = .056 (.050-.062).

contribute to the formation of G differ across developmental phases. To explore this possibility we examined which of the specific process at first testing best predicts G at second testing. In sake of this aim, the G factor at first testing was dropped and the first-order factors were correlated with each other. The second-order G factor at the second testing was preserved and regressed on each of the first-order of the first wave (see Fig. 2). It can be seen (Panel D of Table 2) that in the 8-10 years phase working memory was the only predictor of G (1.00); in the 11-13 years phase attention control was the only predictor (-.62); in the 14-17 years group Raven was the only predictor (1.00). Therefore, the main carrier of effects on G in the 8-10 years phase is working memory; in the 11-13 years phase this role is assumed by control processes involved in attention; finally, in the 14-17 years phase, classic Gf inferential processes (as captured by the Raven test) dominate as the contributors to the formation of G. These findings are partially in line with our third prediction, in that, as expected, working memory and efficiency did appear to be the best predictors of G in the 8-10 and the 11-13 years phases, respectively. However, finding inference as the best predictor of G in the 14-16 years phase is new.

2.3.1. Interactions between processes

Cross-lagged correlations are the method of choice for the specification of interactions between processes across time (Kenny, 1975). This is so because relative differences in the relations between two processes at two testing waves (e.g., Process 1 at Time 1 → Process 2 at Time 2 vs. Process 2 at Time 1 → Process 1 at time 2) can show how processes influence each other in time. For instance, if the first relation is high and the second is low, Process 1 is considered to cause change in Process 2; if the first relation is low and the second is high, Process 2 is considered to cause change in Process 1, but not vice-versa. In the present case, a series of models were ran to test all possible pairs of cross-lagged relations between the five processes involved in the study. These models implement the template model shown in Fig. 5 (correlation matrices for these models are shown in Tables 5-8 of supplementary material). It can be seen that in each run of the model, two processes were taken from the second testing (e.g., Raven and mathematics) and the rest were taken from the first testing (i.e., speed, attention control, and working memory capacity). The two second testing factors were regressed on all three first testing factors. Seven models were run fully exhausting all combinations of pairs of processes at the second testing with trios of processes at the first testing. This approach makes it possible to specify all cross-lagged structural relations between any two abilities at the two testing waves within the context of the total network of relations defined by the set of processes involved in the study. Target processes are often confounded in speeded performance tasks (Stankov & Roberts, 1997). All models fit the data well (all CFIs = .91-.94; all χ^2/df < 2; all RMSEAs < .08).

The cross-lagged relations obtained from these models are shown in Table 3. These suggest that there is a hierarchy of cognitive organization which involves three levels: (i) Processing efficiency, defined by speed and attention control; (ii) representational efficiency, defined by attention control and WMC; (iii) inference, defined Raven and mathematics. Factor-pairs at each level are so tightly connected both within and across testing waves that they obviously represent the same fundamental processes. Adjacent levels in this hierarchy overlap partially as they are interleaved by elements differentially integrated into the fundamental process of each level.

In concern to the processing efficiency level, it can be seen that, in the two younger age phases, there were significant effects from first testing to second testing regardless of which of the two constructs was taken as the predictor. It is noted that in the 8-10 years phase the effect of attention control at first testing on speed at second testing (.50) was considerably higher than the effect of speed at first testing on attention control at second testing (.18). In the 11-13 years phase the relations were very similar (.37 and .35, respectively). These were dropped in the 14-17 years phase, because they were

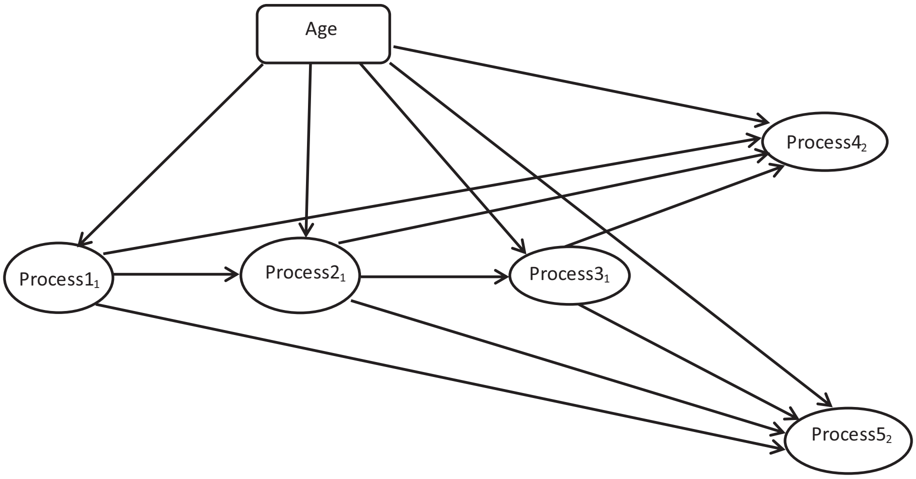

Fig. 5. The prototype model for testing cross-lagged relations between three processes taken from the first testing wave (Process1-3$_1$) and two processes taken from second testing wave (Process4-5$_2$). Note 1. All possible models needed to obtain all cross-lagged correlations between pairs of processes measured at the two waves were tested. Results are presented in Table 2.

very low and non significant. It is noted that the relations between these two factors and the factors representing the other levels, although scarce, always ran top-down. Specifically, WMC at first testing wave did predict speed and attention control at second testing in the 8-10 (-.46 and -.61, respectively) and the 11-13 years phase (-.31 and -.36, respectively). These relations dropped below significance in the 14-17 years phase. The effect of speed and attention control at first testing on working memory at second testing never approached significance. Notably, Raven at first testing exerted an effect on speed (-.25) and attention control at second testing (-.31) in the 14-17 years phase. Also, mathematics at first testing exerted an effect on speed at second testing in the 11-13 years phase (-.42).

The second level geared on representational processes involved in WMC and the inferential processes involved in Raven. It is interesting that in three age phases working memory strongly predicted Raven with increasingstrength (-.41, -.65, and -.85, respectively) but Raven never predicted working memory. However, there were mutual and strong effects between WMC and mathematics in all three age phases, although the effects of WMC at first testing on mathematics at second testing, in all three age phases (.85, .86, and 1.00, respectively), were higher than the effects of mathematics from first testing on working memory from second testing (.75, .75, and .44, respectively).

Raven and mathematics marked the inferential level, exerting effects on each other with increasing strength: The

Table 3
Cross-lagged structural relations between processes across the two testing waves in each age group and auto-regressions (bold in the diagonal).

Processes	Speed	Attent$_2$	Atten$_1$	WM$_2$	WM1	Raven$_2$	Raven$_1$	Maths$_2$	Maths$_1$
Speed$_{1,2}$									
8-10	**.76**	.18	.50	–	-.46	–	–	–	–
11-13	**.85**	.35	.37	–	-.31	–	–	*-.19*	-.42
14-16	**.80**	–	–	–	–	–	-.25	–	–
Attent$_{1,2}$									
8-10		**.86**		–	-.61	*-.12*	–	*.12*	–
11-13		**.87**		–	-.36	–	–	–	–
14-16		**.84**		–	–	–	-.31	–	–
WM$_{1,2}$									
8-10				**1.00**		.41	–	.85	.75
11-13				**.72**		.65	–	.86	.75
14-16				**.79**		.84	–	1.00	.44
Raven$_{1,2}$									
8-10						**.87**		.33	.38
11-13						**.62**		.44	.70
14-16						**.99**		.84	.75
Maths									
8-10								**.83**	
11-13								**1.00**	
14-16								**1.00**	

Note 1: Horizontally, the first value of each process pair shows the structural relation between this process at the second testing and the process shown in the row at the first testing. The second value shows the structural relation between the second testing of the process specified in the raw and this process at the first testing.

effects of Raven at first testing on mathematics at second testing (.33, .44, and .84) and the effects of mathematics at first testing on Raven at second testing (.38, .70, and .75) increased systematically from phase to phase.

In conclusion, WMC appeared to exert strong effects both ways: Top-down, it influenced both speed and attention control in the two younger age phases. Bottom-up, it influenced Raven and interacted on a par with mathematics with increasing strength across the three age phases. Along the same trend, Raven and mathematics intertwined with increasing strength with age. These trends suggest that, with development, control processes are integrated into higher level inferential processes. We will further explicate these trends in the discussion.

2.3.2. Differentiating developmental levels from age phases

The analyses presented so far identified developmental phases with specific age periods, associating each one of them with a specific cognitive profile. Individual differences in developmental rate, however, might spoil this relation. That is, an age period as such may include both fast and slow developers in addition to individuals developing at the expected rate. This state of affairs would confound the relations between the demands of the developmental levels associated with this age period with effects possibly coming from fast or slow development as such. To specify the relations between processing efficiency, WMC, and Gf developmental levels as such, we created level-specific scores for each set of the reasoning batteries used here. Specifically, we scored separately performance on each of the five Raven forms and each of the four levels in each of the two mathematical reasoning batteries. These scores were then transformed into z scores which were averaged across levels (to keep the equivalence of four levels across the three batteries, Series E of the Raven matrices was not included; in fact, performance on this Series was much lower than performance on of level 4 items in the other batteries). These yielded four level-specific mean z scores, one for each level.

Therefore, in principle, each level z score is a powerful measure of Gf, indexing developmentally marked zones of inferential complexity. Assuming that all levels are equally good indicators of Gf implies that all level scores should relate to a common factor rather than to level factors specific to each level. If true, this second possibility would imply that level scores reflect skills or strategies required for handling difficulty, rather than Gf as such. To examine these assumptions, the 12 level specific scores were subjected to a series of confirmatory factor analysis with nested factors.

The first analysis involved only a common factor related to all 12 scores. Expectedly, the fit of this model was poor, $\chi^2(55) = 901.50$, $p = .00$, CFI = .76, RMSEA = .18, model AIC = -791.50, but all score-factor loadings were significant and all but two of them (arithmetic level 1 and SPM level 2) were very high (> .6). A second analysis involved this common factor and also four level specific factors associated with each of the four levels across the domains. The fit of this model, although still poor, $\chi^2(42) = 638.01$, $p = .00$, CFI = .83, RMSEA = .17, model AIC = 554.01, was significantly better than the fit of the first model, $\Delta\chi^2(13) = 239.49$, $p < .001$. It is notable that this improvement resulted from the fact the loadings of level 3 and level 4 SPM and algebra on their respective factors were significant (circa .3). However, the relations of all 12 scores with the common factor remained very high and basically unchanged. A third model included the common factor, the four level-specific factors, and three domain-specific factors, one for each of the three domains, which was related to all four scores representing a domain. The fit of this model was very good, $\chi^2(30) = 103.61$, $p = .00$, CFI = .98, RMSEA = .07, model AIC = 43.61, and significantly better than the second model, $\Delta\chi^2(13) = 534.39$, $p < .001$. The fit of this model to the performance of the second testing was equally good, $\chi^2(30) = 114.99$, $p = .00$, CFI = .97, RMSEA = .08, model AIC = 54.99. The loading of each level score to each factor at the two testing waves are shown in Table 4. It can be seen that all but one (arithmetic Level 1) of the 12 scores were very highly related to the common factor; SPM and algebra did stand up well as domains but arithmetic was weak; all level specific factors were very weak and barely identifiable. Showing that all levels are closely related to Gf regardless of domain was necessary for the exploration of their relations with processing and representational efficiency to be presented below (correlation matrices for these models are shown in Tables 9-12 of supplementary material).

Having verified their association with Gf, level scores were then subjected to two types of SEM (see Fig. 6; (correlation matrices for these models are shown in Tables 9-12 of supplementary material). One set of models involved a speed, an attention control, and a WMC factor from first testing wave and a level factor from the second testing wave. In each of these models, the second-testing level specific factor was regressed on all three first testing factors. Thus, four models were ran, one for each developmental level. In another set of models the time relation was reversed. That is, these models involved the level scores of the first testing wave and the speed, attention control, and WMC scores of the second testing. Each of these second-testing factors were regressed on the first-testing level-specific factor. On the one hand, the first set of models may show how speed, attention control, and WMC may predict attainment in each developmental level one year later, in each age period. On the other hand, the second set of models may show how attainment at each developmental level may predict processing efficiency one year later. Taken together, the two sets of models can reveal the cross-lagged relations between mental processing and Gf at various levels. All models fit the data well (all CFIs = .91-.95; all $\chi^2/df < 2$; all RMSEAs < .08).

The results of these analyses are presented in Table 5. Regarding the first set of models predicting developmental Gf from mental processing functions, the following relations emerged. First, in line with our interpretation above about an inferentially-based tier in cognitive organization, relations of all four Gf developmental levels and WMC were stronger than their relations with any of the other two measures (i.e., speed and attention control), explaining as much as two or three times more variance. However, there was a clear developmental trend in this relation both level-wise and age-wise. In concern to developmental level, the higher the level the stronger the relation with WMC. This is especially clear if level 4 is compared to the rest. In concern to age, this relation decreased systematically with increasing age (from generally high in the 8-10 years period (from .5 to 1.00) to generally moderate in the 11-13 years period (from .35-.65) and non-significant in all but one case in the 14-17 years period.

Table 4
Nested factors for cognitive level scores abstracted from the performance of the whole sample at first/second testing.

Level scores	Gf	Raven	Arithmetic	Algebra	Level 1	Level 2	Level 3	Level 4
Raven 1	.61/.35	.42/.45			.67/.82			
Raven 2	.62/.36	.62/.68				.02/.63		
Raven 3	.71/.68	.51/.52					.04/.24	
Raven 4	.68/.69	.53/.50						.03/.06
Arith. 1	.20/.13		.05/.13		.11/.12			
Arith. 2	.59/.69		.04/.13			.80/.05		
Arith. 3	.59/.60		.80/.53				.06/.08	
Arith. 4	.73/.73		.14/.16					.10/.02
Algebra 1	.76/.58			-.27/.12	.03/.04			
Algebra 2	.82/.80			.29/.31		-.10/.03		
Algebra 3	.88/.85			.36/.45			.32/.05	
Algebra 4	.73/.63			.39/.37				.56/.69

Interestingly, the effects of Gf on working memory capacity were higher than the corresponding inverse effects, especially in the 14-17 years phase, suggesting that top-down influences of inferential possibilities on working memory are very strong.

The top-down flow of causal influences is especially evident in the case of attention control. It can be seen that attention control did not predict Gf at any level but level 1 at 8-10 years; however, Gf predicted attention control at all levels. The pattern of speed → Gf relations was very different. Gf significantly predicted speed at all levels at the 8-10 years phase, they mutually predicted each other in the 11-13 years phase, but their relation waned in the 14-17 years phase. Interestingly, there was a moderate but significant effect of speed on all levels in the 11-13 years phase. This is consistent with earlier findings that speed is a strong index of change in the 11-13 years phase. These results will be analyzed further in the discussion below.

3. Discussion

Performance differences between age groups and testing waves were generally in line with our first prediction. The sequence of levels in each domain were attained, by and large, in the age phase expected and major shifts in processing efficiency occurred within the age windows expected (e.g., changes in different aspects of control). Specifically, in the 8-10 years phase alignment and integration of well specifiable representations dominated. Raven's Series A and B and levels 1 and 2 of arithmetic operations and algebra were grasped in this phase. This is the phase in which inference becomes fluid enough to access individual representations, align them, and bind them together according to underlying relations. We remind that the longitudinal results showed that there was a rather large change in divided attention from 7 to 8 years and a rather large change in selective attention from 8-9 years, reflecting a large improvement in attention control in the first development phase studied here. However, working memory was the pivotal construct in this phase: It was the sole predictor of changes in G_{grand}, driving the acquisition of its two lower phase-specific levels and it carried strong effects on attention control. Thus, WMC operated as an agent that brings attention control under representation control, especially in the 8-10 years phase.

In the next phase, from 11-13 years, G_{grand} was mainly expressed through level C inferential possibilities as addressed by Raven's Series C matrices and level 3 mathematical reasoning problems. Obviously, these problems require representational alignment that is mastered in the previous phase. In addition, however, they also require explicit encoding of the relations generated by alignment into a representational token

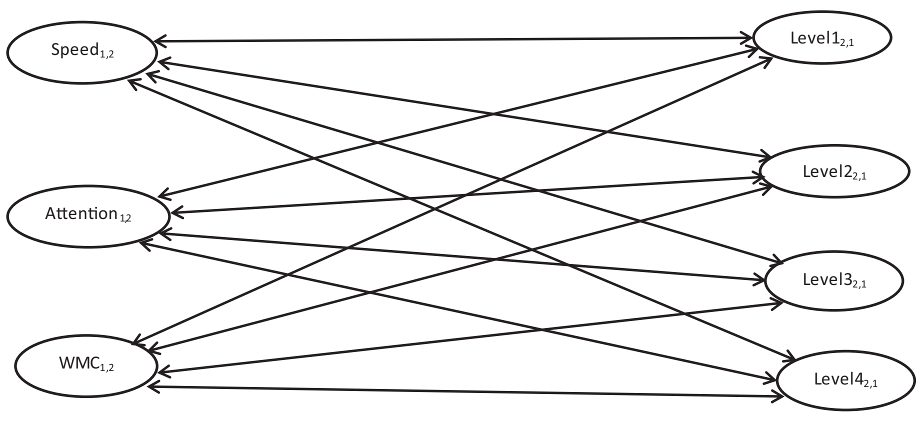

Fig. 6. Prototype model for testing relations between speed, attention control, and working memory capacity at first testing wave and developmental level at second testing (one model for each level) and between developmental level at first testing and speed, attention control , and working memory capacity at second testing (one mode for each level). Note 1. Two-way arrows stand for between factor regressions in each of the two sets of models rather than between factor correlations. Results are presented in Table 3.

Table 5
Structural relations between developmental levels at first or second testing and processing efficiency functions taken from the other testing wave in each age group.

	Cognitive levels at second and first testing wave (2/1)			
Efficiency	Level $1_{2/1}$	Level $2_{2/1}$	Level $3_{2/1}$	Level $4_{2/1}$
Speed$_{1/2}$				
8-10	–/-.46	–/-.50	–/-.44	–/-.46
11-13	-.42/-.44	-28/-.33	-.36/-.35	-.43/-.40
14-16	-*.30*/-.25	-.41/-*.18*	-*.16*/-*.17*	-.34/-.24
Attention$_{1/2}$				
8-10	-.35/-.53	–/-.56	-.18/-.48	–/-.52
11-13	–/–	–/-.16	–/.24	–/-.18
14-16	-*.25*/–	–/-.13	–/-.19	–/-.25
WMC$_{1/2}$				
8-10	.57/.82	.47/.83	.62/.85	1.00/.84
11-13	*.35*/.58	.50/.62	.47/-.64	.65/.61
14-16	–/.38	–/.33	–/.49	.55/.46

Note 1. Mean performance on the tasks of each developmental level across the three domains (SPM, arithmetic operations, and algebra) at second testing were regressed on the three measures of mental processing efficiency measures taken from first testing (first value of each pair) or each of these three efficiency measures at second testing wave were regressed on each developmental level taken from the first testing. Non-significant relations are shown in italics.

of these relations as such. This may be an explicit grasp of the transformation connecting the matrices or the mathematical relation running through a series of mathematical ensembles. It is notable that transition to this phase was associated with a further improvement of selective attention, reflecting an improvement in internally directed stimulus choice and response inhibition. It might be the case that further mental flexibility associated with this change is related to requirements for representational binding posed by principle-based problems. That is, it reflects the mental fluidity required for systematic mental inspection of representations for signs of sought after properties, given the problem goal. It is notable that, on the one hand, change in G_{grand} in this phase was driven by attention control; on the other hand, there was an effect of inferential processes underlying Gf on attention control, especially in its level 3 expression. It is also notable that there was a two-way interaction between inferential domains and working memory. Working memory predicted changes in Raven but mathematics predicted changes in working memory as much as working memory predicted changes in mathematics. This is equivalent to saying that principles in this phase both emerge from but they are also imposed on the data structures available. Interestingly, it was only at this phase the practice effects were observed, probably suggesting that transition to a new cycle raises learning readiness, more than in alignment phases. Finally, in the next phase from 14-18 years these principles may be precisely aligned with each other as suggested by the attainment of the top level in all domains. When this is possible, the inferential processes as such dominate as organizers of change in the phase.

What is then represented by the G_{grand} factor predicted by our second prediction? This factor stands for a common core underlying each phase-specific ensemble of mental possibilities Specifically, in the conceptual alignment phase, from 8-10 years, it allows to systematically track varying stimuli, shift between then, and appropriately align stimuli with response, implement a program of representational re-arrangement covering 3-4 representations as in backward digit span, map 3-4 representations on each other, as in Series B Raven matrices or level 2 arithmetic and algebraic reasoning problems, extract a relation and use it to identify other related representations. In the principles emergence phase following, from 11-13 years, attentional control acquires the mental zooming resolution that is necessary to precisely preselect a process to inhibit (e.g., reading) and another process to implement (e.g., color recognition); the representational re-arrangement program is fluent enough to operate on line during storage and maintenance thereby increasing the capacity of working memory; the inferential process is sharp enough to spot relations and reduce them into representational tokens that may be processed as such. Thus, in the phase following both attention control and working memory operate to the limit (attained in the previous phase). As a result, in this phase, only the inferential processes as such are still improving.

Therefore, it seems that there is a developmental snowball effect in the expansion of the G_{grand} core. That is, there is a functional upgrading of this core in each phase such that new found processes in each next phase sit on the processes acquired in the previous phase and get integrated with them into a smoothly running whole. For instance, the flexibility in stimulus tracking and stimulus-action alignment initiating the 8-10 years phase is embedded into the representational executive program enabling the organization and handling of information in working memory. This may also drive the alignment and binding of representations in problems such as scanning matrices in order to decipher their relation or properly arranging problem solving steps in the various mathematical problems. In the next phase mental attention control takes over as the G_{grant} predictor. In this phase, however, it is more inwardly oriented, enabling zooming on and comparison of representations with each other. It seems that this form of attentional control enables the specification of commonalities of representations and their reduction into a representational token that may be mentally handled as such. This seems to be a prerequisite of inferential control that dominates in the next phase.

The interactions between efficiency measures with developmental levels shed light on how this scaffolding role is implemented in each phase. Specifically, early in each developmental cycle, attention control is important for the construction of representational and inferential processes taking place in the cycle. Working memory is also important for these constructions, especially in the alignment phases when relations between the representations and processes constructed earlier are worked out. However, their importance declines with developmental level (and age to the extent that age is generally associated with developmental level) because specific knowledge and strategies become increasingly important for the construction and smooth functioning of higher levels. In line with earlier findings that speed indexes rather than causes changes in reconceptualization (Demetriou et al., 2013; Demetriou, Spanoudis, & Shayer, 2014; Demetriou, Spanoudis, Shayer, van der Ven, et al., 2014), this study showed longitudinally that reconceptualization predicted speed rather than the other way around, especially in earlier phases and/or developmental levels. This direction of causality demonstrated nicely that ascension along inferential levels tightens mental processing up, making it more efficient. The relation between developmental levels and efficiency indexes, such as speed, weaken at the top of the developmental hierarchy because top levels require much more than processing efficiency. A rich knowledge base and refined operational plans are probably more important than efficient control and working memory at these levels.

These patterns are in line with the model of developing control recently proposed by Demetriou, Spanoudis, & Shayer (2014), Demetriou, Spanoudis, Shayer, van der Ven, et al. (2014). According to this model, in the preschool years, at 3-4 years, executive control is expressed as a "focus-choose-respond" program allowing toddlers to alternate between representations vis-à-vis related actions, as in the go/no go tasks. Later, in primary school, this program is upgraded into a "scan-select-search-shift" program allowing children to perform tasks requiring simultaneous focusing on two stimuli and respond in sequence to them as needed, as in divided attention tasks. In adolescence, this program is extended into *an inferential relevance mastery program.* This allows adolescents, at 12-13 years, to systematically search their representational spaces and choose between representations according to relevance to a goal. In inference, it allows choices between reasoning and/or heuristic processes according to the specificities of the problem at hand and to evaluate relative truth and validity. It is noted that these programs function like scaffolds in building construction: They are necessary for as long as the construction goes on; when the building is complete they can be removed without jeopardizing the stability of the building. This is the meaning of the fact that attention control fully disengaged from inferential processes in the 14-18 years phase and their relations with working memory waned drastically.

Interestingly, Kiyonaga and Egner (2014) showed recently that working memory operates on conflicting information mimicking the Stroop effect in the same way that perception operates on conflicting information in the classical Stroop task. Along the same line, Shimi, Nobre, Astle, and Scerif (2014) showed that improvement of visual working memory from 7 years to adulthood was related to changes in controlled voluntary orienting to the appropriate stimuli. These exert a beneficial effect on the maintenance of information in working memory. These findings seem to have two important implications. First, working memory and attention rely on the same resources and operate over the same representations. Second, the strong top-down effects voluntarily sharpen the focus and deployment of control processes. By definition, awareness and insight into the nature of mental representations is part of this process. Spanoudis et al. (2015) showed recently that indeed awareness of mental representations intervenes between attention and working memory.

In conclusion, this study bears some important implications for developmental and psychometric theories of intelligence. In concern to developmental theories, age boundaries marking transitions in mental organization and functioning are well preserved (i.e., 8, 11, and 14 years in the age span studied here). Therefore, these boundaries may reflect important dynamics in brain and organismic development (Demetriou, Spanoudis, Shayer, 2014; Thatcher, 1994). Each phase involves an executive core embedded into phase-specific programs for accessing, matching, and inter-relating representations. Obviously, meaning and problem solving in each phase depend on learning the specific codes (e.g., logical, heuristic, problem specific, etc.) for what is valid and acceptable in inter-relating representations. Also, there is no single line of causal effects in the assembly of these programs. There are both bottom-up and top-down causal effects, depending upon the phase concerned. Therefore, theories assuming a single chain of effects from simpler to more complex processes during development (e.g., Case, 1985; Halford et al., 1998; Kail, 2007) only grossly captured a part of the causal forces driving intellectual development. Likewise, more global theories (e.g., Bruner, 1966; Piaget, 1970) grossly captured some of the programs surrounding the core in each phase, ignoring the constraints and possibilities afforded by this core.

In concern to psychometric theories, general processes do exist however they are called (e.g., Carroll, 1993; Jensen, 1998; Spearman, 1927), but their profiles varies with developmental phase or the level of competence reached. The discussion about changes in G_{grand} implies that the core of general processes runs through AACog, that is close to the classic psychometric concept of fluid intelligence, processing and representational efficiency processes captured by speed, attention control, and working memory tasks, and self-awareness and self-regulation possibilities not examined here. Obviously, this analysis requires further substantiation at a number of levels. For example, psychological research may examine if systematically varying a particular parameter of G_{grant} at a particular developmental phase would change the other parameters according to the patterns above. Brain research would have to examine if changes in these patterns are reflected in changes in specific neural circuitry. For instance, it may thus be the case that transitions across cycles relate to the establishment of the brain networks which are needed to project representational alignments of an earlier cycle into the more abstract networks capturing the new units that emerged from these alignments. Changes in the connectivity of domain-specific circuitry with general purpose circuitry in the prefrontal cortex may be of this kind. Transitions within cycles may relate to changes in the relations between more localized networks or in the relations between these networks and regions dedicated to working memory as such.

The discussion above bears some practical implications as well. For example, our findings suggest that individual differences in practically important aspects of intelligence, such as learning or dealing with new problems, varies extensively depending upon the cycle or phase concerned. In some phases, processing efficiency and control are very important as predictors; in others, representational or inferential capacity take prominence. Therefore, our diagnostic tools, such as tests of intelligence or intellectual development need to adjust so that they can capture these variations. Obviously, we still we have a long way to go before mapping these new terrains.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.intell.2015.01.004.

References

Anderson, M., Reid, C., & Nelson, J. (2001). Developmental changes in inspection time: What a difference a year makes. *Intelligence*, *29*, 475–486.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Bruner, J. (1966). *Studies in cognitive growth.* New York: Wiley.

Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence*, *47*, 34–43.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Case, R. (1985). *Intellectual development: Birth to adulthood.* New York: Academic Press.

Cowan, N., Morey, C. C., Chen, Z., & Bunting, M. F. (2007). What Do Estimates of Working Memory Capacity Tell Us? In N. Osaka, R. Logie, & M. D'Esposito (Eds.), *Working Memory: Behavioural & Neural Correlates* (pp. 43–58). Oxford, U.K.: Oxford University Press.

Coyle, T. R., Pillow, D. R., Snyder, A. C., & Kochunov, P. (2011). Processing speed mediates the development of general intelligence (g) in adolescence. *Psychological Science*, *22*, 1265–1269.

Demetriou, A., Christou, C., Spanoudis, G., & Platsidou, M. (2002). The development of mental processing: Efficiency, working memory, and thinking. *Monographs of the Society for Research in Child Development*, *67* (Serial No. 268).

Demetriou, A., Pachaury, A., Metallidou, Y., & Kazi, S. (1996). Universal and specificities in the structure and development of quantitative-relational thought: A cross-cultural study in Greece and India. *International Journal of Behavioral Development*, *19*, 255–290.

Demetriou, A., Spanoudis, G., & Shayer, M. (2014). Inference, reconceptualization, insight, and efficiency along intellectual growth: A general theory. *Enfance*(issue 3), 365–396. http://dx.doi.org/10.4074/S0013754514003097.

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 mind. *Intelligence*, *41*, 34–50.

Demetriou, A., Spanoudis, G., Shayer, M., van der Ven, S., Brydges, C. R., Kroesbergen, E., et al. (2014). Relations between speed, working memory, and intelligence from preschool to adulthood: Structural equation modeling of 14 studies. *Intelligence*, *46*, 107–121.

Diamond, A. (2013). Executive function. *Annual Review of Psychology*, *64*, 135–168.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology. General*, *128*, 309–331.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, *87*, 477–531.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *The Behavioral and Brain Sciences*, *21*, 803–864.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences.* Amsterdam: Elsevier.

Johnson-Laird (2013). Mental models and cognitive change. *Journal of Cognitive Psychology*, *25*, 131–138. http://dx.doi.org/10.1080/20445911.2012.759935.

Kail, R. V. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin*, *109*, 490–501.

Kail, R. V. (2007). Longitudinal evidence that increases in processing speed and working memory enhance children's reasoning. *Psychological Science*, *18*, 312–313.

Kail, R. V., & Ferrer, E. F. (2007). Processing speed in childhood and adolescence: Longitudinal models for examining developmental change. *Child Development*, *78*, 1760–1770.

Kenny, D. A. (1975). Cross-Lagged panel correlation: A test for spuriousness. *Psychological Bulletin*, *82*, 887–903.

Kiyonaga, A., & Egner, T. (2014). The working memory Stroop effect: When internal representations clash with external stimuli. *Psychological Science*, *25*, 1619–1629.

Kyllonen, P., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, *14*, 389–433.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.

Mickey, R. M., Dunn, O. J., & Clark, V. (2004). *Applied statistics: Analysis of variance and regression.* Hoboken NJ: Wiley-Interscience.

Nettelbeck, T., & Burns, N. R. (2010). Processing speed, working memory and reasoning ability from childhood to old age. *Personality and Individual Differences*, *48*, 379–384.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica*, *63*, 301–345.

Pascual-Leone, J., & Johnson, J. (2011). A developmental theory of mental attention: its applications to measurement and task analysis. In P. Barrouillet, & V. Gaillard (Eds.), *Cognitive development and working memory: A dialogue between neo-Piagetian and cognitive approaches* (pp. 13–46). New York: Psychology Press.

Pashler, H. E. (1998). *The Psychology of attention.* Massachussetts: A Beadford Book. MIT Press.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's handbook of child development* (pp. 703–732). New York: Wiley.

Rips, L. J. (2001). Two kinds of reasoning. *Psychological Science*, *12*, 129–134.

Shimi, a., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting attention within visual short-term memory: Development and mechanisms. *Child Development*, *85*, 578–592.

Spanoudis, G., Demetriou, A., Kazi, S., Giorgalla, K., & Zenonos, V. (2015s). Embedding cognizance in intellectual development. *Journal of Experimental Child Psychology*, *132*, 32–50.

Spearman, C. (1927). *The abilities of man.* London: MacMillan.

Stankov, L., & Roberts, R. (1997). Mental speed is not the 'basic' process of intelligence. *Personality and Individual Differences*, *22*, 69–84.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654.

Thatcher, R. W. (1994). Cyclic cortical reorganization: origins of human cognitive development. In G. Dawson, & K. Fischer (Eds.), *Human Behavior and the Developing Brain* (pp. 232–266). New York: Guilford.

van der Ven, S. H. G., Kroesbergen, E. H., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*, *82*, 100–119.

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., et al. (2003). The Development of Executive Function in Early Childhood. *Monographs of the Society for Research in Child Development*, *68* (Serial No. 3).