Psychological Test and Assessment Modeling, Volume 56, 2014 (3), 237-254

Measured attention in prolonged over-learned response tasks and its correlation to high level scientific reasoning and school achievement

R. Hotulainen[1], H. Thuneberg[2], J. Hautamäki[2], M.-P. Vainikainen[2]

Abstract

The relationship between attention and academic performance has been of interest starting with early studies on academic success and failure. In this study we examine how attention measured in simple and prolonged over-learned response tasks correlates with and contributes to scientific reasoning and school achievement (GPA). To study attention, the Attention Concentration Test (ACT) was used and to study scientific reasoning, a modified version of Science Reasoning Tasks, tapping control-of-variable schemata, was used. Of special interest was the highest performing attention group (+ 1 *SD*) formed from the ACT results. We gathered our data from Finnish ninth graders (*n* = 358; including 166 girls) from the city located in the eastern part of Finland. Statistically analysed results showed that attention contributed to scientific reasoning, which in turn explained the largest share of the GPA variance. The highest attention group differed from the lowest two attention groups in GPA and from all attention groups in scientific reasoning. For educational practitioners the ACT seems to be a useful tool in assessing exceptional academic learning potential in students.

Keywords: attention, prolonged over-learned working task, scientific reasoning, school achievement, high ability

---

[1] *Correspondence concerning this article should be addressed to:* Risto Hotulainen, PhD, Educational Assessment, University of Helsinki, P.O. Box 9 (Siltavuorenpenger 3A), FI-00014 Helsinki, Finland; email: Risto.Hotulainen@Helsinki.fi

[2] Educational Assessment, University of Helsinki, Finland

Cognitive abilities invite a convergence of pedagogical and developmental research (Adey, Csapó, Demetriou, Hautamäki & Shayer, 2007; Gottfredson, 2002; Demetriou, Spanoudis & Mouyi, 2011; Jensen, 1998; Nisbett, Aronson, Blair, Dickens, Flynn, Halpern & Turkheimer, 2012). In such studies the predicted factors have generally been school achievement, matriculation into selective institutions of higher education, or work related outcomes. On one hand, abilities are used to predict relevant outcomes using one index – intelligence, g-factor, fluid intelligence (gF), or reasoning – or several measures of some of the plenty lower-strata broad and specialized cognitive factors (Carroll, 1993). On the other hand, cognitive abilities have occupied the position of the dependent variable to be understood or explained (Jung & Haier, 2008). In the latter case, the explanatory variables have been, for example, genetic and neural mechanisms (Garlick, 2002; Geary, 2005), attention (Cowan, 2000) and working memory (Blair, 2006). Specifically, there are thousands of studies to be found, for each of the presented notions in their various forms. However, there are only few comprehensive models intended to make an integration of concepts dealing with neural and brain connections with mediating factors like attention and working memory or with concepts dealing with educational outcomes (Gustafsson, 2002; Demetriou et al., 2011).

A useful hypothesis is given by Gray, Chabris & Braver (2003, p. 316). They make explicit the connection between fluid intelligence and attentional control. Heitz, Unsworth and Engle (2004) are on the same line – in relation to Raven – by asking how much of gF can be attributed for the attention.

Given all the above, the study of this integration is important (Adey et al., 2007) and we intend to support this endeavor with specifications of the relations of three domains: attention, reasoning and achievement. Reasoning is studied as an independent factor in relation to school achievement as a mediating factor. In that model, it is possible to propose also the time-order of processes, from attention to reasoning and finally to school marks. In real human development, time-order is wrapped in feedback loops to the extent that it is possible to claim that schooling or cultural institutions cultivate higher mental functions (Adey et al., 2007; Bruner, 1986; Ceci, 1991; Cole, 1996; Luria, 1976; Olson, 2003; Vygotski, 1978).

Attention

In the study, attention is understood as a fundamental attentional capacity (Smit & van der Ven, 1995). Contained within this definition are assumptions such as that the ability to reason and the ability to perform well at school are dependent on the ability to concentrate (attention), and that attention is present in all cognitive tasks requiring mental effort. Attention is considered to be crucial for learning and it is also claimed to be a central factor of intelligence (Baillargeon, Pascual-Leone & Roncadin, 1998; Spearman, 1927; van der Ven, 2001). Still, there is no general agreement how attention should be defined and measured. Logan (2004), for example, emphasizes that paradigmatic consistency and task shifting in attention research has produced substantial cumulative progress in our theoretical understanding of attention as a measurable phenomenon.

Main approaches using reaction times can be roughly divided into two different test traditions: a) speed tests and b) concentration tests (Smit & van der Ven, 1995). Speed tests contain tasks with varying difficulty that can be solved if enough time is given. In the case of concentration tests similar but relatively easy test items have been used, but registration of the reaction times has been different, for example, a subject's time per test item and subject's time of the groupings of test items have been used (Jensen, 1982; Smit & van der Ven, 1995). When trying to capture pure measure of subject's attention the earlier form of testing and related results are easily distracted by the lack of knowledge, the lack of experience and inappropriate mental sets, which indeed, produce qualitative properties of mental ability instead of the needed quantitative one (van der Ven, 2001). This can be avoided when easy and overlearned response tasks are used.

The Inhibition Theory (IT) of Ad van der Ven (2001; van der Ven, Smit & Jansen, 1989) is a useful theoretical model to explain how reaction times fluctuate in prolonged over-learned response tasks, and how this is related to cognitive performance such as intelligence (van der Ven, Gremmen & Smit, 2005). The IT is based on the basic assumption that during the performance of any mental task requiring a minimum of mental effort, the subject actually goes through a series of alternating latent states of distraction (non-work *0*) and attention (work *1*) which cannot be observed and are completely imperceptible to the subject (Smit & van der Ven, 1995). In easy prolonged over-learned response based tasks a subject takes control of his or her progress through given test items and when accordingly measured produces indicator of his or her mental ability (Larson, & Alderton, 1990; Klauer, 1993; Smit & van der Ven, 1995; Spearman, 1927; Steinborn, Flehmig, Westhoff, & Langer, 2008). The Attention Concentration Test used in this study is based on IT theory (van der Ven, 2013).

Finally, it is important to underscore that inhibition theory and related models transcend recent cognitive psychological approaches which describe attention as a cognitive function often categorized for several sub-abilities such as focused, sustained, selective, alternating and divided attention (e.g., Logan, 2004) and also both from the cognitive inhibition (e.g., Houdé, 2000) and the ADHD approaches in which, for example, behavioral inhibition is closely linked to sustained attention and executive functions (e.g., Baddeley, 1986).

School achievement

School achievement is affected by many variables, such as subjects' characteristics, both cognitive and non-cognitive (personality traits, self-perceptions etc.), classroom practices (teacher-student interaction), and contextual variables (home and community contexts). According to the latest studies (Lu, Weber, Spinath & Shi, 2011), the strongest effect is attributed to the cognitive abilities (Gottfredson, 2002; Freudenthaler, Spinath, & Neubauer, 2008) even though contradictory findings have been reported, as well (cf. Spinath, Spinath, Harlaar & Plomin, 2006). Close to our study, Steinborn et al. (2008) reported how self-paced performance, in particular response speed variability, contributes to school achievement. Their study provided support that reaction time coefficient of variation measured via easy serial mental addition and comparison tasks can be used as a predictor of the school achievement, especially, mathematics performance. In Finland, as in other western countries, female students usually perform better when school grades are used as an indicator of school achievement (Hautamäki, Kupiainen, Marjanen, Vainikainen & Hotulainen, 2013). Gender differences emerge especially in language related subjects for the girls and in math-related subjects for the boys, and corresponding self-reported ability perceptions show greater difference than the actual performance level does. The recent PISA results show that math performance is actually higher among girls in Finland as well (Hautamäki et al. 2008; Kupari, Välijärvi, Andersson, Arffman, Nissinen, Puhakka &Vettenranta, 2013). In this study we restrict our interest for the two cognitive components that contribute to school achievement, namely scientific reasoning and attention.

Scientific reasoning

Developmental psychologists have long studied scientific reasoning. One of the most studied constructs in this tradition is the control of variables strategy, which was first introduced by Inhelder and Piaget (1958) as part of the formal operational thinking construct (see Shayer, 1979). Later researchers (see Neimark, 1975; Kuhn, 2002; Shayer, 2008) showed that the control of variables strategy is central to science and an essential skill that is attainable and trainable by the time and well-structured interventions when children are cognitively advancing toward a formal operational level (Adey et al., 2007; Kuhn, 2008; Shayer, 2008).

In this study, scientific reasoning was conceptualized as formal operational thinking, and more specifically on control-of-variable schemata (Scientific reasoning). Regardless of the recent partly controversial developmental views on scientific reasoning and its correspondence to the Piagetian thinking stages (Kuhn 2002; Kuhn, Iordanou, Pease & Wirkala, 2008), in Finland experimental studies have shown that assessment of scientific reasoning shares properties that correspond to the earlier findings concerning formal operation percentage shares within target population (Hautamäki, 1989; Kuusela, 2002). For this reason formal operations were considered to be a valid choice for a study variable. Moreover, the Piaget's thinking stages have shown to correspond with findings of modern neuroscience, as brain scan methods have produced results of neural maturation that parallel the stages (Emick & Welsh, 2005; Shute & Huertas, 1990).

The emergence of formal operations at around 12-15 years of age involves reasoning based on hypotheses, independent of concrete objects. The age variation is considered to be rooted in different intellectual stimuli in children's environments and to be dependent on personal interest and quality and amount of experiences. On the other words, formal thinking is not necessarily applied all the time or across all domains. (Piaget, 1972; 2006). Following Inhelder and Piaget (1958), Shayer and Wylam (1978) reported a survey of 10,000 10- to 16 year-olds, showing that in a class of 12 year-olds (Adey et al., 2007) mental ages ranged from that of average 6 year-old to an above-average 18 year-old. Hautamäki (1989; 2000) in a replication of the Shayer and Wylam (1978) survey reported identical results in Finland, finding that less than one third of 15-year-olds reached at least the earliest formal operational level and only five percent have reached solid level in their thinking as in England.

Here we use scientific reasoning as an indicator of scientific thinking when studying parallel effects. Regardless of the recent controversial findings concerning how control of variable strategy develops and how it can be supported in school setting (Kuhn, 2002; Kuhn, Iordanou, Pease & Wirkala, 2008) one of the goals of this study is to produce additional information concerning prerequisites of the high level scientific thinking.

Purpose of the study

With this study we aim to fill one of the gaps in the literature on the scientific reasoning and school achievement of high able students, that is, the impact of their attentional capacity measured by prolonged over-learned response tasks, on their academic performance. As student's gender is strongly shown to be bound to the school achievement, the effect of gender is studied in the line of study hypotheses.

The following hypotheses resulted from the literature review and our study interest:

– *Hypothesis 1:* Measured attention in prolonged over-learned response task predicts school achievement. Attention also explains performance in scientific reasoning tasks, and the effect of attention on school achievement is partially mediated by performance in scientific reasoning tasks.
– *Hypothesis 2:* Highly able students identified with the measured attention in prolonged overlearned response tasks have both higher scientific reasoning skills and higher school achievement than students who do not show such high attentional capacity.

When classifying some students to belong among highly able we involuntary form other groups having not that high but lower attention ability. For this reason we are simultaneously interested in studying if the lowest performing group differs from the groups having higher attention. This is of interest to gain approximation of the ecological validity of the ACT as an attention screening measure.

Method

Participants

The sample consisted of an entire age cohort of ninth graders attending regular nine-grade comprehensive education in a small town situated in the east of Finland, $N = 358$, 52% male, average age of 16 years (220 months, $SD = 4$ months). All participating schools represent a mixed-ability system having neither entry nor selection criteria.

Instruments

The Formal Operations Test.

This test measures scientific reasoning. Content-wise, the test measures the mastery of control of variables, scientific reasoning schemata. The test is a second-generation modified group-version (Hautamäki, 1989, 2000) of the original Scientific Reasoning Tasks, the Pendulum (Shayer, 1979; Shayer & Wylam, 1978). Formula 1 has been widely used in Finnish studies and the used format (race cars) has shown no gender differences (Hautamäki et al., 2013). In Formula 1, the subjects compared F1 drivers, cars, tires, and race tracks (four variables, all with two given values from which to select: Räikkönen, Schumacher; Ferrari, McLaren; Michelin, Bridgestone; Monaco, Silverstone). Subjects were given a set of values for the four variables (such as Räikkönen, Ferrari, Michelin, Monaco) and asked to construct another set that would clarify the role of a specified variable, for example, tires. Subjects were asked to produce a set of values for all four variables so that the focal variable was studied with an unconfounded pair (see Strand-Cary & Klahr, 2008). In half of the items, the subjects were given a dual set (Räikkönen, McLaren, Michelin, Monaco vs. Räikkönen, Ferrari, Michelin, Monaco) and asked if this was a good test of, to continue the example, the role of tires (in this case, the question was confounded for tires, unconfounded for the nonfocal variable car). Subjects chose their selection from a given set of values for all variables, or answered by checking whether the pair is a good test of driver, car, tires, race (*yes, I do not know*; *no, I do not know* always coded for 0). The nine items were coded dichotomously and summarized into the Scientific Reasoning test score. The reliability of the test was .84. Half the class took the paper-and-pencil formal operation test while the other half performed attention tasks in the computer lab. The groups switched after 30 min. The subjects were encouraged to do their best and to spend the time needed to find solutions on both tests. Neither the time of day nor the order of the tests affected the results.

School achievement.

The Grade point average (GPA) was computed as the mean of nine school subjects, obtained from the school register and common to all study participants. A GPA value of 4 is a failing grade, while 10 is the best grade for any subject, corresponding to A+ in the United States. The reliability of GPA was .86.

The Attention Concentration Test (ACT).

The ACT primarily measures attention or more specifically the level of concentration as measured by variation of series of reaction times. The test is based on the following three assumptions: 1) knowledge should not play a part in the final test score; only simple problems should be used; 2) differences in previous experience with the task should not be allowed, and 3) temporal feelings should not play a part; multiple tries should not be allowed to make the best results.

The ACT has the following options:

Test-length: Beginner (25 bars), Intermediate (35 bars), Expert (45 bars)

Difficulty: Easy (one target), Medium (two targets), Difficult (three targets)

Task: Colors, Positions, Dice

Instruction: Quickly vs. Relax

Stimulus-number: Fixed vs. Random

The first one of each category is the default options. The students performed the test with as Task: Positions, with as Stimulus-number: Random, restricted and with the default version of the test for all other categories. Each bar consisted of the following types of dice:

two eyes horizontal in the middle,

two eyes vertical in the middle,

two eyes horizontal upwards,

two eyes horizontal downwards,

two eyes vertical to the right,

two eyes vertical to the left.

The student had to click on the dice, in which the eyes are horizontal in the middle (the target dice). Figure 1 shows an example of a particular bar.

The dice were displayed in random order. Each bar consisted of 18 dice. For each button of a bar one of the above types of dice was randomly chosen with probability 1/6 and taken as that particular button. If in the resulting bar two consecutive dice were the same or, if the target dice did not occur at least once, then a new bar was created with a new arrangement of dice. This was repeated until no two consecutive dice were the same and the target dice occurred at least one time. Underneath each bar a separate “Next” button was displayed. The instruction read as follows: “Click on the dice, in which the eyes are horizontal in the middle. Work as quickly as possible, making no errors. Click on the dice in the order they are given in the dice bar starting from left to right. In order to go to the next dice bar you must click on the Next button, which is located underneath the dice.” If errors were made, the test was failed and the student had to start all over again.

The actual test consisted of a series of 25 observed consecutive reaction times. For the actual analysis not the original reaction times were used but reaction times which were corrected for possible test effects. The method of correction was as follows: Firstly, for

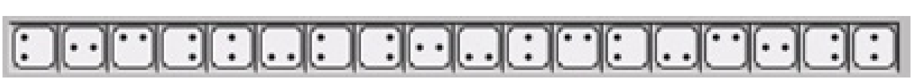

Figure 1:
Task bar and instruction “Click on the dice, in which the eyes are horizontal in the middle”

each reaction time series, a multiple regression analysis was performed with the number of buttons to be pressed and the serial number of the last button as independent variables and the reaction time as the dependent variable (*N* = 25). Next, the original reaction times were corrected by taking the residuals of the multiple regression added by the observed mean of the original reaction times. The analyses to be reported below are all based on these corrected reaction times. When using the ACT as a screening tool for potential attention concentration deficits and potential high achievers among ninth graders, values both above 1 standard deviations from the mean and below -1 standard deviations from the mean respectively, are used as an identification limit (A. G. H. S. van der Ven, personal communication, July 20, 2009). When a student fails to pass the test he or she is classified to the below 1 standard deviation group. We used four tier categorization (below -1 *SD*, between mean and -1 *SD,* between mean and + 1 *SD,* above + 1 *SD,*) to study if those students belonging to either high attention or low attention groups differed from each other and from attention groups falling between on scientific reasoning and school achievement.

Analysis procedures

Descriptive statistics were calculated with SPSS21. Path modelling was applied in AMOS21, using the default method of missing data handling (raw maximum likelihood missing data handling) (Byrne, 2010). The goodness of overall fit of the estimated models was evaluated by four indicators: χ^2 -test, CFI, TLI, RMSEA. The models were considered having a good fit with CFI and TLI > .95, and RMSEA < .06, and an acceptable fit with CFI and TLI > .90, and RMSEA < .08 Group comparisons in path analyses were conducted by fitting the models simultaneously on groups based on the ACT scores, constraining stepwise path coefficients and means equal across groups, and studying the relative changes in the fit indices (see Byrne & Stewart, 2009).

Results

First, we introduce descriptive statistics of the study by gender (Table 1). It is noteworthy that attention scores gained by the ACT results must be read in the opposite direction meaning the lower the value, the smaller the deviation and the better the attention.

The first hypothesis stated that attention would predict school achievement and explain performance in scientific reasoning tasks so that the effect of attention on school achievement would be partially mediated by scientific reasoning skills. To test this, two path models were specified. In the first stage of modeling it was first tested whether attention predicted GPA directly. Then scientific reasoning was added to the model as a mediator, and indirect effects were studied to determine whether the mediation hypothesis held (cf., Zhao, Lynch, & Chen, 2010). The effect of the gender on school achievement was controlled by adding gender in the model as a covariate. Only the students who passed the ACT test were used in this model (*n* = 301).

Table 1:
The descriptives of the study

Measure	n	Items	Min	Max	Mean	SD
			Scientific reasoning			
Boys	187	9	0	9	3.71	2.66
Girls	162	9	0	9	3.86	2.51
Total	349	9	0	9	3.78	2.59
GPA						
Boys	192	11	5.18	9.91	7.69	1.06
Girls	166	11	5.73	9.82	8.06	.94
Total	358	11	5.18	9.91	7.86	1.02
*Attention**						
Boys	157		-1.13	1.18	-.29	.39
Girls	145		-1.31	.98	-.25	.44
Total	302		-1.31	1.18	-.27	.42

Note: *Fifty-six subjects failed the ACT test and constituted a low attention group and due to relative large size of this group we named them as below -1 *SD* group.

The first stage of modeling revealed that attention and gender were not related but as expected, attention predicted GPA (β = -.15, p < .01). When scientific reasoning was added to the model, the direct effect of attention on GPA decreased to β = -.07 (*ns*.), which can be interpreted as a sign of mediation. Gender did not predict scientific reasoning. The final model without the non-significant paths is presented in Figure 2.

When focusing on scientific reasoning and its predictive properties it was found that the effect of scientific reasoning (β = .44, p < .001) on GPA, explained 19 % of the variance of GPA. Furthermore, gender (β = -.18, p < .001) explained 3 % of the GPA variance. Attention (β = -.19, p < .001) explained 3 % of the variance of scientific reasoning. Altogether 22 % of the variance of GPA could be explained with this model which fit the data excellently ($\chi 2$ = 0.88, df = 2, p = 0.64, *CFI* = 1.000, *TLI* = 1.019, *RMSEA* = .000).

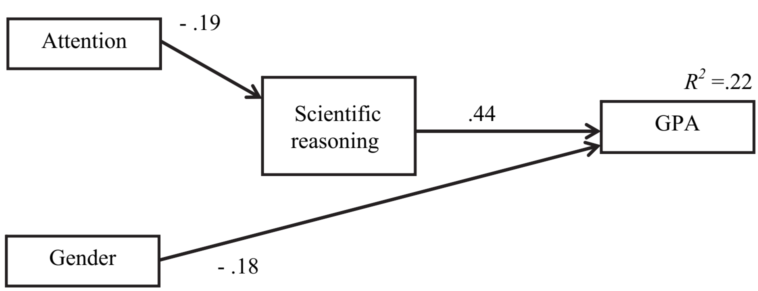

Figure 2:
Model of predicting school achievement

In Model 2, which is depicted in Figure 3, attention scores were divided into four groups: those who failed (*n* = 56) formed below -1 SD group, between -1 *SD* and mean (*n* = 117), between mean and + 1 *SD* (*n* = 121), above +1 *SD* (*n* = 48) to study effect of such grouping. When the effect of gender on GPA was controlled for, GPA was directly predicted by attention grouping and scientific reasoning, and indirectly by attention grouping via scientific reasoning. The direct effect of attention grouping on scientific reasoning (β = .24, p< .001) explained 6 % of the scientific reasoning variance. The total effect of attention grouping on GPA (β = .26) explained 7 % of the GPA variance.

When focusing on scientific reasoning and its predictive properties it was found that the effect of scientific reasoning (β = .38, p < .001) on GPA explained 14 % of the GPA variance. Moreover, gender (β = -.17, p < .001) explained 3 % of the GPA variance. The total model explained 23 % of the GPA variance. The model fit the data well ($\chi 2$ = 0.47, *df* = 2, p = 0.790, CFI = 1.000, TLI = 1.072, RMSEA = .000).

To study second hypothesis *"*Students identified to have high attention have higher scientific reasoning and school achievement than the students who did not show such high attention", another path model was specified without having attention as a predictor. Instead, the earlier introduced grouping based on the ACT measure was used as a grouping variable. The groups were compared by fitting the model to all groups simultaneously and constraining stepwise the paths and means equal across groups. The constrained models showing relevant results are presented in the following Table 2. Changes of the fit indices (CFI and TLI) were used to determine if the constrained models are different from the baseline model (see Table 2). Constraints were added in the following way: in the first model, path coefficients were set equal across groups, followed by the model in which path coefficients and gender (relative proportions of genders) were equal across groups. These models fitted the data well and as the indices changed only minimally, it was concluded that there were no group differences regarding the paths and gender. The next model in where paths and all means were set equal across groups showed an extremely poor fit. This indicated that there were statistically significant differences between groups in either GPA, scientific reasoning or both. These differences were ana-

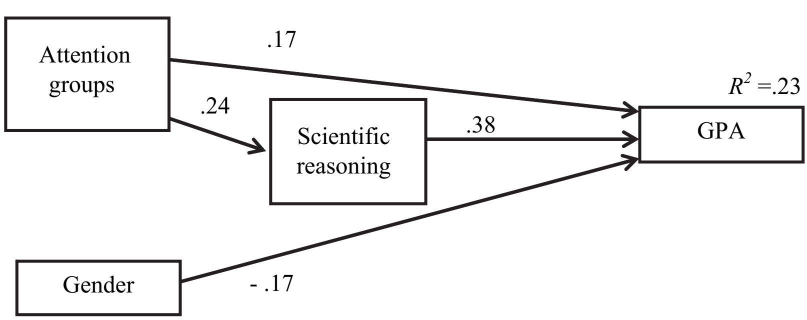

Figure 3:
Model two of predicting school achievement

lyzed in more detail by first constraining the means of the two variables equal across groups one at a time while the other variable was allowed to vary free. As this showed that there were statistically significant group differences in both variables, the groups were then compared by constraining the group means equal pairwise, one variable at a time.

Table 2:
Model comparison of the latent variables and changes in fit indices in group comparisons

Model	$\chi^2\Delta$	$df\Delta$	**CFI**	**TLI**	**RMSEA**	$p\Delta$
GPA and Scientific reasoning by attention groups						
Baseline model (χ^2 results are reported instead of the $\chi^2\Delta$ differences)	.551	4	1.000	1.445	.000	.968
Path coefficients constrained equal across groups	4.800	6	1.000	1.240	.000	.570
Paths coefficients and gender (relative proportions of the two genders) constrained equal across groups	3.612	3	1.000	1.160	.000	.306
Path coefficients, gender and all means constrained equal across groups	32.595	6	.515	.387	.058	<.001
Group comparisons						
GPA (scientific reasoning varies free across groups)						
High able and above mean group are equal	3.752	1	1.000	1.047	.000	.053
High able and below mean group are equal	7.910	1	.938	.894	.024	.005
High able and low performing are equal	10.149	1	.890	.812	.032	.001
Low performing and below mean are equal	1.296	1	1.000	1.138	.000	.255
Low performing and above mean are equal	4.118	1	1.000	1.034	.000	.042
Below and above mean are equal	1.595	1	1.000	1.127	.000	.207
Scientific reasoning (GPA varies free across groups)						
High able and above mean group are equal	5.527	1	.963	.982	.010	.019
High able and below mean group are equal	11.002	1	.872	.780	.032	.001
High able and low performing group are equal	18.528	1	.710	.503	.052	<.000
Low performing and below mean are equal	3.234	1	1.000	1.066	.000	.072
Low performing and above mean are equal	8.488	1	.926	.873	.026	.004
Below and above mean are equal	1.957	1	1.000	1.114	.000	.162

Note: CFI= Comparative fit index, TLI=Tucker-Lewis index, RMSEA=Root mean square error of approximation

Table 3:
Descriptives of the Scientific reasoning and GPA by attention grouping based on ACT-results

Measure	Grouping	*n*	*M*	*SD*	*Min*	*Max*
Scientific reasoning	not passed (low attention)	56	2.77	2.60	0	8
	below average	117	3.57	2.61	0	8
	above average	121	3.99	2.45	0	9
	highest (high attention)	48	4.98	2.42	0	9
	Total	342	3.79	2.60	0	9
GPA	not passed (low attention)	56	7.41	1.01	5.18	8.91
	below average	123	7.75	1.00	5.64	9.82
	above average	124	7.95	.98	5.36	9.82
	highest (high attention)	48	8.37	.91	5.64	9.73
	Total	351	7.85	1.02	5.18	9.82

Group comparisons regarding the both studied variables showed approximately parallel model fit decrease when the comparison advanced towards low performing group. Table 2 shows that there were statistically significant differences between high attention and all the other groups in scientific reasoning. Differences were parallel in school achievement with the exception between high attention group and above mean attention group (see also descriptives in Table 3).

Discussion

In this study we examined how attention measured in prolonged over-learned response tasks predicts scientific reasoning and school performance and if high levels of attention have a significant relationship with high level scientific reasoning and school performance. This was done by fitting path models on the data from 358 Finnish ninth graders. In the models, school achievement was predicted by performance in the Attention Concentration Test and scientific reasoning test, controlling the effect of gender on GPA. Modelling showed that the most powerful single predictor of school performance was scientific reasoning. In general, in the comprehensive model, attention added little to the predictive validity indicated by scientific reasoning. The effect of attention was mainly indirect, scientific reasoning being the mediator. However, results showed that students who belonged to the high attention group possessed higher scientific reasoning scores and school grades (GPA) when compared to their counterparts possessing lower attention level. In other words, exceptional high level attention measured in simple prolonged working task seems to facilitate attainment of high-level scientific reasoning and school achievement.

Differences emerged between students' performance and students' attention level. First, the results showed that students who did not pass the test possessed lower scientific reasoning and lower school grades. These findings can be interpreted in three ways: (a) students who fail the attention test have fundamental deficiencies in their scientific reasoning and learning; (b) students who failed the test may not have adjusted toward tasks measuring their school related behaviour, including scientific reasoning and computer tasks, and (c) students who failed the test require more time to learn and over-learn preconditions of the ACT – which correlates their success in scientific reasoning tasks and school performance. Nevertheless, in addition, these study findings show that unsuccessful ACT performance is associated with lower level scientific reasoning and school performance. This finding in turn has both theoretical and practical implications; when do these differences in attention, scientific reasoning and school achievement emerge and how do they develop? How attention measured by prolonged over-learned response tasks is associated with the working memory measures that have often stated to be in the central role in the cognitive performance (cf. Bühner, Mangels, Krumm & Ziegler, 2005). If there are differences in children's attention before school-age how does this affect their learning in informal and formal learning situations?

Secondly, two additional conclusions can be drawn from the results, at least in general terms. First, mediation analysis revealed that attention did not have a direct effect on school performance itself but via scientific reasoning. Second, when the ACT results were used as a screening measure to identify either low or high level attention, as van der Ven (personal communication, July 20, 2009) have recommended, the formed four-tier categorization showed clear between groups differences for both the high and low attention groups both in scientific reasoning and school achievement.

To deepen the discussion, our results for the relationship of formal operation to achievement align well with previous results from intelligence test scores, that indicate 25 % of variation in achievement is explained by intelligence (Steinmayr & Spinath, 2008), and even more with Freudenthaler et al.'s (2008) results, demonstrating 23 % of the achievement variance was due to the intelligence factor. However, in this study we conclude that the 19 % of explained variance in GPA is accorded to scientific reasoning. Indeed, high level attention measured in prolonged over learned response tasks seems to facilitate acquisition of the higher level scientific reasoning which in turn is associated with school performance. Our study does not clearly show the extent to which high level attention is required for the highest level of scientific reasoning but it certainly facilitates learning the most challenging tasks.

This study has numerous practical implications. It is obvious that by this method one can identify low-performing students. The test parameters can be adjusted for the age-level suitable for the kindergarten-aged children onwards. Additionally, due to the fact that the ACT assesses attention in a consistent way, it could be used as an additional method along with the more traditional and often interpretive ways (e.g. check-lists, ICD 10) to reliably identify attention capacity and its disorders (e.g. ADD and ADHD). The ACT could help to make the important distinction whether ADHD has a neurological origin, or whether observed secondary attention and behavioral difficulties are products of poor environmental-fit. Thus, by using ACT as a screening method we might likely find those children who will benefit from further diagnostics and possible medical treatment, and on the other hand, those, whose attention difficulties are products of deficits in their environment, and could be supported simply by altering prevailing conditions.

We also can suggest that the ACT be used in intervention and training purposes. Though it seems obvious, those who have high scores in ACT possess the ability to perform academically at the higher level. Thus, the ACT provides a method to identify both potential high achievers and underachievers. Results of the ACT test could provide added value in educational counseling situations, especially when assigning high-performing students to academic tracks. In future studies it would be important to determine if people with the highest attention levels achieve on the highest levels in other ways, for example in extracurricular activities such as music, sport and arts.

In summary, this study revealed that simple prolonged working tasks provide predictive values of both scientific reasoning and school achievement. This finding is based on evidence that when measuring cognitive performance and its components very simple attentional tasks should be used to measure characteristics that are idiosyncratic to the learner and to his or her cumulative learning (cf. Gustafsson & Carlstedt, 2006; Heitz, Unsworth & Engle, 2004; Smit van der Ven, 1995). Individual differences in simple attentional tasks seem to provide promising additional value when building comprehensive models explaining cognitive performance such as intelligence. When knowing that environment and motivational factors strongly effect learning and development of expertise we do not argue that attention measured by the tools, such as the ACT, explain the greatest proportion of those achievements, but we assume that differences in attention may orient towards best fit between individual and environment and only exceptional deficiencies in attention can create overwhelming barriers for learning (Stoeger, 2013; Ziegler & Phillipson, 2013). However, to study how applicable this model is outside of the academic field calls for further research (see also Ziegler, Heckel & Ziegler, 2013).

To sum up, this study suggests how we may identify academic low and high achievers at the end of Finnish comprehensive school by wise use of the ACT. Despite concerns regarding the comprehensive nature of the Finnish school system the success level in the ACT produced statistically significant differences in relation to the examined study variables, scientific reasoning and school achievement. Consequently, according to these studied variables, the potential academic giftedness and potential academic low-achievers could be identified in the future with the help of the ACT. It is noteworthy that with a poor ACT test result it is nearly impossible to be a high achiever in both in scientific reasoning and school achievement. Simultaneously, many students end up as low achievers in measured outcome variables regardless of their good or even excellent level of attention. According to the latest PISA findings (Kupari et al., 2013) and according to findings that report a negative Flynn effect in developed western countries (Australia, Finland, Great-Britain, New Zealand, USA, etc.) including Finland, students' cognitive levels in such measures are declining when comparing performance of earlier age cohorts at all achievement levels (Dutton & Lynn, 2013). As Kuhn (2007) reminds us, in today's school the most challenging task is to make school sensible to the young people for whom they are designed, including the high achievers.

References

Adey, P., Csapo, B., Demteriou A., Hautamäki J., & Shayer, M. (2007). Can we be intelligent about intelligence? Why education needs the concept of plastic general ability. *The Educational Research Review,* 2(2), 75–97.

Baillargeon, R., Pascual-Leone, J., & Roncadin, C. (1998). Mental-attentional capacity: Does cognitive style make a difference? *Journal of Experimental Child Psychology, 70*(2), 143–166.

Baddeley, A. D., (1986). *Working Memory.* Oxford: Oxford University Press.

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences, 29*(2), 109–124.

Bruner, J. (1986). Actual minds, possible worlds. Cambridge, Mass.; Harvard University Press.

Bühner, M., Mangels, M., Krumm, S., & Ziegler, M. (2005). Are working memory and attention related constructs? *Journal of Individual Differences, 26*(2), 287-321.

Byrne, B. M. (2010). *Structural equation modelling with AMOS: Basic concepts, applications, and programming* (2nd Ed.). New York: Taylor & Francis.

Byrne, B.M. & Stewart, S.M. (2006). The MACS approach to testing for multigroup I invariance of a second-order structure: A walk through the process. Structural Equation Modeling. *A Multidisciplinary Journal, 13*(2), 287-321.

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

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental psychology, 27(*5), 703–722.

Cole, M. (1996). *Cultural Psychology – a once and future discipline.* Cambridge: Harvard University Press.

Cowan, N. (2000). Processing limits of selective attention and working memory: Potential implications for interpreting. *Interpreting, 5*(2), 117–146.

Demetriou, A., Spanoudis, G., & Mouyi, A. (2011). Educating the developing mind: Towards an overarching paradigm. *Educational Psychology Review, 23*(4), 601–663.

Dutton, E., & Lynn, R. (2013). A negative Flynn effect in Finland, 1997–2009*, Intelligence, 4*(6), 817–820.

Emick, J., & Welsh, M. (2005). Association between formal operational thought and executive function as measured by the tower of Hanoi-Revised. *Learning and Individual Differences, 15*(3), 177–188.

Freudenthaler, H. H., Spinath, B., & Neubauer, A. C. (2008). Predicting school achievement in boys and girls. *European Journal of Personality, 22*(3), 231–245.

Garlick, D. (2002). Understanding the nature of the general factor of intelligence: the role of individual differences in neural plasticity as an explanatory mechanism. *Psychological review, 109*(1), 116.

Geary, D. C. (2005). *The origin of mind.* Washington, DC: American Psychological Association.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*, 316–322.

Gottfredson, L. S. (2002). G: Highly general and highly practical. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The general factor of intelligence. How general is it?* (331–380). Mahwah, NJ: Erlbaum.

Gustafsson, J.-E. (2002). Measurement from a hierarchical point of view. In H. I. Braun, D. N. Jackson, & D. E.Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 73– 95). Mahwah: Lawrence Erlbaum.

Gustafsson, J. E., & Carlstedt, B. (2006). Abilities and grades as predictors of achievement:

The encapsulation theory. In Paper presented at the symposium "The investment theory of intelligence: New evidence, new challenges" given at the annual meeting of the American Psychological Association, New Orleans, August 10–13, 2006.

Hautamäki, J. (1989). The application of a Rasch model on Piagetian measures of stages of thinking. In P. Adey, J. Bliss, J. Head, & M. Shayer (Eds.), Adolescent development and school science (pp. 342–349). London: Falmers.

Hautamäki, J. (2000). Formaali ajattelu yleisenä oppimisvalmiutena [Formal operations as a general learning competence]. In J. Hautamäki, P. Arinen, A. Hautamäki, . . ., & P. Scheinin (Eds.), *Oppimaan oppiminen yläasteella [Learning to learn in lower secondary education]*. Evaluation 7/2000 (pp. 57–71). Helsinki: National Board of Education.

Hautamäki, J., Harjunen, E., Hautamäki, A., Karjalainen, T., Kupiainen, S. et al. (2008). *PISA06 Finland: Analyses, reflections, explanations. Publication No. 44*. Helsinki, Finland: Ministry of Education.

Hautamäki, J., Kupiainen, S., Marjanen, J., Vainikainen, M-P., & Hotulainen, R. (2013). Oppimaan oppiminen peruskoulussa. Tilanne vuonna 2012 ja muutos vuodesta 2001. [Learning to learn at the end of basic education. Results in 2012 and changes from 2001]. *Department of Teacher of Education Research Report, No 347*. Helsinki: University Press.

Heitz, R. P., Unsworth, N. & Engle, R. W. (2004). Working Memory Capacity, Attention Control, and Fluid Intelligence. In: O. Wilhelm & R.W. Engle (Eds.), *Handbook of Understanding and Measuring Intelligence* (pp. 61–77). London: Sage Publications.

Houde´, O. (2000). Inhibition and cognitive development: object, number, categorization and reasoning. *Cognitive Development, 15*(1), 63–73.

Jensen, A. R. (1982). Reaction time and psychometric "g". In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93–132). New York: Springer-Verlag.

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

Klauer, K. C. (1993). *Belastung und Entlastung beim Problemlösen. Ein Theorie des deklarativen Vereinfaches. [Cognitive load in problem-solving. Theory of declarative decision-making]*.Göttingen: Hogrefe.

Kuhn, D. (2002). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 371–393). Oxford, England: Blackwell.

Kuhn, D. (2007). How to produce a high-achieving child. *Phi Delta Kappan, 88*(7), 757–763.

Kuhn, D. (2008). Formal operations from a twenty-first century perspective. *Human Development, 51*(1), 48–55.

Kuhn, D., Iordanou, K., Pease, M., & Wirkala, C. (2008). Beyond control of variables: What needs to develop to achieve skilled scientific thinking? *Cognitive Development, 23*(4), 435–451.

Kupari, P., Välijärvi, J., Andersson, L., Arffman, I., Nissinen, K., Puhakka, E., & Vettenranta, J. (2013). *PISA12 ensituloksia. [PISA12 preliminary results].*Helsinki: Publications of Finnish Ministry of Education and Culture.

Kuusela, J. (2002). *Tieteellisen paradigman mukaisen ajattelun kehittyminen peruskoulussa. [The development of thinking towards a scientific paradigm in the comprehensive school].* Department of Teacher Education. Research Report 221. Helsinki: Unversity Press.

Larson, G. E., & Alderton, D. L. (1990). Reaction Time Variability and Intelligence: A "Worst Performance" Analysis of Individual Differences. *Intelligence, 14*(3), 309–325.

Logan, G. D. (2004). Cumulative progress in formal theories of attention. *Annual Review of Psychology, 55*(2), 207–234.

Lu, L., Weber, H. S., Spinath, F. M., & Shi, J. (2011). Predicting School Achievement from Cognitive and Non-Cognitive Variables in a Chinese Sample of Elementary School Children. *Intelligence, 39*(2–3), 130–140.

Luria, A. R. (1976). *Cognitive development. Its cultural and social development.* Cambridge, Mass.: Harvard University Press

Neimark, E. (1975). Intellectual development during adolescence. In F. Horowitz (Ed.), *Review of child development research* (pp. 541–594). Chicago: University of Chicago Press.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: new findings and theoretical developments. American *Psychologist, 67*(2), 130–140.

Olson, D. R. (2003). *Psychological theory and educational reform: How school remakes mind and society*. Cambridge University Press.

Piaget, J. (1970). Structuralism. New York: Harper & Row.

Piaget, J. (1972). *Psychology and epistemology: Towards a theory of knowledge.* Harmondsworth: Penguin.

Piaget, J. (2006). Reason. *New Ideas in Psychology, 24*(1), 1–29.

Shayer, M. (1979). Has Piaget's construct of formal operational thinking any utility? British *Journal of Educational Psychology, 48*(1), 62–70.

Shayer, M. (2008). Intelligence for Education: As described by Piaget and measured by psychometrics. *British Journal of Educational Psychology, 78*(1), 1–29.

Shayer, M., & Wylam, H. (1978). The distribution of Piagetian stages of thinking in British middle and secondary school children 11 – 14/16-year olds and sex differentials. *British Journal of Educational Psychology, 48*(1), 62–70.

Shute, G., & Huertas, V. (1990). Developmental variability in frontal lobe function. *Developmental Neuropsychology, 6*(1), 1–12.

Smit, J. C., & van der Ven, A. H. G. S. (1995). Inhibition in Speed and Concentration Tests: the Poisson Inhibition Model. *Journal of Mathematical Psychology, 39*(3), 265–273.

Spearman, C. (1927). *The Abilities of Man.* London: MacMillan.

Spinath, B., Spinath, F. M., Harlaar, N., & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence, 34*(4), 363–374.

Strand-Cary, M., & Klahr, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development, 23*(4), 488–511.

Steinborn, M. B., Flehmig, H. C., Westhoff, K., & Langer, R. (2008). Predicting school achievement from self-paced continuous performance: Examining the contributions of response speed, accuracy, and response speed variability. *Psychology Science Quarterly, 50*(4), 613–634.

Steinmayr, R., & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? European Journal of Personality, 22(2), 185–209.

Stoeger, H. (2013). Learning as a creative process. In A.-G. Tan (Ed.). *Creativity, talent, and excellence* (pp. 3–11). Singapore: Springer.

Ven, A. H. G. S. (2013, December 12). *Welcome to the Attention Concentration Test.* Retrieved from http://www.ru.nl/pwo/olo/test-cbo/tests/act_english_welcome/

Ven, A. H. G. S. van der (2001). A Theoretical Foundation of Speed and Concentration Tests. In F. Columbus (Ed.). *Advances in Psychology Research IV*, (pp. 315–354). Huntington: Nova Science Publishers, Inc.

Ven, A. H. G. S. van der, Gremmen F. M., & Smit, J. C. (2005). A Statistical Model for Binocular Rivalry. *British Journal of Mathematical and Statistical Psychology, 58*(2), 97–116.

Ven, A. H. G. S. van der, Smit, J. C., & Jansen, R. W. (1989). Inhibition in prolonged work tasks. *Applied Psychological Measurement, 13*(2), 177–191.

Vygotsky, L. S. (1978). *Mind in society – the development of higher psychological processes*. Cambridge, Mass: Harvard University Press.

Zhao, X., Lynch, J. G. & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*, 197–206

Ziegler, A., & Phillipson, S. N. (2012). Towards a systemic theory of gifted education. *High Ability Studies, 23*(1), 3–30.

Ziegler, A. Heckel, A., & Ziegler, A. 2013. Mathematical foundations of and empirical investigations into the dynamic of top positions: Stabilization Effect, Reversed Matthew Effect, and Heraclitus Effect. *Psychological Test and Assessment Modeling, 55*(4), 401–414.