

# PLOS Medicine

## Inattention in primary school is not good for your future school achievement - a pattern classification study --Manuscript Draft--

<b>Manuscript Number:</b>	
<b>Full Title:</b>	Inattention in primary school is not good for your future school achievement - a pattern classification study
<b>Short Title:</b>	Inattention and school achievement
<b>Article Type:</b>	Research Article
<b>Keywords:</b>	Inattention; academic achievement; longitudinal; pattern classification; machine learning
<b>Corresponding Author:</b>	Astri Johansen Lundervold, PhD University of Bergen Bergen, NORWAY
<b>Corresponding Author Secondary Information:</b>	
<b>Corresponding Author's Institution:</b>	University of Bergen
<b>Corresponding Author's Secondary Institution:</b>	
<b>First Author:</b>	Astri Johansen Lundervold, PhD
<b>First Author Secondary Information:</b>	
<b>Order of Authors:</b>	Astri Johansen Lundervold, PhD Tormod Bøe Arvid Lundervold
<b>Order of Authors Secondary Information:</b>	
<b>Abstract:</b>	<p><b>Objective.</b> Inattention in childhood has been associated with academic problems later in life. The contribution of specific aspects of inattentive behaviour is, however, less known. We investigated the importance of primary teacher reports on nine aspects of inattentive behaviour in predicting future academic achievement.</p> <p><b>Methods.</b> Primary school teachers of 2397 children (7 - 9 years) rated nine items reflecting different aspects of inattentive behaviour. Academic achievement scores were available from an official school register about ten years later. All scores were at a categorical level. Two multivariate statistical methods (in R), a multinomial logistic regression analysis and a Classification and Regression Trees (CART) analysis, were included to predict high school academic achievement from the nine inattention items, gender and their primary school class level (grade) when evaluated by their teachers (2nd to 4th grade, 7 - 9 years).</p> <p><b>Results.</b> Problems related to sustained attention and distractibility were found to be important predictors of low level of academic achievement in high school. Overall, inattention was rated as more severe in boys, who also obtained lower achievement scores at high school than girls. However, if reported to have no problems related to sustained attention and distractibility in the 4th primary school class level, the percentage of boys with the highest level of the achievement score was as high as in girls (&gt; 40%).</p> <p><b>Conclusion</b> Primary school teachers' reports of problems related to sustained attention and distractibility were strong predictors of low academic achievement in high school. Identification and follow-up procedures of primary school children showing this pattern of inattentive behaviour should be prioritised to prevent future academic and</p>

	occupational failure.
Suggested Reviewers:	
Opposed Reviewers:	
Additional Information:	
Question	Response
<p><b>Financial Disclosure</b></p> <p>Please describe all sources of funding that have supported your work. <b>This information is required for submission and will be published with your article, should it be accepted.</b> A complete funding statement should do the following:</p> <p>Include <b>grant numbers and the URLs</b> of any funder's website. Use the full name, not acronyms, of funding institutions, and use initials to identify authors who received the funding.</p> <p><b>Describe the role</b> of any sponsors or funders in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. If the funders had <b>no role</b> in any of the above, include this sentence at the end of your statement: "<i>The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.</i>"</p> <p>However, if the study was <b>unfunded</b>, please provide a statement that clearly indicates this, for example: "<i>The author(s) received no specific funding for this work.</i>"</p> <p>* typeset</p>	<p>The author(s) received no specific funding for this work.</p>
<p><b>Competing Interests</b></p> <p>You are responsible for recognizing and disclosing on behalf of all authors any competing interest that could be perceived to bias their work, acknowledging all financial support and any other relevant financial or non-financial competing interests.</p> <p>Do any authors of this manuscript have competing interests (as described in the <a href="#">PLOS Policy on Declaration and Evaluation of Competing Interests</a>)?</p>	<p>The authors have declared that no competing interests exist.</p>

If **yes**, please provide details about any and all competing interests in the box below. Your response should begin with this statement: *I have read the journal's policy and the authors of this manuscript have the following competing interests:*

**If no authors have any competing interests to declare, please enter this statement in the box: *"The authors have declared that no competing interests exist"***

\* typeset

## Data Availability

PLOS journals require authors to make all data underlying the findings described in their manuscript fully available, without restriction and from the time of publication, with only rare exceptions to address legal and ethical concerns (see the [PLOS Data Policy](#) and [FAQ](#) for further details). When submitting a manuscript, authors must provide a Data Availability Statement that describes where the data underlying their manuscript can be found.

Your answers to the following constitute your statement about data availability and will be included with the article in the event of publication. **Please note that simply stating 'data available on request from the author' is not acceptable. If, however, your data are only available upon request from the author(s), you must answer "No" to the first question below, and explain your exceptional situation in the text box provided.**

Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?

No - some restrictions will apply

Please describe where your data may be found, writing in full sentences. **Your answers should be entered into the box below and will be published in the form you provide them, if your manuscript is accepted.** If you are copying our sample text below, please ensure you replace any instances of **XXX** with the appropriate details.

he R markdown notebook, implementing our analysis, will be available on GitHub [address TBA]. The data are also planned to be published in its anonymous form, after being clarified by the data owner.

If your data are all contained within the paper and/or Supporting Information files, please state this in your answer below. For example, "All relevant data are within the paper and its Supporting Information files."

If your data are held or will be held in a public repository, include URLs, accession numbers or DOIs. For example, "All XXX files are available from the XXX database (accession number(s) XXX, XXX)." If this information will only be available after acceptance, please indicate this by ticking the box below. If neither of these applies but you are able to provide details of access elsewhere, with or without limitations, please do so in the box below. For example:

"Data are available from the XXX Institutional Data Access / Ethics Committee for researchers who meet the criteria for access to confidential data."

"Data are from the XXX study whose authors may be contacted at XXX."

\* typeset

Additional data availability information:

**Dear Editor of Plos One**

Thank you for considering the enclosed manuscript entitled “Sustained attention and distractibility in primary school predict academic achievement in high school - a pattern classification study” for publication in Plos One. We know from previous studies that inattention in early childhood may be crucial for present and future academic success. Most of these studies have included samples of individuals with Attention Deficit Hyperactivity disorder and have defined inattention as the sum score across a set of items. Symptoms of inattention are, however, not restricted to individuals with a specific diagnostic label, and the items used to define inattention indeed reflect different aspects of cognitive function. The present study asked if we could identify specific features of inattentive behavior in primary school children acting as primary predictors of academic achievement in high school. To accomplish that, we included data from 2397 children participating in the longitudinal, population-based Bergen Child Study. Inattentive behavior was rated on a set of nine items on a modified Likert scale by primary school teachers when the participants were 7 – 9 years old. Academic achievement level was obtained for all these participants about 10 years later, when they attended high school. Two statistical methods were included – a multinomial logistic regression analysis and a Classification and Regression Trees analysis - to handle a set of categorical predictors and an outcome variable of three levels of academic achievement. Primary school teachers’ reports of problems related to sustained attention and distractibility turned out to be the most important predictors of academic achievement 10 years later, emphasizing the importance of identification and help to children struggling with these problems.

Sincerely,



Astri J. Lundervold (astri.lundervold@uib.no)

# Inattention in primary school is not good for your future school achievement - a pattern classification study

Astri J. Lundervold<sup>1,2\*</sup>, Tormod Bøe<sup>3</sup>, Arvid Lundervold<sup>4</sup>

- 1 Department of Biological and Medical Psychology University of Bergen, 5009 Bergen, Norway
- 2 K.G. Jebsen Center for Research on Neuropsychiatric Disorders, University of Bergen, Bergen, Norway
- 3 Regional Centre for Child and Youth Mental Health and Child Welfare, Uni Research Health, Bergen, Norway
- 4 Neuroinformatics and Image Analysis Laboratory, Department of Biomedicine, University of Bergen

\* Corresponding author: [astri.lundervold@uib.no](mailto:astri.lundervold@uib.no)

## Abstract

*Objective.* Inattention in childhood has been associated with academic problems later in life. The contribution of specific aspects of inattentive behaviour is, however, less known. We investigated the importance of primary teacher reports on nine aspects of inattentive behaviour in predicting future academic achievement.

*Methods.* Primary school teachers of 2397 children (7 - 9 years) rated nine items reflecting different aspects of inattentive behaviour. Academic achievement scores were available from an official school register about ten years later. All scores were at a categorical level. Two multivariate statistical methods (in R), a multinomial logistic regression analysis and a Classification and Regression Trees (CART) analysis, were included to predict high school academic achievement from the nine inattention items, gender and their primary school class level (grade) when evaluated by their teachers (2<sup>nd</sup> to 4<sup>th</sup> grade, 7 - 9 years).

*Results.* Problems related to sustained attention and distractibility were found to be important predictors of low level of academic achievement in high school. Overall, inattention was rated as more severe in boys, who also obtained lower achievement scores at high school than girls. However, if reported to have no problems related to sustained attention and distractibility in the 4<sup>th</sup> primary school class level, the percentage of boys with the highest level of the achievement score was as high as in girls (> 40%).

*Conclusion* Primary school teachers' reports of problems related to sustained attention and distractibility were strong predictors of low academic achievement in high school. Identification and follow-up procedures of primary school children showing this pattern of inattentive behaviour should be prioritised to prevent future academic and occupational failure.

## Introduction

Inattention in early childhood has been linked to a wide range of behavioural and social problems [2,11], including poor academic achievement. This has been shown in several

1  
2  
3

studies of individuals with Attention Deficit Hyperactivity disorder (ADHD, see e.g. [18, 22]), but is also widely documented in studies including community samples [13, 14, 17, 20, 25, 26, 30]. In most of these studies, inattention is defined as a sum score across a set of items.

Inattention is, however, a multidimensional concept, where the items reflect impairment of sustained and focused attention, impaired working memory, distractibility, forgetfulness, as well as impaired ability to organise and plan activities and tasks. These aspects of inattention have been described as independent at a biological level [5], but may be extremely difficult to disentangle behaviourally. They rather tend to occur as patterns of behaviour. For example, most children may be distracted by external stimuli in a classroom situation [28], and these distractions will probably be especially hard to handle by a child having problems in maintaining attention and engagement in a task. Thus, it may not be the total inattention score, but rather specific patterns of inattentive behaviour that have the most detrimental effect on the child's present and future function at school. Identification of important features of inattentive behaviour will therefore be of great importance when developing remediation procedures.

Primary school teachers' skills are crucial in the work to detect and help a child struggling with inattention. They observe their pupils on a regular basis and in a wide range of situations where inattention tends to have negative effects on performance. At the same time, one should be aware of the risk of biases. Primary school teachers may for example be more tolerant to the behaviour of a child in the lowest class levels, and previous studies have shown that teachers tend to rate girls as less impaired than boys, even when the girls exhibit problematic behaviour in the classroom [1, 8, 21]. The child's gender and age should therefore be taken into account when evaluating teacher ratings of inattentive behaviour.

The aim of the present study was to further investigate the importance of primary school teachers' reports of inattentive behaviour. To that end, we included data from the Bergen Child Study, where primary school teachers completed a questionnaire including nine items reflecting different aspects of inattentive behaviour when the children were between 7 and 9 years old. About ten years later, when the children had become high school students, academic achievement scores from the official school registry of Norway were available for 2397 of the children from the original sample. Each of the nine inattention items were rated on a Likert scale with three response alternatives, and the outcome variable, academic achievement, was discretised into three intervals, including an almost equal number of participants in each category. Teacher scores on each of the nine items were used as predictors together with gender and primary school class level (a proxy for age) to answer the following questions: (1) which feature of inattentive behaviour in primary school represents the strongest predictor of academic achievement in high school? (2) are there specific patterns of associations between these predictor variables?, and (3) are gender and the age of the child when evaluated by their primary school teachers of importance to the prediction?

In this context, two statistical approaches were selected according to the following criteria: (i) the methods must handle multiple predictors with a small set of response alternatives, and with a small set of outcome categories; (ii) the methods should be generic and of interest to other similar data analysis situations and prediction challenges occurring in the behavioural sciences, and (iii) the methods should produce results that are easy to interpret at a clinical level. Based on these criteria we selected the following two statistical methods: *multinomial logistic regression* (MLR), and *classification and regression trees* (CART). The classification and regression tree is a multivariate analysis method belonging to the class of recursive partitioning algorithms. These methods have recently become a valuable tool for exploring complex datasets in psychology where the algorithms generate decision trees that aim to correctly classify members of a sample.

This is obtained by splitting them into subsamples based on several binary decisions operating hierarchically on the values of the independent predictor variables. These decisions are easy to interpret and visualise. Together they reveal class predictive patterns of the independent variables in the sample, such that the decision tree can generalise and be used to classify new cases. Being trained and used for such predictive purposes, both MLR and CART are well established methods in the field of machine learning.

## Materials and methods

The first wave of the Bergen Child Study (BCS) was launched in October 2002 and included the total population of 9,430 children attending second to fourth grade (7–9 years old, born in 1993, 1994 and 1995) in all public, private, and special schools in Bergen. During the initial screening phase, parents and teachers were asked to complete a four-page BCS questionnaire, including, among other scales, a somewhat modified Swanson, Nolan, and Pelham Questionnaire - Fourth Edition (SNAP-IV); [33]. Sample protocols of the first wave have been described in several previous publications from the Bergen Child Study group [16, 19, 31].

A fourth wave of BCS was conducted when the youth were between 16 and 19 years old; it comprised a more comprehensive sample than in the original BCS sample, including all adolescents born between 1993 and 1995 living in the county of Hordaland. This county includes the city of Bergen, and the BCS sample was thus nested within this Hordaland sample. A total of 10,222 adolescents completed a questionnaire asking for information about school attendance and psychiatric diagnosis. The BCS was approved by the Regional Committee for Medical and Health Research Ethics (REC), Western Norway (2015/800 Barn i Bergen/ung@hordaland). Parents gave written consent for participation in the first wave of the study. In accordance with the regulations from the REC and Norwegian health authorities, adolescents aged 16 years and older can make decisions regarding their own health (including participation in health studies), and thus gave consent themselves to participate in the fourth wave of the study. Parents/guardians have the right to be informed, and in the current study, all parents/guardians received written information about the study in advance.

## The sample

The present study included 2397 participants (1141 boys) with teacher reports on all selected SNAP-IV items when they were 7 to 9 years old (primary school class levels 2, 3, or 4), and information about gender and academic achievement when they attended high school (16 to 19 years old). The percentages of children in the 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> class levels when evaluated by their teachers were 41.9%, 34.7% and 23.4%, respectively.

## Teacher reports

Items representing *Inattention* were selected from the SNAP-IV questionnaire [33], including the symptoms used to define the inattentive symptoms of the Attention Deficit Hyperactivity Disorder (ADHD) according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (APA, 2013). The original SNAP-IV uses four levels to evaluate each item, whereas in our study, the parents and the teachers evaluated each answer on a 3-level item Likert-type scale (*not true*, *somewhat true*, or *certainly true*) in order to follow the response pattern of the remaining scales included



in the first wave of the BCS. Each answer was assigned a value 0, 1, or 2. The nine inattention items from SNAP-IV are listed in Table 1.

**Table 1. SNAP items, scored as *not true* (0), *somewhat true* (1), and *certainly true* (2)).**

SNAP 1:	Often fails to give close attention to details or makes careless mistakes in schoolwork, work, or other activities
SNAP 2:	Often has difficulty sustaining attention in tasks or play activities
SNAP 3:	Often does not seem to listen when spoken to directly
SNAP 4:	Often does not follow through on instructions and fails to finish schoolwork, chores, or duties
SNAP 5:	Often has difficulty organizing tasks and activities
SNAP 6:	Often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort
SNAP 7:	Often loses things necessary for tasks or activities (e.g., toys, school assignments, pencils, books, or tools)
SNAP 8:	Often is distracted by extraneous stimuli
SNAP 9:	Often is forgetful in daily activities

The total score across the nine SNAP-IV items was statistically significant higher in boys ( $M = 1.67$  ( $SD = 2.8$ )) than in girls ( $M = .59$  ( $SD = 1.6$ ),  $t(1776.099) = 11.36$ ,  $p < .001$ ). The percentages of score values within the three SNAP-IV response categories are given in Table 2, showing that the frequency of girls reported with no inattention (*not true*) was significantly higher than in boys.

**Table 2. Percentage of children obtaining a given response from their teachers on each inattention item (SNAP-IV).**

	<i>Not true</i>			<i>Somewhat true</i>			<i>Certainly true</i>		
	All (%)	Girls	Boys	All (%)	Girls	Boys	All (%)	Girls	Boys
SNAP 1	86.7	91.1	81.9**	11.3	7.6	15.4	1.9	1.3	2.7
SNAP 2	88.3	93.9	82.1**	9.6	5.6	14.0	2.1	.5	3.9
SNAP 3	91.8	96.6	86.6**	7.6	3.2	12.4	.6	.2	1.1
SNAP 4	92.5	96.2	88.4**	6.8	3.6	10.4	.7	.2	1.1
SNAP 5	91.4	95.9	86.3**	7.3	3.6	11.5	1.3	.5	2.2
SNAP 6	91.6	96.2	86.5**	7.1	3.4	11.1	1.3	.4	2.4
SNAP 7	96.5	98.5	94.2**	3.0	1.2	5.1	.5	.3	.7
SNAP 8	74.8	84.3	64.4**	21.3	14.3	29.0	3.9	1.4	6.6
SNAP 9	89.4	93.3	85.0**	9.5	6.3	13.1	1.1	.4	1.9

*Note:* Total number of children = 2397, girls = 1256, boys = 1141. \*\*:  $p$  value  $< .001$  according to a chi-square test comparing a “not true” report in boys and girls.

## Academic achievement

Academic achievement scores were provided by the official registers from the Hordaland County. In Norway, secondary schools use a scale spanning from 1 to 6, with 6 being

the highest grade (outstanding competence), 2 the lowest passing grade (low level of competence), and 1 being a *fail*. The academic achievement scores that were available to our study was the mean value of the grades during the previous semester, comprising all school subjects except for physical education (gym). The mean score for girls was statistically significant higher ( $M = 4.13$  ( $SD = .72$ )) than for boys ( $M = 3.91$  ( $SD = .72$ ),  $t(2371.956) = 7.23$ ,  $p < .001$ ). For the present study, the academic achievement scores were categorised into three levels, calculated to generate groups with similar number of participants (see details below).

## Statistical analysis

The data analysis was divided into three parts: (a) data preparation, (b) explorative data analysis, and (c) pattern classification using two different multivariate statistical methods (MLR and CART). To perform these steps, explained in detail below, we used R (ver. 3.2.3) with selected packages in the RStudio environment, with an exception for Fig. 3 where MATLAB (R2015b) was used. The R markdown notebook, implementing our analysis, will be available on GitHub [address TBA].

## Data preparation

The original data provided to us as a SPSS-file was imported into the R environment. For the analysis we used the sample of  $n = 2397$  children having complete data on the 11 predictor variables and academic achievement as outcome variable as illustrated in Fig. 1.

For classification purposes, the mean average academic achievement scores were discretised into three intervals (level of academic achievement) constructed such that prior probability was non-informative, i.e. about the same number of participants in each of the categories; *low* ( $[1.000 - 3.750]$ ,  $n = 779$ ), *medium* ( $[3.750 - 4.429]$ ,  $n = 818$ ), and *high* ( $[4.429 - 5.900]$ ,  $n = 800$ ). The distribution across the three levels - from low to high - was 39.2%, 34.1% and 26.7% for boys, and 26.4%, 34.2% and 39.4% for girls, confirming the overall higher academic scores achieved by the girls. Depiction of the data values (gray scale heatmaps) and the classification methods being used are given in Fig. 2.

## Multinomial logistic regression model (MLR)

The multinomial logistic regression analysis included the following set of variables on a nominal level: the three levels of academic achievement scores as the outcome variable, and teacher reports on the nine inattention items, gender and primary school class level as predictors. Generally, the multinomial logistic regression model relates a set of explanatory variables  $x_1, \dots, x_p$  to a set of log-odds,  $\log(\pi_2/\pi_1), \dots, \log(\pi_J/\pi_1)$  according to

$$\log(\pi_j/\pi_1) = \beta_{j0} + \beta_{j1}x_1 + \dots + \beta_{jp}x_p \quad (1)$$

for  $j = 2, \dots, J$ . Here,  $j = 1$  represents the base level category,  $\pi_j = P(\text{academic achievement level} = j)$ ,  $\pi_j/\pi_{j'}$  denotes the odds of category  $j$  relative to  $j'$ , and  $\sum_{j=1}^J \pi_j = 1$  (see e.g. [6] for details). In our case, we let the base level category  $j = 1$  be the *low* mean academic achievement. For computations we used the `mlogit()` function in the R package **mlogit**.

## Classification trees (CART)

The nine Snap-IV items were included together with demographics (*Gender* and *primary school glass level (grade)*) as predictor variables in a classification and regression tree analysis (CART) [7] used to predict level of academic achievement score {*low, medium, high*}.

In brief, the *root* of the classification tree is the top node and input patterns are passed down the tree such that decisions are made at each node until a terminal node (a *leaf*) is reached. At each non-terminal node a question is posed on which a binary split is made such that the “child” nodes are on average “purer” than their “parent”. A measure of “impurity” is low (i.e. close to 0) if the probability of the occurrence of a class at a given node for all subsets of patterns reaching that node is concentrated on that class. The “impurity” is maximal if the class probabilities at that node is uniform. A common algorithmic approach is to use the *Gini index* as a measure of impurity, which can be interpreted as the expected error rate if the class label is chosen randomly from the class distribution at that node [29,32]. In our analysis we used the **rpart** package in R for growing the classification tree (cf. Fig. 3).

## Results

### Multinomial logistic regression model (MLR)

The chi-square test, indicating how much new variance is explained by the baseline 0-model, is statistically significant (255.7,  $p < .001$ ).

Gender significantly predicted whether a child obtained a low rather than a high academic achievement score in high school ( $b = -.49$ ,  $p < .001$ ) as well as a low rather than a medium score ( $b = -.23$ ,  $p = .003$ ). The odds ratio tells us that going from a female to a male reduced the odds of obtaining a low compared to a medium (.80) and a high (.61) score. In other words, boys were overall somewhat more likely to obtain a low academic achievement level in high school than the girls.

Two of the teacher reported inattention items significantly predicted a low rather than a medium academic achievement score. The strongest effect was found for an item reflecting problems related to sustained attention (SNAP2,  $b = -.65$ ,  $p < .001$ ). An odds ratio of .52 tells us that for each unit change in the score given by the teacher, the child was almost two times less likely to obtain a medium compared to a low academic achievement score ( $1/.52 = 1.9$ ). The second item reflects distractibility (SNAP8,  $b = -.29$ ,  $p = .002$ ), with an odds ratio of .75 leaving the child with a somewhat increased odds (1.3) of obtaining a low score.

The predictions from these two items were even stronger when comparing high to low academic achievement scores, with the highest estimate of the sustained attention item (SNAP2,  $b = -.96$ ,  $p < .001$ ) followed by the distractibility item (SNAP8,  $b = -.60$ ,  $p < .001$ ). The odds ratios show that the child was 2.6 times more likely to obtain a low score in high school for each more severe step in problems reported on SNAP2 (OD = .38) and 2.4 times more likely for each step on SNAP8 (OR = .41).

The prediction of low rather than high academic achievement score was also significant for two other items reflecting problems related to sustained attention, SNAP1 ( $b = -.58$ ,  $p < .001$ ) and SNAP6 ( $b = -.66$ ,  $p = .002$ ). With ORs of .56 and .52, the increase was almost twofold (1.9 and 1.8, respectively).

To sum up the results from the MLR, inattentive behaviour associated with problems related to sustained attention and distractibility predicted low rather than high academic achievement levels in high school, with an overall higher odds-ratio in boys than in girls.

Table 3. Multinomial logistic regression model.

	Estimate	SE	z	P> z	2.5%	OR	96.5%
1:(intercept)	0.86	0.20	4.20	0.00	1.58	2.35	3.50
2:(intercept)	0.93	0.21	4.45	0.00	1.68	2.53	3.80
1:gender	-0.23	0.11	-2.14	0.03	0.65	0.80	0.98
2:gender	-0.49	0.11	-4.48	0.00	0.50	0.61	0.76
1:grade	-0.16	0.07	-2.51	0.01	0.75	0.85	0.96
2:grade	-0.10	0.07	-1.52	0.13	0.79	0.90	1.03
1:snap1	-0.05	0.14	-0.40	0.69	0.72	0.95	1.24
2:snap1	-0.58	0.18	-3.20	0.00	0.39	0.56	0.80
1:snap2	-0.65	0.19	-3.34	0.00	0.36	0.52	0.76
2:snap2	-0.96	0.27	-3.59	0.00	0.23	0.38	0.65
1:snap3	0.02	0.20	0.12	0.90	0.70	1.02	1.50
2:snap3	-0.08	0.25	-0.32	0.75	0.57	0.92	1.50
1:snap4	0.17	0.24	0.70	0.49	0.74	1.18	1.89
2:snap4	0.11	0.32	0.33	0.74	0.59	1.11	2.09
1:snap5	0.04	0.22	0.19	0.85	0.68	1.04	1.60
2:snap5	0.47	0.27	1.70	0.09	0.93	1.59	2.73
1:snap6	-0.17	0.20	-0.85	0.40	0.56	0.84	1.25
2:snap6	-0.66	0.29	-2.29	0.02	0.29	0.52	0.91
1:snap7	0.16	0.26	0.62	0.53	0.70	1.18	1.97
2:snap7	-0.19	0.43	-0.45	0.65	0.35	0.82	1.92
1:snap8	-0.29	0.13	-2.25	0.02	0.58	0.75	0.96
2:snap8	-0.60	0.15	-4.05	0.00	0.41	0.55	0.73
1:snap9	-0.06	0.16	-0.39	0.69	0.68	0.94	1.29
2:snap9	-0.24	0.20	-1.21	0.23	0.53	0.79	1.16

1: = medium versus low achievement; 2: = high versus low academic achievement  
SE = standard error; OR = Odds ratios.

## Classification trees (CART)

### Tree classification and feature importance

The CART analysis generated five terminal nodes (Fig. 3). The first split is identified on SNAP 2, assessing if the child has difficulty sustaining attention. Teacher reports of *somewhat true* or *certainly true* on this item are associated with a low academic achievement score at high school (node #3). A total of 12 % of the sample was allocated to this node, where only 9 % obtained the highest and 65 % the lowest achievement level. If teachers reported no problems on SNAP 2, there was a second split on SNAP 8, assessing distractibility. A teacher report of *somewhat true* or *certainly true* on this item allocates 43 % of the children to the lowest academic achievement level in high school (node #5, including 14 % of the sample).

For the children with *not true* reports on both the SNAP 2 and SNAP 8, 43% of the girls obtained the highest, 35% the medium and 23% the lowest grade within the terminal node #8 (including 44% of the full sample). Reports of no problems were also obtained by 30% of the boys, but their class level when evaluated by their teachers were of importance for their future academic achievement. If reported with no problems by their teachers in the fourth grade (7% of the sample), 44% obtained the highest academic achievement level (node #18), while this number was still 30% in the 24% of

boys evaluated without problems related to sustained attention and distractibility at lower class levels.

To sum up the results from the CART, reports of problems on items reflecting sustained attention and distractibility were strong predictors of a low academic achievement score in high school. This was true for 26% of the sample, reflecting that the majority of children were reported to have no problems related to these two items. Most of the girls within this group obtained academic achievement scores at the two highest levels, which was also true for the small subgroup of boys reported without problems when attending the 4<sup>th</sup> primary school class level.

## Discussion

### Summary of results

The present study asked if specific features of inattentive behaviour in primary school - as reported by teachers - act as predictors of academic achievement in high school. The time span between the two events was about 10 years, and two types of multivariate analyses were used to handle a set of categorical variables. Overall, primary school teachers reported inattentive behaviour in about 10% of the children - more frequently in boys than girls, and the statistical models selected items reflecting problems related to sustained attention and distractibility as primary predictors. The MLR showed the strongest effect when comparing adolescents allocated to the low and high academic achievement levels. In the CART analysis, the first two splits were made between children with teacher reports of no versus those with minor or severe problems on SNAP items reflecting sustained attention and distractibility, respectively. Problems on either of these two items strongly predicted low academic achievement scores in high school. Gender was identified as a strong predictor by the multinomial logistic regression analysis, with lower academic achievement scores in boys than in girls. The CART analysis showed that 44% of the girls and 31% of the boys were reported with no problems on both the sustained attention and distractibility items. The age when evaluated by the primary school teachers was of importance to future academic achievements in boys. In a small subsample of boys (7%), more than 40% reported with no problems when they attended the 4<sup>th</sup> primary school class level obtained a high achievement score, an achievement that was comparable to the one shown for girls with no problems on these items. Taken together, the results suggest that the majority of the boys reported with problems related to sustained attention and distractibility at this age were allocated to terminal nodes characterised by a predominance of low academic achievement scores in high school.

### Early predictors of academic achievement in high school

The main contribution of the present study was the finding of sustained attention and distractibility in primary school as strong predictors of high school academic achievement. By this, the results partly overlapped with findings previously reported in a study by Holmberg et al. [17], where teacher reports of failure to finish a task was found to be one of the main factors in explaining academic outcome. Our study add to this by revealing the importance of problems related to distractibility. The MLR showed that this problem was associated with an almost two-fold increased odds-ratio of an achievement score in the lower than higher end of the scale, and its importance as a predictor was clearly supported by the CART analysis. In a class situation, the relation between the two is obvious. A child with the ability to stay focused on a task over a longer period of time is expected to be less disturbed by habits and cues in the

environment than a child with poor vigilance. This enables the child to obtain the basic skills and knowledge that are of importance to the academic achievement scores as the curriculum become more complex at high school level.

Inattention is one of the core symptoms of ADHD, but inattentive behaviour are definitely not restricted to a diagnostic category. This is underscored by the results of the present study, showing that inattention may have a detrimental effects on future academic function even in a population-based sample. However, although a high proportion of children obtaining a low academic achievement score were reported as inattentive by their primary school teachers, the prediction accuracy across all academic achievement levels was still poor. This reflects both the instability of inattentive behaviour and the large number of co-existing and new challenges influencing a child through childhood and adolescence. Identification of these factors awaits further longitudinal studies.

Taken together, the present results should inspire assessment and treatment efforts in primary school children vulnerable to distractibility and problems to sustain their attention in their school-related work. The close relationship between inattentive behaviour and cognitive function [3, 4] have lead to the popularity of presenting cognitive training programs to school children with ADHD (see e.g., [27, 34]). A sole focus on cognitive training of the child is, however, not expected to lead to successful alleviation of the inattentive behaviour described in the present paper. This is supported by the conclusion in the meta-analysis presented by Cortese and collaborators [12], where cognitive training procedures were shown to have limited effects on ADHD symptoms when the assessment was based on blinded measures. Positive contributions from parents and teachers are essential (see e.g., [24]). Whereas parent-focused training produces improvements in negative parenting and impairment at home, incorporation of child skill training and teacher consultation may be necessary to produce improvements at school [15].

Gender turned out to be another important predictor. Girls were reported by their primary school teachers to have less inattention symptoms and to obtain higher academic achievement in high school than boys. Boys showed an increased odds ratio when compared to girls for obtaining the lowest compared to the highest academic achievement level. Although gender was identified as one of the main predictors of academic achievement scores in MLR, the CART analysis showed a more differentiated picture. Age at the time of primary teacher reports turned out to be an important factor for boys. If reported with no problems at the 4<sup>th</sup> primary school class level, more than 40% of the boys obtained the highest level of academic achievement. A similar effect of age was not found in girls, supporting that predictors of academic achievement may be different for girls and boys at different ages. Further gender balanced longitudinal studies of functional outcomes of early inattentive behaviour are therefore warranted.

## Strengths and Limitations

The large population-based sample of high-school students followed from childhood, inclusion of a standardised questionnaire assessing inattention, and inclusion of academic achievement scores from official National registers are main strengths of the present study. Another strength is the inclusion of statistical methods - the MLR and CART. By this we could handle predictor variables with few categories and possible hidden multidimensional relationships, and still generate behavioural patterns that was easy to interpret. Furthermore, the relevance of the present analytic approach is clearly not restricted to the topic of the present study, in that questionnaire data with only a few response categories are commonly used in psychological research.

In spite of the strengths and the importance of the findings generated, several limitations must be mentioned. Inclusion of very few predictor variables when assessing

predictions over a nine-years period as in the present study, is an obvious limitation of the present study. A stronger model could have been obtained by including results from psychometric test assessing vigilance and distractibility, similar to the one developed by Cassuto et al. [10], or more ecological valid virtual reality test as the one described by Pelham et al. [23]. Inclusion of teacher reports only may also be considered as a limitation. Although parent reports obviously are of importance, teacher reports were selected due to the focus on academic achievements. An even more sophisticated analysis would be to include repeated inattention measures to understand the trajectory from early symptoms of inattention to function in adolescence and adulthood. Its importance was demonstrated in a study by Pingault and collaborators [25], showing that increase in symptoms of inattention during childhood really matters when it comes to school graduation failure. Such studies are important and should include analysis of behavioural patterns, because a specific pattern of vigilance and distraction was suggested by the present study.

**Figure 1** The data organisation for the sample. *gender* is 0 (girls) or 1 (boys), *grade* is primary school class level, 2, 3, or 4. The three-level Likert items from Snap-IV: SNAP1, ..., SNAP9 each have three possible values: 0 = *not true*; 1 = *somewhat true*; 2 = *certainly true*)

**Figure 2** The predictor data (explanatory variables), the academic achievement outcome, and the types of classification analyses being performed. Data values are represented as grey level heat maps. See Fig. 1 for explanation of variables. MLR = Multinomial logistic regression, CART = Classification and regression trees.

**Figure 3** Fitted classification tree (CART analysis), including the predictor variables SNAP-IV items 1 to 9 (0 = *not true*; 1 = *somewhat true*; 2 = *certainly true*); gender (0 = *girl*; 1 = *boy*); grade (primary school class level 2, 3, 4) and the academic achievement outcome (H = *high*, L = *low*, M = *medium*, where the three occurrence frequencies in the node boxes are given in alphabetical order). The percentage in each node box denote the percentage of samples routed to that particular node - where the root node will contain 100% of the samples, and a leaf node will contain the least number of samples along a rooted path in the decision tree. The node numbers are given on top of each node box. For each split decision, '*yes*' denotes that the corresponding statement is *false* and then pointing to the left child node (that is either a new internal decision node or a final leaf node), and '*no*' denotes that the corresponding statement is *true* and then pointing to the right child node (that is either a new internal decision node or a final leaf node). The tree is plotted using the `fancyRpartPlot()` function from Graham Williams **rattle** package (<http://rattle.togaware.com>)

## Acknowledgement

The present study was supported by the Centre for Child and Adolescent Mental Health and Welfare, Uni health, Uni Research, Bergen, Norway, and was also funded by the University of Bergen, the Norwegian Directorate for Health and Social Affairs, and the Western Norway Regional Health Authority. We are grateful to the children, parents and teachers participating in the Bergen Child Study (BCS) and members of the BCS project group for making the study possible. A special thank to Professor Christopher Gillberg, who initiated this study.



## References

1. Stephen P. Becker, Aaron M. Luebke, Paula J. Fite, Leilani Greening, and Laura Stoppelbein. Oppositional defiant disorder symptoms in relation to psychopathic traits and aggression among psychiatrically hospitalized children: ADHD symptoms as a potential moderator. *Aggressive Behavior*, 39(3):201–211, 2013b.
2. C. J. Bellanti and K. L. Bierman. Disentangling the impact of low cognitive ability and inattention on social behavior and peer relationships. conduct problems prevention re search group. *Journal of Clinical Child Psychology*, 29(1):66–75, Mar 2000.
3. Itai Berger, Anna Remington, Yael Leitner, and Alan Leviton. Brain development and the attention spectrum. *Frontiers in Human Neuroscience*, 9:23, 2015.
4. Itai Berger, Ortal Slobodin, Merav Aboud, Julia Melamed, and Hanoch Cassuto. Maturational delay in ADHD: evidence from CPT. *Frontiers in Human Neuroscience*, 7:691, 2013.
5. Anne S. Berry, Elise Demeter, Surya Sabhapathy, Brett A. English, Randy D. Blakely, Martin Sarter, and Cindy Lustig. Disposed to distraction: genetic variation in the cholinergic system influences distractibility but not time-on-task effects. *Journal of Cognitive Neuroscience*, 26(9):1981–1991, Sep 2014.
6. Chrisopher R. Bilder and Thonas M. Loughin. *Analysis of Categorical Data with R*. CRC Press, 2015.
7. Leo Breiman, Jerome Friedman, Charles J. Stone, and R. A. Olshen. *Classification and Regression Trees*. Taaylor and Francis, 1984.
8. Regina Bussing, Bonnie T Zima, Faye A Gary, and Cynthia Wilson Garvan. Barriers to detection, help-seeking, and service use for children with ADHD symptoms. *The journal of Behavioral Health Services & Research*, 30:176–189, 2003.
9. Matthew R Capriotti and Linda J Pfiffner. Patterns and predictors of service utilization among youth with ADHD-predominantly inattentive presentation. *Journal of Attention Disorders*, page 1087054716677817, January 2017.
10. Hanoch Cassuto, Anat Ben-Simon, and Itai Berger. Using environmental distractors in the diagnosis of ADHD. *Frontiers in Human Neuroscience*, 7:805, 2013.
11. Laura L. Connors, Jennifer Connolly, and Maggie E. Toplak. Self-reported inattention in early adolescence in a community sample. *Journal of Attention Disorders*, 16(1):60–70, Jan 2012.
12. Samuele Cortese, Maite Ferrin, Daniel Brandeis, Jan Buitelaar, David Daley, Ralf W Dittmann, Martin Holtmann, Paramala Santosh, Jim Stevenson, Argyris Stringaris, Alessandro Zuddas, Edmund J S Sonuga-Barke, and European ADHD Guidelines Group (EAGG). Cognitive training for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *Journal of the American Academy of Child and Adolescent Psychiatry*, 54:164–174, March 2015.



13. Annie A. Garner, Briannon C. O'Connor, Megan E. Narad, Leanne Tamm, John Simon, and Jeffery N. Epstein. The relationship between adhd symptom dimensions, clinical correlates, and functional impairments. *J Dev Behav Pediatr*, 34(7):469–477, Sep 2013. 403–406
14. Sarah A O. Gray, Alice S. Carter, Margaret J. Briggs-Gowan, Stephanie M. Jones, and Robert L. Wagnmiller. Growth trajectories of early aggression, overactivity, and inattention: Relations to second-grade reading. *Developmental Psychology*, 50(9):2255–2263, Sep 2014. 407–410
15. Lauren M Haack, Miguel Villodas, Keith McBurnett, Stephen Hinshaw, and Linda J Piffner. Parenting as a mechanism of change in psychosocial treatment for youth with ADHD, predominantly inattentive presentation. *Journal of Abnormal Child Psychology*, September 2016. 411–414
16. Einar Heiervang, Kjell M. Stormark, Astri J. Lundervold, Mikael Heimann, Robert Goodman, Maj-Britt Posserud, Anne K. Ullebø, Kerstin J. Plessen, Ingvar Bjelland, Stein A. Lie, and Christopher Gillberg. Psychiatric disorders in norwegian 8- to 10-year-olds: an epidemiological survey of prevalence, risk factors, and service use. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46(4):438–447, Apr 2007. 415–420
17. Kirsten Holmberg and Sven Bölte. Do symptoms of ADHD at ages 7 and 10 predict academic outcome at age 16 in the general population? *Journal of Attention Disorders*, 18(8):635–645, Nov 2014. 421–423
18. Steve S Lee and Stephen P Hinshaw. Predictors of adolescent functioning in girls with attention deficit hyperactivity disorder (ADHD): the role of childhood ADHD, conduct problems, and peer status. *Journal of Clinical Child and Adolescent Psychology*, 35:356–368, September 2006. 424–427
19. Astri J. Lundervold, Steinunn Adolfsdottir, Helene Halleland, Anne Halmøy, Kerstin Plessen, and Jan Haavik. Attention network test in adults with ADHD—the impact of affective fluctuations. *Behavior and Brain Function*, 7:27, 2011. 428–431
20. Lindsay A Metcalfe, Elizabeth A Harvey, and Holly B Laws. The longitudinal relation between academic/cognitive skills and externalizing behavior problems in preschool children. *Journal of Educational Psychology*, 105:881–894, August 2013. 432–434
21. Jeneva L Ohan and Troy A W Visser. Why is there a gender gap in children presenting for attention deficit/hyperactivity disorder services? *Journal of Clinical Child and Adolescent Psychology*, 38:650–660, September 2009. 435–437
22. Elizabeth B Owens, Christine Zalecki, Peter Gillette, and Stephen P Hinshaw. Girls with childhood adhd as adults: Cross-domain outcomes by diagnostic persistence. *Journal of Consulting and Clinical Psychology*, April 2017. 438–440
23. William E Pelham, Jr, Daniel A. Waschbusch, Betsy Hoza, Elizabeth M. Gnagy, Andrew R. Greiner, Susan E. Sams, Gary Vallano, Antara Majumdar, and Randy L. Carter. Music and video as distractors for boys with ADHD in the classroom: comparison with controls, individual differences, and medication effects. *Journal of Abnormal Child Psychology*, 39(8):1085–1098, Nov 2011. 441–445
24. Linda J Piffner and Lauren M Haack. Behavior management for school-aged children with ADHD. *Child and Adolescent Psychiatric Clinics of North America*, 23:731–746, October 2014. 446–448

25. Jean-Baptiste Pingault, Sylvana M. Cote, Frank Vitaro, Bruno Falissard, Christophe Genolini, and Richard E. Tremblay. The developmental course of childhood inattention symptoms uniquely predicts educational attainment: a 16-year longitudinal study. *Psychiatry Research*, 219(3):707–709, Nov 2014.
26. Tinca J C. Polderman, D. I. Boomsma, M. Bartels, F. C. Verhulst, and A. C. Huizink. A systematic review of prospective studies on attention problems and academic achievement. *Acta Psychiatrica Scandinavica*, 122(4):271–284, Oct 2010.
27. Mark D. Rapport, Sarah A. Orban, Michael J. Kofler, and Lauren M. Friedman. Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? a meta-analytic review of cognitive, academic, and behavioral outcomes. *Clinical Psychology Review*, 33(8):1237–1252, Dec 2013.
28. Leslie Rescorla, Thomas M. Achenbach, Masha Y. Ivanova, Levent Dumenci, Fredrik Almqvist, Niels Bilenberg, Hector Bird, Anders Broberg, Anca Dobrea, Manfred Düpfner, Nese Erol, Maria Forn, Helga Hannesdottir, Yasuko Kanbayashi, Michael C. Lambert, Patrick Leung, Asghar Minaei, Mesfin S. Mulatu, Torunn S. Novik, Kyung-Ja Oh, Alexandra Roussos, Michael Sawyer, Zeynep Simsek, Hans-Christoph Steinhausen, Sheila Weintraub, Christa Winkler Metzke, Tomasz Wolanczyk, Nelly Zilber, Rita Zukauskienė, and Frank Verhulst. Epidemiological comparisons of problems and positive qualities reported by adolescents in 24 countries. *Journal of Consulting and Clinical Psychology*, 75(2):351–358, Apr 2007.
29. Brian D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996.
30. Julie Salla, Gregory Michel, Jean Baptiste Pingault, Eric Lacourse, Stephane Paquin, Cedric Galera, Bruno Falissard, Michel Boivin, Richard E Tremblay, and Sylvana M Cote. Childhood trajectories of inattention-hyperactivity and academic achievement at 12 years. *European Child and Adolescent Psychiatry*, 25:1195–1206, November 2016.
31. Børge Sivertsen, Nick Glozier, Allison G. Harvey, and Mari Hysing. Academic performance in adolescents with delayed sleep phase. *Sleep Medicine*, 16(9):1084–1090, Sep 2015.
32. Carolin Strobl, James Malley, and Gerhard Tutz. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, 14(4):323–348, Dec 2009.
33. J.M. Swanson. *School-based assessments and interventions for ADD students*. KC publishing, 1992.
34. Leanne Tamm, Jeffery N Epstein, Richard E A Loren, Stephen P Becker, Sarah B Brenner, Morgan E Bamberger, James Peugh, and Jeffrey M Halperin. Generating attention, inhibition, and memory: A pilot randomized trial for preschoolers with executive functioning deficits. *Journal of Clinical Child and Adolescent Psychology*, pages 1–15, January 2017.

Figure 1

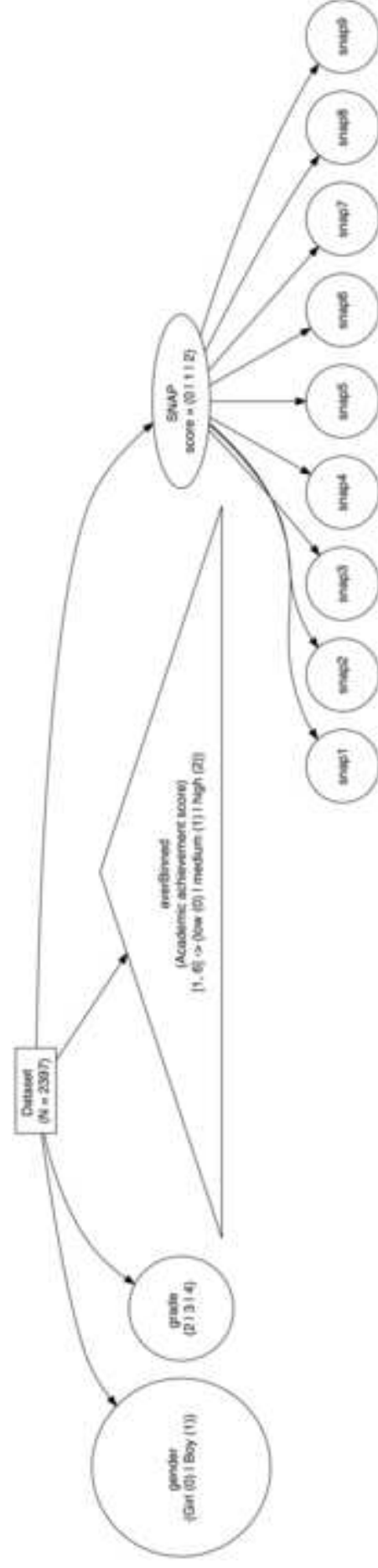
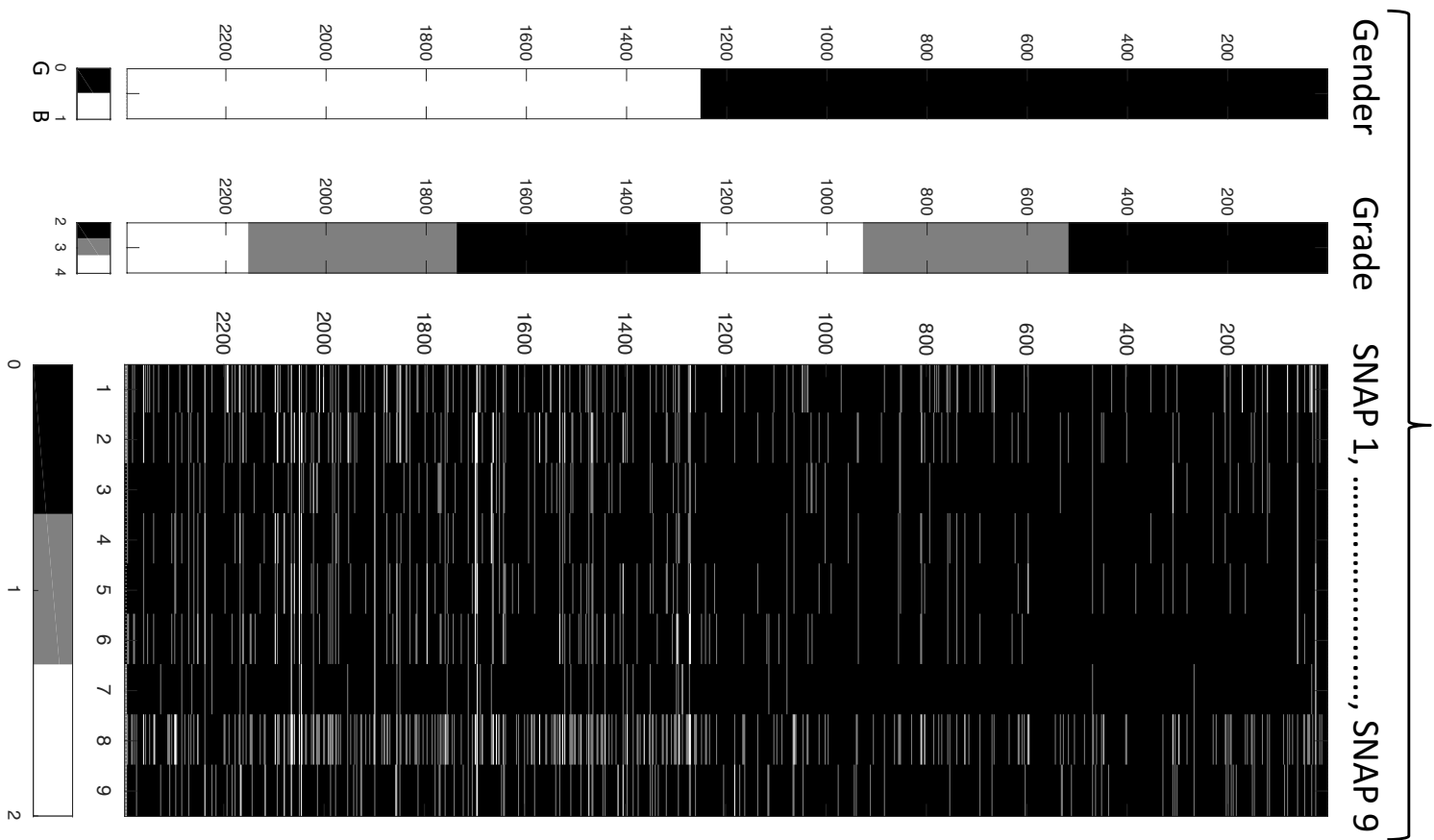


Figure 2

DATA (predictors)



[Click here to download Figure Fig2\\_data\\_to\\_classes.pptx.pdf](#)

OUTCOME

Academic achievement

Continuous    Discretized

Prediction  
MLR  
CART

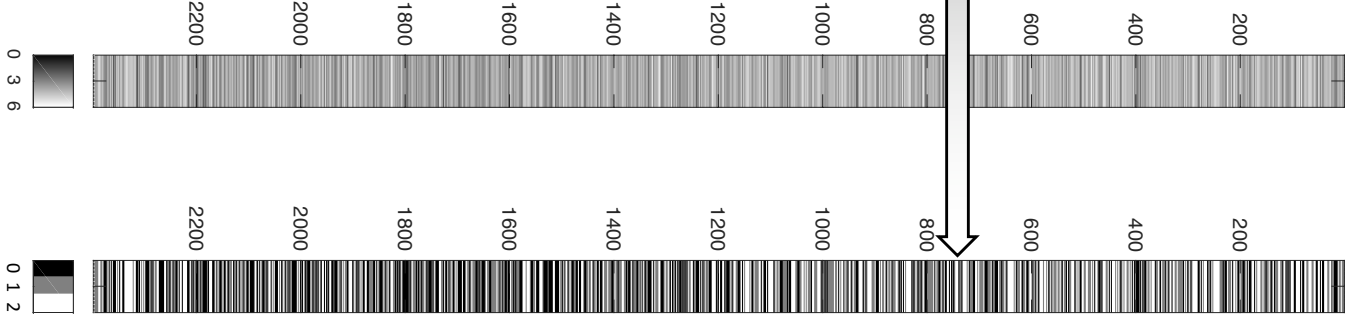


Figure 3

