

PROBLEM STUDENTS: A CONTEXTUAL PHENOMENON?

CORA J.M. MAAS

University of Utrecht, The Netherlands

G. WIM MEIJNEN

University of Amsterdam, The Netherlands

The number of referrals to different types of special education has been increasing for years in the Netherlands. In particular the number of children with Learning and/or Behavioral Difficulties (LBD) has increased sharply. The teacher usually describes them as students with problems or with special educational needs. This study investigates which factors influence teachers' judgement in this respect, with a particular focus on contextual factors at class level. The study looked at 2,340 students from 114 primary school classes in the Netherlands. Multilevel analyses show that 13% of the total variation in teachers' judgement is group (context) level variance and 87% individual (student) level variance. Contextual factors which influence teachers' judgement are: the average level of achievement in the class, the amount of disruptive behavior in the whole teaching group and the percentage of boys in the class. Various characteristics of individual teachers turn out to be irrelevant.

In the Netherlands, the number of students referred to institutions for special education is still increasing. In this country, about 5% of the total population in primary education attends schools for special education. The largest groups comprise educable mentally retarded (40%) and children with learning and/or behavior difficulties (40%). Provisions for the latter group alone account for nearly all the growth in the special education system of the Netherlands. (Meijer, Pijl & Hegarty, 1994).

In the Netherlands, the percentage of children who fail to meet the standards for clear physical reasons has remained more or less constant over time. In contrast,

Cora J.M. Maas, Department of Methodology and Statistics, University of Utrecht, The Netherlands and G. Wim Meijnen, SCO-Kohnstamm Institute, University of Amsterdam, The Netherlands.

Please address correspondence and reprint requests to: Cora J.M. Maas, Ph.D., Department of Methodology and Statistics, University of Utrecht, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands.

the percentage of children who are judged to be 'deviant' due to learning and/or behavioral difficulties has shown a continuous rise. This phenomenon was observed also in the U.S. (Thurlow, Christenson & Ysseldyke, 1989). Peterson (1996) presents various studies which indicate an increasing prevalence of students with Educational and Behavioral Difficulties (EBD) in that country. Since, in making a diagnosis, it is difficult to differentiate between EBD and Learning Difficulties (LD), and between the latter and lighter forms of Mental Retardation (MR) (Fletcher, et al., 1994; Macmillan, 1996), Peterson notices also a rise when the different groups are merged under the term 'Learning and/or Behavioral Difficulties (LBD)'. These problems in defining exactly what LBD-students are lead to a variation of criteria in time and space. Research in the Netherlands has shown that there are regional differences between special schools for children with learning and behavioral problems as regards the 'identity' of their students (Van der Wissel, 1984; Meijer, Pijl & Rispens, 1986). Peterson (1996) indicates that in the U.S. there is an ongoing discussion among scientists, and between specialists from the field, policymakers and scientists, concerning the definitions of - and the criteria for - the groups which fall under the LBD category.

Consequently, teachers have a great deal of freedom to give their own interpretation to LBD. Most of the time they do not use this term, describing these students as 'problem students', 'students with special educational needs' or 'students at risk'. Specialist diagnosis then establishes with exactly which kind of learning and behavioral problems the teacher is confronted.

The fact that, due to a lack of univocal criteria, the teacher has the freedom to let his/her own insights prevail in perceiving problem students, makes it probable that there will be some degree of teacher variance regarding the above-mentioned judgment. In other words, comparable student behavior will not be categorized in the same way by the teachers. This variance has three sources, namely individual teacher characteristics, context characteristics and the interaction between these sources. Individual teacher characteristics are: age, experience, gender, ethnicity, etc. Class characteristics such as class size, social composition of the class, achievement level of the class, etc., and also school characteristics, such as school culture, school size, social composition of the school, aptitude level of the students, composition of the teaching staff etc., are regarded as contextual characteristics. So it is expected that experienced teachers will be more able to handle problem students, and therefore, will perceive deviant behavior less often as a problem. The same experienced teacher, however, will tend to perceive deviant behavior more often as a problem in a class with 30 students than in one containing only 20 students, because of the increased working pressure in the former. Of course, there can also be interactions between contextual factors and individual characteristics.

In this study, we investigate which factors lead the teacher to define a child as 'a problem'. In particular, we focus on the influence of contextual factors at

classroom level. We anticipate that the teacher's judgement of the individual students is partly dependent on the characteristics of the group which s/he is teaching (size, uniformity, level etc.). The research question addressed by this study is therefore: *what are the effects of class characteristics on teachers' perception and labelling of children as "problem students"?*

To be able to trace the net effect of these contextual factors one must first examine what influence the characteristics of individual students have. A large body of research shows that teachers' judgement is chiefly influenced by 'objective' child characteristics such as low intelligence, disruptive behavior etc. If, in a contextual analysis, one does not control for the effect of individual student characteristics, too much variation will wrongly be ascribed to the effect of context on teachers' judgment.

BACKGROUND

CHILDREN WITH LEARNING AND BEHAVIORAL DIFFICULTIES

According to Peterson (1996) students who face problems in regular educational settings can be classified in groups with (a) learning disabilities, (b) speech or language impairments, (c) mental retardation, (d) physical disabilities (deafness, visual impairments, etc.) and (e) emotional and behavioral difficulties.

In the Netherlands, 80% of the students with special educational needs who are referred to special education start their educational careers in regular schools. Teachers in primary schools are not confronted by children with severe physical or mental disabilities. As these children are directly referred to special schools when they reach the age of four years. However teachers in regular schools are confronted with children with learning disabilities, emotional and behavioral difficulties, reading difficulties (e.g. dyslexia) and moderate degrees of mental retardation.

Combinations of factors often cause learning problems. So, a committee of the Health Council of the Netherlands (Gersons-Wolfensberger & Ruijssenaars, 1997) has recently published a report about the definition and treatment of dyslexia. When both single presentation (dyslexia alone) and complex presentation forms, and both severe and lighter disorders are included in the definition of dyslexia, then, according to the committee mentioned above, the estimates regarding the occurrence of dyslexia can vary between 2% and 10%. As regards Emotional and Behavioral Difficulties, with the same variation in criteria, the estimates of the prevalence in the U.S. vary between 2% and 25% of the total school population (Peterson, 1996). Many of these students are, however, placed in learning disability programs because of the reluctance of schools to classify children as emotionally disturbed (Lopez, 1996).

Progress in educational achievement represents the core task of the school. In primary education this mainly concerns children's progress in basic skills: reading, writing and number work. The lack of progress in educational achievement often goes hand in hand with behavioral problems. Finally there are, of course, students with fair educational achievements, but who exhibit disruptive behavior in the classroom to the extent that they become problem students for their teacher. Brophy and Good (1974) asked teachers who taught in the second grade of primary education, and other teachers, to indicate which types of students they hoped would leave the school. The teachers circumscribed these as aggressive, impulsive, impudent and lazy. Brophy and Rohrkemper (1981) distinguish no fewer than 12 types of 'difficult' or 'problem' students. The 12 patterns are defined so as to be mutually exclusive, although several could coexist in the same student. Out of these 12 types, only three directly indicate learning disabilities. The rest of them denote behavioral difficulties or combinations of difficulties.

Helton and Oakland (1977) observed teachers in primary schools - examining their feelings of preference, concern or rejection with respect to their students. The teachers tended to prefer rigid, conforming, orderly, passive and dependent children. The children who were rejected were flexible, non-conforming and aggressive. Safran and Safran (1985) found that teachers condemn behavior directed against other students or against the teacher more than behavior which is restricted to the student him/herself (inattention, laziness). Merrett and Wheldall (1984) reported that teachers have the most serious problems with students who molest their fellow students, or who frequently chat with their classmates. Hutton (1984) asked female teachers in primary schools which kind of student behavior worried them most. Aggressive interaction and frightened response on the part of other students were reported most often.

More recently, the British government commissioned a report in which 1,200 primary school teachers were asked which kind of problem behavior they were confronted with, and how frequently (Elton Report, 1989). The teachers reported no fewer than 14 types of disruptive behavior. The most frequent were: (a) disruptive talking, chattering and shouting, (2) disturbing other students, (3) making unnecessary noise, (4) physical aggression against fellow students, (5) wandering around without permission, (6) avoiding school work, (7) exhibiting rough or wild behavior.

A lack of progress in learning and 'acting out' appear to teachers to be the worst problems. Very often these two problems coincide. The question of cause and effect cannot yet be answered (McMichael, 1979; Mortimore et al., 1988; Sturge, 1982). Boys demonstrate very often this type of coincidence. Badly behaved boys are mostly low achievers (Hudley, 1993). As a consequence, the gender ratio of students attending schools for children with learning and/or behavioral problems

in the Netherlands is 3.0. Girls are a minority in this type of education (CBS, 1995).

THEORETICAL ASSUMPTIONS ABOUT SCARCITY AND STANDARDS

Teachers try to give their students the opportunity to achieve according to their abilities, however, students are not all the same and teachers have to make choices as to how they divide their time and attention among their (individual) students. Time and attention, however, are scarce commodities according to utility theory (Brown & Saks, 1987), which asserts that teachers judge the outcome of their actions in terms of the performance of their (individual) students.

From the literature on teacher thinking (Letiche, Van der Wolf & Plooy, 1991) we know also that teachers' behavior is not guided exclusively by academic arguments. They use methods of working and routines which are geared to making or keeping complex situations manageable. In this context, one can speak of 'situated cognition' (Brown, Collins & Duguid, 1989). Technical know-how is considered useful only when it fits into the existing repertoire of teaching methods. Academic knowledge is introduced in the daily context of teaching. So, contextual factors also will play an important role in numerous judgments, decisions and actions of the teacher. This holds true also for the definition of what is experienced as a problem, as well as for the choice of how to deal with a difficult student. In this study, we confine ourselves to contextual factors as they relate to the diagnosis of problem behavior.

Contextual factors at classroom level determine, *mutatis mutandis*, the amount of time and attention given by the teacher to the individual students. Under more favorable circumstances, the amount of time and attention given by the teacher to potential problem students will increase. The assumption is then that the teacher will be less inclined to define his/her students as problem students.

Teachers also gear their standards with regard to the achievement which must be attained by their students to what they themselves can achieve by putting an average amount of effort into the class. In a class with students who are expected - or have been proved - to achieve well, the teacher's standards will be high. The teacher's judgment of which students should be considered as problem students will therefore depend also on the standard developed by the teacher for the whole group.

The present study verifies that, from this line of argument, a number of contextual factors at class level will influence teachers' judgment as to which students should be considered as problem students. These factors are: (a) *mean level of intelligence*, (b) *mean level of achievement*, (c) *mean level of disruptive behavior*, (d) *mean level of learning ability*, (e) *mean level of parental pressure*, (f) *variation in school achievement*, (g) *variation in children's cultural background*, (h) *class size*. *Mean intelligence level of the class*. If a class contains a large number of intelligent

students, the teacher will not have to put as much effort in teaching as s/he would in a class where the students have great problems in understanding the instruction. Consequently, teachers in classes with fewer intelligent students will raise the achievement standard and will, *mutatis mutandis*, be quicker to label the less intelligent as problem students.

Mean achievement level of the class. Since teachers are generally not acquainted with the IQ level of their students, the effect of the mean intelligence level could be overshadowed by the achievement level of the students. Since there is no one-to-one relationship between IQ and educational achievement, it is advisable to examine the achievement level separately. Again, the hypothesis is that, *mutatis mutandis*, in high achieving classes, teachers will be quicker to label lower achievers as problem students.

Mean level of learning ability. The amount of attention given by teachers to their individual students will depend on the expected learning ability. Teachers' expectation with regard to their students' learning ability is influenced by characteristics ascribed by the teacher to the individual student - such as motivation, intelligence, perseverance etc. When a teacher notices that many students in his/her class have high learning abilities, *mutatis mutandis*, he/she will be quicker to label those students who have lower learning abilities as problem students.

Mean level of disruptive behavior of the class. In classes with large numbers of students showing disruptive behavior, teachers will base their standards on this, which means that, in this context, teachers will be less quick to label disruptive students as problems.

Mean level of parental pressure. Parents differ with regard to their expectations regarding their children's school achievement. They differ also with regard to the amount of pressure which they exert on the teacher to pay more attention to their children (West, 1993; O'Sullivan, 1992). Time spent on one child reduces the teaching time spent on the other students. Consequently, the attention received by a student from his/her teacher shows also the coefficient of pressure exerted on the teacher by the parents of the other students. Therefore, students whose parents do not demand any attention from the teacher have a greater chance of being labelled as problem students in classes where many parents demand extra attention for their children.

Variation in school achievement. Classes with a great deal of variation in abilities between children are more difficult to handle than classes which are more homogeneous in this respect. Other things being equal, the chance of being labelled as a problem student will be greater in a heterogeneous than in a homogeneous class.

Class size. In classes with larger numbers of students, teachers have fewer possibilities of giving extra time and attention to their individual students. Teachers indicate this explicitly in various studies dealing with this subject (Vaughn et al.,

1996). In classes with larger numbers of students, children have a greater chance of being labelled as problem students than in smaller classes.

Variation in cultural background. Children from different cultural backgrounds behave differently. If they come from lower SES families, they are often not very familiar with school practices, and if they belong also to an ethnic minority, then there will be language problems too. These factors hamper the instruction process. As a consequence, teachers in high poverty schools are often skeptical about the abilities of their students. Teachers' beliefs are generally consistent with their practices, even when the school's poverty level and the mean achievement level of the students are statistically controlled (Solomon, 1996; Stevens & Palincsar, 1992). So, in classes with children from an heterogeneous cultural background, students have a greater chance of being labeled as problem students.

The contextual characteristics mentioned above may be responsible, at least in part, for the variation found in teachers' judgement. If this can be proved, we will have some indications as to why 'objective' student characteristics such as IQ or achievement scores do not predict teachers' judgment very accurately.

METHOD

SAMPLE

In the Netherlands, students start primary education at the age of four. Primary education consists of 8 grades. Referral to schools for children with learning and behavioral difficulties usually takes place in grade 4, when the children are eight or nine years old. We, therefore, decided to investigate the influence of contextual factors on teachers' judgment in grade 4.

From a file of 566 schools in the Dutch province of Utrecht, a sample of 116 primary schools was drawn (schools were excluded if they were very small or had non-standard ways of grouping). From each school, one grade-4 class was selected. In some schools there were no classes with only grade-4 students, so, a 'vertically grouped class' including grade-4 students was selected. After removing school classes (and students) with too many missing variables, the total number of students was 2,340 in 114 classes. These students were tested by trained personnel.

INDEPENDENT VARIABLES

The *intelligence test* administered was a revised version of the Primary Mental Ability 4-6 test (PMA 4-6) of Thurstone and Thurstone. The test consists of four subtests: verbal meaning, number facility, spatial relations, perceptual speed. The information is processed into one IQ-score. The test is standardized (z-score) on this sample with a mean of 100 and a deviation of 15. The *achievement test* contains tests for arithmetic and language. Data reduction and standardization carried out

according to the same procedures, provide us with a variable with a mean score of 100 and a deviation of 15.

An interview was held with the grade-4 teachers to gather information about the characteristics of the students and the teachers themselves. The teacher was asked to make a judgment of the progress a student could make if he or she could receive more individual instruction or attention (*learning ability*). The teacher had to rate the student on a 5-point scale (would make very little progress: 9.4%, little progress: 10.3%, reasonable progress: 33.6%, much progress: 27.6%, a great deal of progress: 19.0%). The children were also assessed by the teacher for the amount of *disruptive behavior* they exhibited on a 5-point scale (a great deal: 3.2%, much: 8.5%, quite a lot: 18.8%, little: 27.6%, very little: 41.9%). A high score on this variable means little disruptive behavior. The teachers were also asked which parents had asked them to enhance the teaching given to their children (*parental pressure*). The parents who did this sometimes were given a score of one (5%), those who did this regularly were given a score of two (1%) and those who did this frequently were given a score of three (0.1%). The teachers were asked also which children came from abroad, and which parents were not native speakers (*ethnicity*) (9.4%) which kind of work the parents did (*socioeconomic status; SES*). The information was coded into a 6-point scale (unskilled and semi-skilled: 11.5%, skilled: 20.9%, lower administrative: 9.1%, self-employed: 9.4%, medium and higher administrative: 26.0%, higher professions: 23.2%).

The contextual variables *mean level of intelligence of the class*, *mean level of achievement of the class*, *mean level of disruptive behavior*, *mean level of learning ability at class level* and *mean level of parental pressure* were calculated by adding up the scores of the individual students and dividing the sum by their number. The contextual variable *variation in student achievement* is defined as the *standard deviation* of the variable achievement of the class.

Furthermore, the *age*, *teaching experience*, *additional educational training* and *gender* of the teachers were registered. Research into the influence of class characteristics often includes the aggregated scores of *gender*, *ethnicity* and *SES*. Although theoretically it is difficult to identify these variables, they have been included in the analysis in order to isolate their possible influence in relation to the theoretically-derived contextual factors differentiated in our study.

Finally, the data on the contextual variables *class size* and *variation in cultural background* were extracted from school records. The variation in cultural background of the students in one school is measured by the so-called "OVB-score" (Educational Priority Policy Score). We used the 3-point version of this score. In classes with a score of one, most children are native Dutch and come from middle and higher social economic strata. In schools with a score of two, most children have a 'blue collar' background, while in schools with a score of three, most children come from abroad and have poorly-educated parents. Of the

114 classes in our research, 64% had a score of one, 21% had a score of two and 15% had a score of three.

DEPENDENT VARIABLE

Finally, the teachers were asked to rate each student with regard to the extent in which he or she could be labelled as a problem student. We asked the teachers to rate all the students on a 10-point scale ranging from 'very difficult' to 'not difficult at all'. *We did not give them any instructions, hints or clues about one or another concept of difficult behavior* in an attempt to do justice to their own perceived reality. We did not impose any formal judgment upon the teachers. They were given the freedom to process their own experiences and judgments in the score given to each individual student.

ANALYSIS

To isolate the effects of contextual factors there had to be a control for effects at individual level. Otherwise, too much variance would be attributed to the contextual level. We therefore analyzed the model by introducing student characteristics first, followed by contextual factors and teacher characteristics.

In this type of research, one is dealing with hierarchically nested data. Children are grouped in classes, classes in schools and so on. If we want to analyze such data, a general linear model is not appropriate. When the linear model is applied to hierarchically-nested data, with the dependent variable on the lowest individual level, the nested data structure implies that the independent assumption of the linear model has been violated. The reason for this is that there are within-group correlations of residuals, since individuals within groups have things in common.

A solution to this problem is provided by the recently developed random coefficient models. In these models, not only variables at individual level, but also variables at higher levels are thought to be influential on the dependent variable. The variance of the dependent variable can be divided into two (or more) parts - one part which is attributed to differences between students, and one part which is attributed to differences between groups (Goldstein, 1987; Longford, 1988). For this analysis, we used the multilevel program VARCL developed by Longford.

RESULTS

HOW MANY PROBLEM STUDENTS?

First, we will introduce the results concerning the dependent variable: the teacher's judgment of the extent to which a child must be labelled as a problem student.

If we consider students with a rating of 5 or less as difficult, we see that almost 13% are labelled as such. Of these 13%, more than one half are slightly difficult

(score 5), 25% rather difficult (score 4) and one fifth very difficult (score lower than 4). The distribution of the number of students in a group who are considered as difficult by the teacher is given in table 2.

TABLE 1
TEACHER RATINGS ABOUT BEING A PROBLEM STUDENT

	Number	Percentage
1. very difficult	4	.2
2.	9	.4
3.	26	1.1
4.	73	3.1
5.	183	7.8
6.	487	20.8
7.	701	30.0
8.	551	23.5
9.	245	10.5
10. not difficult at all	61	2.6
Total	2340	100

TABLE 2
NUMBER OF PROBLEM STUDENTS IN A GROUP

	Number	Percentage
0	21	18.4
1	20	17.5
2	23	20.2
3	21	18.4
4	10	8.8
5	7	6.1
6	3	2.6
7	5	4.4
8	1	.9
9	1	.9
10	1	.9
11	1	.9
Total	114	100

Between groups, the number of problem students differs considerably. Twenty-one teachers think that they have no difficult students in their class. Twenty other teachers say they have only one difficult student. Classes with two or three difficult students also were found very frequently: 23 and 21 times respectively. The numbers 4 up to 7 occur less frequently, while the numbers 8 or more are rare.

INTELLIGENCE, ACHIEVEMENT AND BEHAVIOR OF PROBLEM STUDENTS

First, we will discuss the characteristics of the problem students (score 5 or less). If we compare the intelligence level of the problem students with that of the other students, we find that the former score significantly lower on intelligence.

The mean of the intelligence scores of the problem students in our study was 92.05, while the mean of the other students was 101.09 (t -value = -9.89, $p = .00$).

The achievement levels of the problem students were different too. Problem students were proved to achieve significantly lower than the other students. The mean achievement of the problem students was 84.06, while the mean achievement of the other students was 102.50 (t -value = -20.06, $p = .00$).

The learning ability of the problem students, rated on a 5-point scale, was also significantly lower (3,16) than that of the other students (2.55; t -value = 8.36, $p = .00$).

Disruptive behavior in the classroom also was rated on a 5-point scale. The mean of the problem students was 3.28, while the mean of the other students was 4.06. This difference is significant (t -value = -9.97, $p = .00$).

We see that problem students differ from the other students on all four variables mentioned above. The correlations between these four variables and teachers' ratings of the students with regard to difficult behavior is given in table 3. We did not use the dichotomous form of the dependent variable for this calculation, but instead used the original scores.

TABLE 3
CORRELATIONS BETWEEN THE VARIABLES: LABELING CHILDREN AS PROBLEM STUDENTS, INTELLIGENCE, ACHIEVEMENT, DISRUPTIVE BEHAVIOR AND LEARNING ABILITY

	rating	IQ	achiev.	disr.beh	learn.ab
rating	1.00				
IQ	.32**	1.00			
achievement	.59**	.47**	1.00		
disr. behav.	.32**	.08**	.19**	1.00	
learn. abil.	.25**	.10**	.16**	.13**	1.00

** $p < .01$

If we look at the results, we see that all student characteristics correlate moderately or strongly with teachers' ratings. The highest correlation is that between the scores on the achievement test and the ratings. This confirms the results from other studies. Teachers regard lack of progress as the greatest problem. However, the correlation with disruptive behavior is also marked, about the same as the correlation with the IQ-score. A (weaker) correlation with learning ability indicates that teachers have more faith in the benefits of giving good students more time and attention. The correlations among the four variables are weak however, with the exception of that between intelligence and educational achievement. These data lead to the conclusion that the cognitive dimension of problem behavior is seen as the most difficult aspect, but that the social dimension also is a significant factor in its own right.

BACKGROUND CHARACTERISTICS OF PROBLEM STUDENTS

To describe the differences in background characteristics we again divided the children into problem students (a score of 5 or less) and those who are not a problem (6 or more).

The SES of the problem students, rated on a 6-point scale, was significantly lower (3.23), than the mean of the other students (3.96 t -value = -6.65, $p = .00$). About 13% of the problem students came from abroad or had parents who were not native speakers, against 9% of the other students who came from abroad or had parents who were not native speakers. This difference is significant (Chi^2 -value = 4.66, $p = .03$).

If we compare the gender ratio of the problem students with that of the other students, we see that 59% of the problem students are boys, while 50% of the other students are boys. This difference is significant (Chi^2 -value = 8.74, $p = .00$). The ratio 3 (boys) to 2 (girls) is, however, considerably smaller than the gender ratio in special schools for children with learning and behavioral difficulties (3:1). It is possible that many 'difficult' boys had already been referred to special education in the lower grades. Another explanation would be that boys who are found to be difficult are referred more often than are girls with similar problems.

MULTILEVEL ANALYSIS

Because we wanted to test the influence of both individual and group characteristics we used the multilevel program VARCL developed by Longford. In the above-mentioned analysis the dependent variable (being a problem student) was dichotomized. As we did not want to lose any information, we used the original scores for the dependent variable in the multilevel analysis.

To construct the analysis in a logical way, it might be thought desirable to draw up a model of the effects of only individual characteristics first. Then, the effects at group level could be added subsequently. This, however, presents some problems - because we are working with the hypothesis that judging students is a process involving comparisons with the whole group. A model containing only individual characteristics produces results which differ sharply from those in a more all-embracing model. That is why we analyzed individual characteristics in conjunction with group characteristics.

To investigate the extent to which variation in teachers' judgment is caused by individual student characteristics, and how much of that variation arises from characteristics at group level we analyzed an empty model first. That showed that 87.4% of the variation is located at individual level and 12.6% at group level. So, about one eighth of the influences which affect teachers' judgment are caused, not by characteristics of the individual students, but by characteristics which belong to the context. That is comparable to, or slightly higher than, the percentage figures for variation in educational achievement between different schools in the Netherlands (Scheerens, 1992).

Subsequently, the data on individual student characteristics and aggregated group characteristics were entered into the empty model (table 4).

TABLE 4
THE INFLUENCE OF INDIVIDUAL AND AGGREGATED STUDENT CHARACTERISTICS ON LABELLING CHILDREN AS PROBLEM STUDENTS

Individual effects		
- intelligence	.01*	(.08, .00)
- achievement	.05*	(.57, .00)
- learning ability	.18*	(.15, .02)
- disruptive behavior	.21*	(.16, .02)
- SES	.08*	(.10, .01)
- ethnicity	-.05	(-.01, .09)
- gender	.05	(.02, .04)
Aggregated effects		
- mean intelligence	-.01	(-.05, .01)
- mean achievement	-.05*	(-.23, .01)
- mean learning ability	-.10	(-.05, .07)
- mean disruptive behavior	.20*	(.07, .10)
- mean SES	.11	(.07, .06)
- mean ethnicity	.18	(.03, .31)
- mean gender	-.81*	(-.07, .36)
R ² individual	48%	
R ² intercept	22%	
Deviance	6558.03	
Df	14	

* $p < .05$

between brackets: beta and standard errors

The variables intelligence, achievement, learning ability and disruptive behavior were proved to have significant effects. This means that students who are more intelligent are higher achievers, demonstrate less disruptive behavior, have higher learning abilities, and have less chance of being labelled as problem students. Each variable shows an independent effect, which has been controlled for other effects. The results in table 3 remain essentially valid.

When we enter also background variables into the analysis, we see that only the SES variable has a significant, positive effect. This means that children from higher SES families, other things being equal, have less chance of being perceived as problem students by their teachers.

Of the aggregated *theory based, contextual* characteristics which were distinguished, the *average achievement level of the class* and the *level of disruptive behavior* are relevant. From the accepted proxies, the percentage of boys is significant. Of course, this was again controlled for other effects. The higher the achievement level of the class, the greater is the risk that an individual student will be labeled a problem student. That confirms the hypothesis which was formulated. The hypothesis on disruptive behavior as a group characteristic also was confirmed:

the less disruptive the class in the eyes of the teacher, the less chance a child has to become a problem student.

Somewhat surprising is the fact that the chance of being labelled as a problem student increases in classes with a higher percentages of boys. This appeared to be true after controls were entered for the individual characteristic of gender and the group characteristic of disruptive behavior. Evidently, as the number of boys in a class rises, processes come into play which the teacher experiences as a nuisance, causing him/her to label his/her students as problem students more readily. In contrast, the socioeconomic and ethnic composition of the class does not affect teachers' judgment significantly, if one controls for factors such as average achievement level and level of disruptive behavior. Neither are intelligence scores at group level significant. That can probably be attributed to the fact that the variable 'average educational performance' has already taken up much of the corresponding variation. Finally, at class level, estimated learning abilities appear not to have an influence either. Here too, one could argue that the variable 'average achievement level' accounts for this.

The contextual variables analyzed above are all group characteristics based on the aggregation of individual student scores. Some of them belong to the category which was theoretically derived in the present study, others are proxies which are often used in the literature and which, therefore, have been included as some kind of control variables. The table below (table 5), adds the remaining *contextual*

TABLE 5
THE INFLUENCE OF INDIVIDUAL AND GROUP CHARACTERISTICS ON THE LABELLING CHILDREN AS PROBLEM STUDENTS

Individual effects	
- intelligence	.01* (.08, .00)
- achievement	.05* (.57, .00)
- learning ability	.18* (.15, .02)
- disruptive behavior	.22* (.17, .02)
- SES	.08* (.10, .01)
Group effects	
- mean intelligence	-.06* (-.25, .01)
- mean disruptive behavior	.21* (.08, .10)
- mean gender	-.91* (-.08, .36)
- variation in achievement	-.00 (-.00, .01)
- class size	-.00 (-.00, .01)
- variation in cultural background	-.00 (-.00, .07)
- mean parental pressure	-.01 (-.01, .01)
R ² individual	47%
R ² intercept	19%
Deviance	6565.22
Df	13

* $p < .05$
between brackets: beta and standard errors

variables which have been theoretically derived in our study to the significant context variables given in table 4 (whether the latter are theoretically funded or not).

None of the features affects the results. The heterogeneity of the classes in terms of both achievement and cultural background of the children turns out not to influence teachers' judgment. That goes also for the number of students the teacher has and the amount of pressure put on the teacher by the parents. Moreover, one has to remember that allowance has already been made for the influence of both individual student characteristics and aggregated student characteristics. Because the paths between variables cannot be specified in a VARCL analysis, it is not possible to investigate whether or not the added variables have any indirect effects.

Finally, we added teacher characteristics to the analysis.

TABLE 6
THE INFLUENCE OF INDIVIDUAL, GROUP AND TEACHER CHARACTERISTICS ON LABELING CHILDREN AS PROBLEM STUDENTS

Individual effects		
- intelligence	.01*	(.08, .00)
- achievement	.05*	(.57, .00)
- learning ability	.18*	(.15, .02)
- disruptive behavior	.22*	(.17, .02)
- SES	.08*	(.10, .01)
Group effects		
- mean achievement	-.06*	(-.25, .01)
- mean disruptive behavior	.22*	(.08, .10)
- mean gender	-.90*	(-.08, .35)
Teacher effects		
- gender	.09	(-.03, .10)
- experience	-.00	(-.01, .01)
- additional training	-.06	(-.03, .08)
R ² individual	48%	
R ² intercept	21%	
Deviance	6563.45	
Df	13	

* $p < .05$
 between brackets: beta and standard errors

As stated in the introduction, the literature does not provide any clear indications concerning relevant characteristics of teachers which may be associated with referral behavior. We added *gender, experience and additional training* to the analysis. None of these factors turned out to relate significantly to teachers' judgment of students' difficult behavior. This confirms the hypothesis that contradictory research findings can probably be blamed on failure to specify the models used sufficiently, or on sampling errors.

TABLE 7
THE INFLUENCE OF INDIVIDUAL AND AGGREGATED STUDENT CHARACTERISTICS ON LABELING CHILDREN AS PROBLEM STUDENTS

Individual effects	
- intelligence	.01* (.08, .00)
- achievement	.05* (.57, .00)
- learning ability	.18* (.14, .02)
- disruptive behavior	.22* (.17, .02)
- SES	.08* (.10, .01)
Aggregated effects	
- mean achievement	-.06* (-.25, .01)
- mean disruptive behavior	.22* (.07, .10)
- mean gender	-.91* (-.08, .36)
R ² individual	47%
R ² intercept	19%
Deviance	6565.49
Df	10

* $p < .05$
 between brackets: beta and standard errors

As reported above, 87.4% of the variation is found at individual level.

Table 7 shows the results of an analysis in which the calculation of all non-significant variables from the previous analyses was left out. This table shows that 47% of that is explained by the measured variables at individual level. In total, 41% of the variation in teachers' judgement is accounted for by the measured individual student characteristics. Clearly, there will be other individual characteristics which have not been identified, which, taken together will be at least as important.

This goes also for the contextual level. Only 19% of the 12.6% variation which is present at contextual level can be explained. Over 80% of the variance at contextual level has therefore not been identified, which provides opportunities for further research and theory formulation. All that remains to be said, is that the possibility of interaction between the separate variables was investigated. One cannot rule out the possibility that a combination of two variables may be important, while each variable on its own is not. If that were the case, then the theoretical model could be specified further. Analyses carried out to look at this did not produce any significant findings. In other words, the value of the characteristics (taken separately or combined) identified in this study to explain the problem under investigation is confined to the results presented here.

DISCUSSION

Referral to special schools is initiated by teachers who feel that certain children are problem students. In other words, if particular students are unable to participate

successfully in the education being offered to them, or are disrupting it badly, the teacher begins to have doubts about whether this student belongs in his class. For a long time, it has been assumed that teachers base this judgement solely on 'objective' observed grounds. One teacher might have a somewhat sharper perception of individual children's potential than others but, in principle, every teacher should be able to make an objective assessment of the learning behavior and potential of his students. Behavior which is regarded as difficult at school X would also be judged in the same way by the teachers at school Y. However, as time has gone by, research and reports given by practicing teachers have indicated that contextual factors probably also play a role in the process whereby teachers form judgments. The most frequently-quoted example has always been the role of the average level of achievement of a class: a child runs a greater risk of being labeled as a problem student in a class with a large number of high achievers than in a class where many of his/her fellow students are not performing very well.

This study investigated the influence of several such contextual factors. The use of a multilevel analysis resolved the technical problem of nested data, which bedeviled many earlier researchers in this field.

The analysis showed that about 13% of the variation in teachers' judgement is found at contextual level. This is comparable to the level of variation in educational achievement between different primary schools in the Netherlands. Although at first sight this might seem quite small, it could lead to noticeable differences between one school and another.

This research shows that the chance of being labeled as a problem student is greater when s/he is (a) less intelligent, (b) a low achiever, (c) the teacher thinks that her/his learning ability is low, (d) s/he shows disruptive behavior, (e) s/he comes from a family with a lower socioeconomic status, (f) the average achievement level in the class is higher, (g) the level of disruptive behavior among the students is higher, and (h) there are more boys in the class.

Factors which do not show an effect are: (a) the average level of intelligence in the class, (b) the average level of learning ability as estimated by the teacher, (c) the social composition of the class, (d) the ethnic composition of the class, (e) the size of the class, (f) the heterogeneity of students' performance, (g) the amount of pressure put on the teacher by the parents to help their child perform better, (h) the gender of the teacher, (i) his or her teaching experience and (j) his or her additional training.

If parents want to prevent their children from running a greater than average risk of being labeled as a problem student, then their school choice is certainly important. They should look out for the achievement level of the students, their level of disruptive behavior and their gender ratio.

In accordance with most previous studies, the above-mentioned data suggest that individual teacher characteristics do not correlate with students' labeling.

Furthermore, contextual factors which are most closely related to instruction appeared to be significant: mean achievement level and mean level of disruptive behavior. Remarkable and unexplainable is the influence of the gender ratio: it has already been controlled for all other influences, including disruptive behavior, both at individual and class level.

From the results of the multilevel analysis, one can conclude also that the individual student characteristics which were isolated account for about half of the variation at individual level, which amounts to about 87% of the total variation. Clearly, there will be other student characteristics which have not been identified, which also have an important influence on teachers' judgement. This is even more true for variation at group level. Variation at group level amounts to 13% of the total variation, of which only 19% is explained in the study. The search for causes at contextual level needs to be continued therefore. It seems likely that the indicators isolated in our research are still too 'proxy' to be the actual processes which influence teachers' judgement. Observations of the processes of group dynamics in the classroom would probably constitute a step in the right direction.

REFERENCES

- Brophy, J.E. (1995). *Elementary Teachers' Perceptions of and Reported Strategies for Coping with Twelve Types of Problem Students*. Michigan State University: Institute for Research on Teaching.
- Brophy, J.E. & Rohrkemper, M.M. (1981). The influence of problem ownership on teachers' perceptions of and strategies for coping with problem students. *Journal of Educational Psychology*, 73(3), 295-311.
- Brophy, J. & McCaslin (1992). Teachers' Reports on How They Perceive and Cope with Problem Students. *The Elementary School Journal*, 93, 3-68.
- Brophy, J.E. & Good, T.L. (1974). *Teacher-student relationships - Causes and Consequences*, New York: Holt: Rinehart & Winston.
- Brown, J.S., Collins, A. & Duguid, P. (1989). Situated knowledge and the culture of learning. *Educational Researcher*, 18, 32-44.
- Brown, B.W. & Saks, D.H. (1987). The Microeconomics of the Allocation of Teachers' Time and Student Learning. *Economics of Education Review*, 6, 319-332.
- Carrier, J.C. (1986). Sociology and special education: differentiation and allocation in mass education. *American Journal of Education*, 33, 281-313.
- Centraal Bureau voor de Statistiek (CBS) (1995). *Statistical Yearbook 1995*. Den Haag, Sdu.
- Elton Report (1989). *Discipline in schools*. London: HMSO.
- Fletcher, J.M., Shaywitz, S.E., Shankweiler, D.P., Katz, L., Liberman, I.Y., Stuebing, K.K., Francis, D.J., Fowler, A.E. (1994). Cognitive profiles of reading Disability: Comparisons of Discrepancy and Low Achievement Profiles. *The Journal of Educational Psychology: including Experimental Pedagogy, Child Psychology and Hygiene and Educational Statistics*, 86(1), 6-22.
- Gersons-Wolfensperger, D.C.M. & Ruijsenaars, A.J.J.M. (1997). Definition and Treatment of Dyslexia: A Report by the Committee on Dyslexia of the Health Council of the Netherlands. *Journal of Learning Disabilities*, 30(2), 209-213.
- Goldstein, H. (1987). *Multilevel Models in Educational and Social Research*. London, Oxford University Press.

- Helton, G.B. & Oakland, Th.D. (1977). Teachers' attitudinal responses to differing characteristics of elementary school students. *Journal of Educational Psychology*, 69(3), 261-265.
- Hudley, C.A. (1993). Comparing Teacher and Peer Perceptions of Aggression: An Ecological Approach. *Journal of Educational Psychology*, 85(2), 377-384.
- Hutton, J.B. (1984). Teacher ratings of problem behaviors: Which student behaviors "concern" and "disturb" teachers? *Psychology in the Schools*, 21, 482-484.
- Letiche, H.K., Wolf, J.C. van der & Plooy, F.X. (eds.) (1991). *The Practitioners' Power of Choice in Staff-Development and In-Service-Training*. Lisse, Zwets & Zeitlinger.
- Longford, N.T. (1988). *VARCL manual*. Princeton New Jersey, Educational Testing Service.
- Lopez, M.F. (1996). Children with Attention Deficit Hyperactivity Disorder and Emotional or Behavior Disorders in Primary Grades: Inappropriate Placement in the Learning Disability Category. *Education and Treatment of Children*, 19(3), 286-299.
- McMichael, P. (1979). The hen or the egg? Which come first - Antisocial emotional disorders or reading disability? *British Journal of Educational Psychology*, 49, 226-238.
- Macmillan, D.L. (1996). The Labyrinth of IDEA: School Decisions on Referred Students with Subaverage General Intelligence. *American Journal of Mental Retardation*, 101(2), 161-174.
- Merrett, F. & Wheldall K. (1984). Classroom behavior problems which junior school teachers find most troublesome. *Educational Studies*, 10, 87-92.
- Meijer, C.J.W. (1986). *Verwijzing en terugplaatsing* [Referral and Re-Referral]. Groningen: RION.
- Meijer, C.J.W., Pijl, S.J. & Hegarty, S. (eds.) (1994). *New Perspectives in Special Education*. London & New York: Routledge.
- Meijer, C.J.W., Pijl, S.J. & Rispen, J. (1986). *Beslissen over verwijzen en toelaten* [Deciding on Referral and Admission]. Lisse: Swets & Zeitlinger.
- Mortimore, P., Sammons, P., Ecob, R., Stoll, L. & Lewis, D. (1988). *School matters: the junior years*. Salisbury: Open Books.
- O'Sullivan, J.T. (1992). *Reading beliefs and Reading Achievement: A Development Study of Students from Low-Income Families. Report Nr 6. Summary Reports of Paths to Literacy and Illiteracy in Newfoundland and Labrador*. St. John: Memorial University, 25 p.
- Patrick, H. (1995). *Personality Characteristics, Social Competence, and Early School Adjustment: A Contextual and Developmental Perspective*. Paper presented at the Biennial Meeting of the Society for Research in Child development. Indianapolis, March 30 April 2, 19 p.
- Safran, S.P. & Safran, J.S. (1985). Classroom context and teachers' perceptions of problem behaviors. *Journal of educational Psychology*, 77, 20-28.
- Solomon, D. (1996). *Teacher Beliefs and Practices in School Serving Communities that differ in Socio-economic Level*. Paper presented at the Annual Meeting of the American Educational Research Association. New York: April 8-12, 28 p.
- Scheerens, J. (1992). *Effective schooling: Research, theory and practice*. London: Cassell.
- Stevens, D.D. & Palincsar, A.S. (1992). *Urban Teachers' Beliefs and Knowledge about Literacy Teaching and Learning: An Examination from Mechanistic and Contextualistic Perspectives*. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, April 20-24, 40 p.
- Sturge, C. (1982). Reading retardation and antisocial behavior. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 23, 21-31.
- Thurlow, M.L., Christenson, S. & Yseldyke, J.E. (1983). *Referral research: an integrative summary of findings. Research Report 141*. Minneapolis: Institute for Research on Learning Disabilities, University of Minneapolis.
- Vaughn, S., Schumm, J.S., Jallad, B., Slusher, J. & Saumell, L. (1996). Teachers' Views of Inclusion. *Learning Disabilities Research and Practice*, 11(2), 96-106.

- West, J. (1993). *Readiness for Kindergarten: Parent and Teacher Beliefs. Statistics in Brief*. Washington DC: National Center for Education Statistics, 11 p.
- Wissel, A. van der (1984). *Schooluitval* [Early school leaving]. Groningen, University of Groningen.

