DOCUMENT RESUME

ED 306 292 TM 013 159

AUTHOR Wagner, Mary; Shaver, Debra M.
TITLE Educational Programs and Achievements of Secondary Special Education Students: Findings from the National Longitudinal Transition Study.
INSTITUTION SRI International, Menlo Park, Calif.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.
PUB DATE Mar 89
CONTRACT 300-87-0054
NOTE 52p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; Academic Failure; Disabilities; *Longitudinal Studies; Mainstreaming; *National Surveys; School Demography; *School Surveys; Secondary Education; *Secondary School Students; *Special Education; Student Characteristics; Transitional Programs; Vocational Education
IDENTIFIERS *National Longitudinal Transition Study

ABSTRACT

Initial findings regarding the educational programs and other secondary special education services surveyed during the congressionally mandated National Longitudinal Transition Study (NLTS) of Special Education Students are presented. The study, started in 1987 under a Department of Education contract with SRI International, addresses issues concerning disabled youths' school programs, services, social integration, educational achievements, and independent living and employment experiences. Three major areas are addressed: (1) what educational programs and other services are provided to secondary special education students; (2) how well to these students perform in school; and (3) what student characteristics are related to school performance, as measured by receipt of failing grades, among special education students. Data for a nationally representative sample of more than 8,000 youth (aged 13 to 23 years) who attended special education in the 1985-86 school year, were collected in 1987 via telephone interviews with parents, a survey of schools youth attended, and students' school records. Specific findings are presented on types of educational programs, nature and size of schools attended, educational achievement, school characteristics, participation in special education and regular education courses, enrollment in vocational education, other services received, demographic characteristics, academic achievement, and characteristics related to school performance, as measured by failing grades. Fourteen data tables are provided. An overview of the NLTS is appended. (TJH)

EDUCATIONAL PROGRAMS AND ACHIEVEMENTS OF SECONDARY SPECIAL EDUCATION STUDENTS: FINDINGS FROM THE NATIONAL LONGITUDINAL TRANSITION STUDY

March 1989

Prepared by:

Mary Wagner, Ph.D., Director
National Longitudinal Transition Study
SRI International

and

Debra M. Shaver
Center for Health, Education, and
Social Systems Research
SRI International

Prepared for presentation to the Special Education Special Interest Group at the meetings of the American Educational Research Association, San Francisco, California, March, 1989.

This research was supported by contract number 300-87-0054 from the Office of Special Education Programs, U.S. Department of Education. The findings presented in this paper do not necessarily reflect the views or policies of the U.S. Department of Education.

SRI International®
333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 326-6200 • TWX: 910-373-2046 • Telex: 334-486

EDUCATIONAL PROGRAMS AND ACHIEVEMENTS OF SECONDARY SPECIAL EDUCATION STUDENTS: FINDINGS FROM THE NATIONAL LONGITUDINAL TRANSITION STUDY

In the 1986-87 school year, more than 1.5 million secondary school students received special education services under the Education of the Handicapped Act (U.S. Department of Education, 1988). What programs and services are provided to these secondary-age special education students? How well do these students achieve in school?

Responding in part to the lack of information to answer such questions, the U.S. Congress mandated in 1983 that the U.S. Department of Education conduct a national study of youth in the years of transition from secondary school to adult living (Sec. 8, section 618(e), PL 98-199). The Office of Special Education Programs (OSEP) of the U.S. Department of Education contracted with SRI International to develop a study design and student sample; in 1987, under a second contract, SRI began the National Longitudinal Transition Study of Special Education Students. The study addresses issues concerning disabled youths' school programs, services, social integration, educational achievements, and independent living and employment experiences.

This paper presents the first findings regarding the educational programs and other services and the secondary school achievement of special education students nationwide. We address three major questions:

- What educational programs and other services are provided to secondary special education students?
- How well do secondary special education students do in school?
- What student characteristics are related to school performance, as measured by receipt of failing grades, among secondary special education students?

The following sections of this paper present findings related to these questions based on National Longitudinal Transition Study (NLTS) data for a nationally representative sample of more than 8,000 youth who were ages 13 to 23 and in special education in the 1985-86 school year. Data were collected

in 1987 from telephone interviews with parents, a survey of the schools youth attended, and from the students school records. (See the appendix for a further description of the NLTS and descriptive statistics regarding the demographic characteristics of youth in the sample.)

We will first present descriptive findings, then multivariate analyses of effects of student background factors on one measure of student achievement.

Educational Programs and Other Services Provided to Secondary Special Education Students

"Educational program" is a complex construct. It encompasses aspects of school setting and climate, courses taken, lesson content, curriculum, and instructional method. Additional services, too, can include a complex combination of several kinds of therapies and support services provided to help students benefit from their educational programs. Capturing this complexity in detail for secondary special education students nationally is beyond the scope of the NLTS. However, we can paint, in broad strokes, major aspects of students' educational programs and the kinds of additional services they are reported to receive. This section presents descriptive data on five aspects of the educational programs and services of secondary youth with disabilities:

- The nature and size of the schools attended.
- Participation in special education.
- Involvement in regular education courses.
- Enrollment in vocational education courses.
- The nature of additional services provided by the schools and others.

School Characteristics

The school environment is an important factor in understanding the experiences youth with disabilities have in school. Two aspects of the school environment are described in Table 1*: the types and sizes of schools that youth with disabilities attended, as reported by school administrators.

Most youth with disabilities (89%) attended comprehensive secondary schools whose student bodies were primarily nondisabled students. However, 8% of secondary youth attended special schools for youth with disabilities. The rate at which youth attended special schools varies considerably between disability categories. For example, 63% of youth in the deaf category and 94% of those who are deaf/blind attended special schools, a significantly higher percentage than for youth with emotional disturbances or mental retardation, for example (12% and 17%; $p \leq .01$). Youth with visual or multiple disabilities also had relatively high rates of attending schools for disabled students (35% and 41%, respectively), compared to youth in such categories as speech or health impaired (4% and 10%; $p \leq .01$).

These figures on attendance at special schools by secondary-age students are quite similar to the rates reported by the federal government for all age groups for 1985-86 (U.S. Department of Education, 1988). For example, federal data indicate that 7% of all special education students attended public or private day or residential schools for youth with disabilities, compared to the NLTS rate of 8%. Similarly, federal figures indicate 15% of youth with mental retardation and 18% of youth with orthopedic impairments attended special schools, compared to NLTS rates of 17% and 14%. Only for youth who have emotional disturbances or who are deaf/blind, do NLTS rates substantially exceed federal figures; i.e., NLTS rates of 35% and 94% exceed federal rates of 24% and 51% for these groups.

---

* In Tables 1 through 10, percentages are weighted to represent youth in each primary disability category and age group (see appendix). Sample sizes are unweighted. Primary disability category is based on reports from schools or school districts.

Table 1: TYPES AND SIZES OF SCHOOLS ATTENDED BY SECONOARY STUDENTS WITH OISABILITIES

Type of School Attended	Total	Primary Oisability Category: Learning Disabled	Emotionally Oisturbed	Mentally Retarded	Speech Impaired	Visually Impaired	Hard of Hearing	Oeaf	Oeaf/ Blind	Orthoped- ically Impaired	Health Impaired	Multiply Handi- capped
Percentage of youth attending:												
Comprehensive school	88.8	95.2	82.5	80.4	93.5	62.7	87.0	36.1	5.9	93.1	88.0	53.4
Special school for students with disabilities	8.0	1.6	12.4	17.2	4.1	34.7	9.0	63.2	94.1	14.4	10.2	40.8
Magnet school	0.2	0.2	0.7	0.0	0.8	0.4	0.8	0.2	0.0	1.1	0.0	0.0
Vocational technical school	1.6	2.0	0.9	1.2	0.8	0.5	2.8	0.4	0.0	0.7	0.8	1.5
Continuation or alternative school	1.4	1.0	3.6	1.2	0.8	1.7	0.5	0.1	0.0	0.8	1.0	4.3
(Number of respondents)	6781	955	588	948	477	761	629	774	90	595	368	596
Percentage of youth attending schools with an average daily attendance of:												
Fewer than 500 students	27.7	22.5	29.5	38.1	20.5	41.4	19.4	66.5	94.8	20.7	24.7	58.7
501 to 1100 students	38.7	40.5	33.7	40.3	37.4	24.1	36.6	10.8	2.1	32.6	23.4	24.8
More than 1100 students	33.6	37.1	36.8	21.6	42.1	34.5	44.0	22.8	3.1	46.6	51.9	15.5
(Number of respondents)	6696	940	580	930	460	752	627	773	90	592	361	591

Using a 2-tailed test, the sampling errors at the 95% confidence level for type of school attended by the full sample are under ±1%. For individual disability categories, confidence intervals for attendance at comprehensive schools range from ±1% for the LD category to ±4% for the multiply handicapped category.

Source: mail survey of administrators in schools attended most recently by sample youth.

The relative advantages of schools of different sizes have long been debated in the educational arena. Some large schools are able to provide a broader range of course offerings, placements, support services, and specialized staff than small schools. Smaller schools may provide more opportunities for individual attention and a more manageable environment for exploring and exercising students' skills, roles, and responsibilities. Table 1 demonstrates that, overall 28% of youth with disabilities attended schools with fewer than 500 students, 39% attended schools with between 500 and 1100 students, and about one-third attended schools with more than 1100 students.

The distribution of special education students overall masks variation by disability category in the size of schools attended. Youth who are deaf, deaf/blind, or have multiple impairments were significantly more likely to attend schools with fewer than 500 students than were youth with speech or learning disabilities, for example ($p \leq .001$). This reflects the smaller size of the special schools attended more often by youth in these categories than by other students; the average daily attendance at special schools was 182, compared to 1,216 for comprehensive secondary schools attended by youth with disabilities. Youth with mental retardation or visual impairments were also more likely to attend smaller schools than youth with emotional disturbances or physical impairments, for example (around 40%, compared to 20% to 30%, $p \leq .01$).

Participation in Special Education

The common adage that special education is a one-way street--once in special education, always in special education--has been challenged in recent research, which reports a 1-year declassification rate of 17% for elementary students in 3 urban school districts (Singer, 1988). This rate for elementary students appears higher than for youth in upper grades. NLTS data in Table 2 indicate that about 5% of secondary youth were declassified from special education in their most recent year in secondary school, as reported in school records. This rate is the same as the 1-year declassification rate for elementary and secondary students together reported by the Council of the Great City Schools for its member districts (CSGCS, 1986).

Table 2: STUDENTS DECLASSIFIED FROM SECONDARY SPECIAL EDUCATION

Primary Disability Category	% Declassified from Special Education	Sample Size
All conditions	4.7	6182
Learning disabled	5.2	881
Emotionally disturbed	7.1	551
Mentally retarded	1.4	925
Speech impaired	18.0	406
Visually impaired	3.6	648
Hard of Hearing	2.3	563
Deaf	.3	714
Deaf/blind	.0	72
Orthopedically impaired	4.5	558
Health impaired	7.0	306
Multiply handicapped	.2	558

Using a 2-tailed test, the sampling error at the 95% confidence level for youth in all conditions is ±.5%. Confidence intervals for individual disability categories range from ±1% to ±4%.

Source: Students' school records for their most recent year in secondary school.

Secondary students with speech impairments were declassified at a rate of 18%, which is almost 3 times the rate at which youth in any other category were declassified ($p \leq .01$). NLTS data reveal that about 7% of youth with health impairments and emotional disturbances were declassified during their most recent year in secondary school. Fewer than 2% of students with disabilities such as mental retardation, hearing impairments, and multiple disabilities were declassified. We did not find significant differences in the rates of declassification based on grade level of the students.

Enrollment in Regular Education Courses

The degree to which students are served in settings which inhibit or encourage interaction with nondisabled youth and the regular instructional program is important in understanding the educational experiences of youth with disabilities. Students who are served only in special education classes made up solely of students with disabilities have different experiences than students who are more integrated with the regular instructional program and with nondisabled peers. The concept of "least restrictive environment," a

cornerstone of the Education of the Handicapped Act (EHA), reflects the intent of special education to maximize integration to the extent appropriate for individual students.

Table 3 describes the level of enrollment in regular education courses in the most recent year in school of secondary special education students who attended schools that also served nondisabled youth. Overall, 17% of disabled youth in schools with nondisabled students were enrolled exclusively in special education courses. Not surprisingly, the extent to which youth were in completely self-contained special education courses varies greatly by disability category. For example, students in the deaf/blind and multiply handicapped categories were much more likely than other youth to be in special education classes exclusively, with about 70% taking no regular education courses ($p<.01$). However, even among youth in such categories as learning disabled, speech impaired, or hard of hearing, about 1 in 10 youth were not enrolled in any regular education courses in their most recent year in secondary school.

Almost 1 in 4 students were mainstreamed for nonacademic* subjects only. This was the most common program for youth with mental retardation; 42% of these youth were mainstreamed only for nonacademic courses, a significantly higher percentage than for other categories, such as emotionally disturbed or deaf, for example ($p\leq.01$). Overall, 44% of youth were mainstreamed for some academic subjects, and 9% were mainstreamed for all courses. Youth with visual, speech, and health impairments were significantly more likely than other youth to be enrolled entirely in regular education courses ($p\leq.01$). About half of youth with learning disabilities, emotional disturbances, visual impairments, or who are hard of hearing were mainstreamed for part of their academic subjects, but continued to take some of their coursework in special education classes.

---

* Academic courses include courses in English, mathematics, science, social science, or foreign language. Other classes that do not fall in the academic category include courses in home economics or life skills, the arts, vocational education, physical education, study hall, health, driver's education, and other some additional electives.

Table 3: ENROLLMENT IN REGULAR EDUCATION COURSES BY STUDENTS WITH DISABILITIES ATTENDING REGULAR SECONDARY SCHOOLS

		Primary Disability Category:										
	Total	Learning Disabled	Emotionally Disturbed	Mentally Retarded	Speech Impaired	Visually Impaired	Hard of Hearing	Deaf	Deaf/ Blind	Orthoped-ically Impaired	Health Impaired	Multiply Handi-capped
Percentage of youth enrolled in:*												
No regular education classes	16.9	9.5	18.3	31.9	12.1	15.9	11.5	34.2	72.5	28.1	27.1	69.1
Regular education for nonacademic courses only	24.0	20.0	16.6	41.6	9.4	6.3	23.5	19.6	25.5	13.9	11.5	10.0
Some regular education courses (subjects unknown)	5.7	6.3	4.8	4.6	7.1	3.0	5.7	2.8	0.0	6.8	3.1	8.6
Regular education for academic courses	44.1	54.1	47.9	19.6	45.1	49.9	50.0	39.8	0.0	36.6	33.3	10.1
All regular education classes	9.3	10.2	12.4	2.3	26.4	28.9	9.2	3.6	2.1	14.6	25.8	2.1
(Number of respondents)	5170	872	503	828	405	425	543	410	22	509	287	366

Using a 2-tailed test, the sampling errors at the 95% confidence level for the full sample range from ±.6% to ±1.4%. For disability categories, they range from ±2% to ±5% for most categories. For the deaf/blind category, they range up to ±19%.

Source: students' school records.

Enrollment in Vocational Education

Vocational education as a field has recently emphasized the recruitment of students with special needs, as reflected in the Carl D. Perkins Vocational Education Act of 1984 (PL 98-524). Table 4 reports students' enrollment in vocational education in their most recent year in secondary school, as indicated in school records. Overall, about 60% of youth with disabilities took at least one vocational education course in their ost recent year in school. Data recently reported from the National High School Transcript Study suggests 96% of special education students attending regular high schools took some vocational education courses in their 4-year high school career (Hayward, 1989). NLTS data suggest that those who were enrolled in vocational education spent an average of 6.8 hours per week in these courses during the most recent school year.

Although participation in vocational education ranged from 48% of youth with multiple disabilities to 76% of youth who are deaf ($p \leq .01$), for most disability categories, about 5 or 6 of 10 youth were enrolled in some vocational education courses in their most recent year in secondary school. The average number of hours spent in vocational education does not vary greatly by disability category, being between 5 and 7 hours per week for most groups.

Rates of participation in vocational education steadily increased from grade level to grade level. Among youth with disabilities in 7th and 8th grades, the rate of enrollment in vocational courses was 51%, compared to about 71% of youth in 9th grade and 86% for youth who were in 11th or 12th grade ($p \leq .01$). The intensity of involvement in vocational education also increases by grade level. For example, 9th grade vocational students averaged 5 hours per week in vocatinal courses during the year; seniors who took vocational education averaged 9 hours in those courses during the year.

Table 4: VOCATIONAL EDUCATION PARTICIPATION OF YOUTH WITH DISABILITIES

Student Characteristics	Vocational Education Enrollment in Most Recent School Year		Average Hours Per Week of Vocational Education in Most Recent School Year	
	%	N	%	N
Primary disability category				
All conditions	59.5	7766	6.8	4432
Learning disabled	59.2	1103	6.9	665
Mentally retarded	65.9	1113	6.9	711
Emotionally disturbed	51.8	726	6.1	376
Speech impaired	50.9	557	5.4	270
Visually impaired	57.3	807	5.9	427
Hard of hearing	60.2	720	6.5	404
Deaf	76.5	834	7.6	600
Deaf/blind	60.0	83	10.0	41
Orthopedically impaired	51.4	707	6.6	361
Other health impaired	55.2	434	5.5	227
Multiply handicapped	47.8	688	7.0	350
Grade level				
All grades	59.5	7766	7.1	3874
7th-8th grade	51.0	629	1.3	99
9th grade	70.8	962	5.0	546
10th grade	78.5	974	6.5	626
11th grade	86.4	1036	8.0	751
12th grade	86.2	1426	9.0	1088
Ungraded	65.2	1027	8.5	549

Using a 2-tailed test, the sampling error at the 95% confidence level for vocational education participation for youth in all conditions is ±1%. Confidence intervals for most categories range from ±3% to ±5%. The confidence interval for the deaf/blind category is ±11% because of its small sample size.

Source: Students' school records.

When we examine the participation rates in vocational education by grade level for individual disability categories (table not included), we see one possible explanation for differences between disability categories. Lower participation by youth in some categories appears to result from the higher incidence of youth in ungraded courses. Youth in ungraded programs have a lower participation rate in vocational courses; a higher incidence of such students in a disability category, such as is the case for youth with multiple handicaps, lowers the overall participation rate for that category.

Other Services Received

Under EHA, special education students who need them are entitled to specific kinds of services to help them benefit from their educational programs. Table 5 reports data on the percentage of youth whose parents or schools reported they received selected services.

More than half of students with disabilities (53%) reportedly did not receive from their school any of the services that we investigated in the previous year. Occupational therapy or life skills training was the most common service reported, with 23% of secondary students receiving it in the previous year, largely through instructional courses rather than supplemental therapy. Speech or language therapy was provided by the school to 16% of secondary students. This compares to a rate of 20% for students in all grade levels reported in the National Special Education Expenditure Study (Moore, et al., 1988). Personal counseling and aides that gave tutoring, reading, or interpreting services were reportedly provided to 15% and 13% of students, respectively. Transportation assistance was provided to 10% of secondary students, compared to a rate 3 times as high reported for youth at all grade levels (Moore, et al., 1988). Physical therapy and hearing-loss therapy were less common. Speech therapy, logically, was most commonly provided to youth with speech, hearing, or multiple impairments. Personal counseling was most often provided to youth in the emotionally disturbed category (31%). Physical therapy or mobility training was most often provided to youth with physical, visual, health, or multiple impairments.

Table 5: SERVICES RECEIVED BY SECONDARY STUDENTS WITH DISABILITIES

Service	Total	Primary Disability Category: Learning Disabled	Emotionally Disturbed	Mentally Retarded	Speech Impaired	Visually Impaired	Hard of Hearing	Deaf	Deaf/ Blind	Orthoped-ically Impaired	Health Impaired	Multiply Handi-capped
Percentage of youth receiving in the past year from or through their school:												
No additional services	52.8	61.0	54.3	40.0	43.4	39.6	30.1	26.0	30.4	32.6	44.0	16.7
Speech or language therapy	16.5	9.6	6.4	27.8	44.6	10.6	50.2	56.5	25.8	20.1	15.9	57.6
Personal counseling or therapy	14.6	12.1	31.0	13.7	5.1	15.9	13.8	27.4	14.2	13.6	14.7	23.0
Occupational therapy or life skills training	22.8	17.0	15.5	36.9	16.6	32.1	20.9	39.1	41.0	34.1	27.7	53.3
Help from a tutor/reader/interpreter	13.0	13.9	9.3	10.8	6.9	23.6	32.9	45.1	22.8	15.5	15.4	12.8
Physical therapy/mobility training	4.9	2.0	1.8	9.5	1.4	18.0	3.4	8.7	32.2	35.4	10.3	32.6
Hearing-loss therapy	1.2	0.0	0.2	1.0	1.0	2.2	41.6	52.7	54.1	0.6	1.1	6.1
Help in getting or using transportation	9.5	2.0	6.2	22.4	3.7	31.1	21.1	24.9	41.8	45.4	19.1	55.5
(Number of respondents)	8169	1152	762	1165	573	850	748	393	96	748	460	722
Percentage of youth not receiving services from their school in the past year who received them from other sources:												
Speech or language therapy	.7	.3	.6	1.5	1.1	.5	1.4	2.2	2.0	1.9	.7	5.1
Personal counseling or therapy	5.0	3.9	10.5	4.9	4.5	4.2	3.6	3.2	1.9	5.3	8.5	5.4
Occupational therapy or life skills training	1.5	.9	1.7	2.4	1.1	2.8	1.0	2.1	7.5	3.2	1.7	6.8
Help from a tutor/reader/interpreter	2.8	2.7	2.8	2.8	3.4	4.7	4.8	11.4	7.6	1.8	3.1	5.4
Physical therapy/mobility training	1.1	.2	.2	2.2	.2	6.0	.8	.9	9.1	8.6	7.9	9.0
Hearing-loss therapy	.2	.0	.0	.3	.0	.0	3.2	4.2	5.0	0.0	.1	2.2
Help in getting or using transportation	54.6	59.6	50.0	48.1	61.6	45.1	45.1	44.1	27.8	29.7	45.3	31.3
(Number of respondents)	8169	1152	762	1165	573	850	748	893	96	748	460	722

Using a 2-tailed test, sampling errors at the 95% confidence level for the full sample are ±1% or lower. For disability categories, they range from below ±1% to ±5%.

Although these services were concentrated in the categories to which they would seem most appropriate, only a minority of youth in most categories were reported to have received any of these services in the previous year. For example, 31% of youth with emotional disturbances received personal counseling in the previous year from the school. The second half of the table reveals that an additional 10% of youth in this category received counseling from another source, for a total of 42% of youth with emotional disturbances receiving counseling from any source. Similarly for youth classified as speech impaired, 44% received speech therapy from their school and another 1% received services from another source, leaving a majority of speech impaired youth receiving no speech therapy. This does not mean, however, that these youth received no help with their disabilities; for example, as part of their special education instructional program, they may have been enrolled in language-oriented classes or classes specifically for youth with emotional disturbances, rather than being provided speech therapy or counseling as adjunct services.

The second half of Table 5 further demonstrates that the school was the primary provider of all related services except transportation. Excepting transportation, the percentage of youth who reportedly received services from the school exceeds those who received services from other sources for all services and all categories of youth.

Academic Achievement

The previous sections of this paper have demonstrated considerable variation in the educational programs and services experienced by secondary school youth with disabilities. Students' levels of achievement in secondary school also vary widely. Here, we examine four measures of school achievement, based on information taken from students' school records and/or parent reports:

- Whether youth who were in graded programs* received a failing grade in any course in their most recent year in secondary school.
- Whether youth in graded programs who were to continue in school were promoted to the next grade level.
- Whether youth who were subject to minimum competency tests passed them.
- Whether youth completed secondary school by graduating, dropping out, or exceeding the school age limit.

The extent to which youth who were in graded programs received failing grades in school is revealed in Table 6. Almost 1 in 3 youth with disabilities (31%) who were in graded programs received a failing grade in 1 or more classes in their most recent school year. Youth with emotional disturbances were significantly more likely than youth in any other category to have received a failing grade (45%; $p \leq .01$). They were also generally more likely to be failing more courses when they were failing. For example, 19% of youth with emotional disturbances received a failing grade in 6 or more classes, compared to 8% of youth with learning disabilities and 6% of youth with speech impairments ($p \leq .01$).

Failing grades were more likely to be given to youth in lower grades, as demonstrated in Table 7. The percentage of youth receiving at least 1 failing grade is fairly stable from 7th to 10th grade, but then decreases significantly, from 42% of 9th and 10th grade students to 34% of 11th grade students ($p \leq .05$) and to 19% of 12th graders ($p \leq .01$). Twelfth graders were also more likely to be failing only one course when they failed than were students in earlier grades. To the extent that failing in school leads to dropping out of school (Butler-Nalin and Padilla, 1989; Wagner, 1989), the relationship between age and failing in school may result from the fact that many failing students left school before they reached the upper grades. Alternatively, teachers may have been more lenient with older students, reasoning that they were close to the end of their secondary school careers and that little was to be gained by forcing a youth to repeat a class by assigning him or her a failing grade.

---

* Youth are considered in a graded program if school records indicated they were designated as at a specific grade level or received a grade for at least one of the classes in which they were enrolled.

Table 6: RECEIPT OF FAILING GRADES IN MOST RECENT SCHOOL YEAR, BY CATEGORY OF DISABILITY

Receipt of Failing Grades	Total	Primary Disability										
		Learning Disabled	Emotionally Disturbed	Mentally Retarded	Speech Impaired	Visually Impaired	Hard of Hearing	Deaf	Deaf/ Blind	Orthoped-ically Impaired	Health Impaired	Multiply Handi-capped
Percentage of youth receiving grades who received a failing grade in one or more courses in the most recent year in secondary school	31.3	34.8	44.6	21.8	35.0	17.1	21.2	8.1	4.0	15.2	25.8	6.5
(Number of respondents)	5683	812	506	864	366	567	518	688	71	473	287	531
Of those receiving a failing grade, percentage failing:												
1 course	42.6	44.4	33.8	41.2	46.7	60.4	51.4	62.6	--	46.5	46.1	37.2
2 courses	22.9	25.5	20.2	18.3	13.6	18.1	13.1	24.0	--	27.7	15.6	22.4
3 courses	11.8	12.0	9.5	12.1	18.9	4.9	10.8	2.8	--	10.8	6.4	9.2
4 courses	5.5	4.0	8.9	7.0	9.6	3.5	10.2	5.4	--	8.1	4.3	7.8
5 courses	6.7	5.6	8.6	9.0	5.6	5.3	7.1	1.4	--	5.0	18.3	6.2
6 or more courses	10.5	8.4	19.0	12.4	5.5	7.8	7.3	3.8	--	2.0	9.2	17.2
(Number of respondents)	1181	255	208	169	120	91	99	60	2	65	74	38

Using a 2-tailed test, the sampling error at the 95% confidence level for receipt of failing grades for students in all conditions is ±1%. Confidence levels for individual disability categories range from ±2% for the deaf category to ±5% for the other health impaired category. For the number of courses failed, the confidence intervals range from ±1% to ±3%. For individual disability categories, confidence intervals range from ±1% to ±15% for youth in the multiply handicapped category failing 1 course.

Source: students' school records.

Table 7: RECEIPT OF FAILING GRADES, BY GRADE LEVEL

Receipt of Failing Grades	Total	Grade Level 7 or 8	9 or 10	11	12
Percentage of youth in graded programs receiving a failing grade in 1 or more courses in the most recent year in secondary school	31.3	33.9	41.7	33.7	19.0
(Number of respondents)	5649	551	1177	959	1312
Of those receiving a failing grade, percentage failing:					
1 course	42.6	37.1	37.2	47.5	63.8
2 courses	22.9	27.6	23.5	21.6	20.3
3 courses	11.8	20.9	9.2	12.4	11.5
4 courses	5.5	3.1	6.5	5.6	1.5
5 courses	6.7	3.4	8.8	5.8	1.6
6 or more courses	10.5	7.9	14.8	7.1	1.2
(Number of respondents)	1181	152	572	233	179

Using a 2-tailed test, the sampling error at the 95% confidence level for receipt of failing grades for all students is ±1% and by grade level, ranges from ±3% to ±4%. By number of courses, the confidence levels range from ±1% to ±2%. By grade level, they range from ±2% to 8%.

Source: students' school records.

Another measure of students' performance is whether or not they successfully completed the school year and were promoted to the next grade level. In Table 8, we show the percentage of youth in each disability category who were promoted to the next grade level at the end of the school year. (Students in 12th grade and students who were in ungraded programs are not included in this table.) A large majority of youth (74%) were successfully promoted to the next grade level, with promotion rates being above 75% for most categories. When lower rates of promotion are apparent for a category, it is often indicative of a larger proportion of youth with a status of "other" at the end of the school year, which includes youth who dropped out. Findings for youth in the hard of hearing, learning disabled, multiply handicapped, and mentally retarded categories show that youth at the lower grade levels were more likely than older youth to experience grade retention. Although this pattern is not apparent across all disability categories, it is consistent with the findings for the other achievement variables showing that youth in the higher grade levels were less likely to receive failing grades than youth in the lower grades.

Table 8: PROMOTION RATES OF SECONDARY STUDENTS WITH DISABILITIES

Primary Disability Category	Percentage of Youth* Who: Were Promoted	Were Not Promoted	Other**	Sample Size
All conditions	74.3	6.1	19.6	3082
Learning disabled	76.9	4.6	18.5	503
Emotionally disturbed	60.3	10.8	28.9	311
Mentally retarded	69.7	8.3	22.0	387
Speech impaired	78.4	8.2	13.4	247
Visually impaired	87.7	8.2	4.9	333
Hard of hearing	88.2	3.8	8.0	342
Deaf	89.7	1.6	8.7	398
Deaf/blind	--	--	--	--
Orthopedically impaired	88.6	4.0	7.4	252
Health impaired	78.1	7.9	13.8	179
Multiply handicapped	81.0	10.2	8.8	128

* Youth in 12th grade and ungraded programs are not included in the sample on which these figures are based.

** The "Other" category largely includes youth who dropped out or withdrew. It also includes a minority of youth who moved or were suspended, expelled, institutionalized, or incarcerated.

-- Too few deaf/blind students in graded programs to be included in this analysis.

Using a 2-tailed test, the sampling errors at the 95% confidence level for youth in all conditions were ±1%. For disability categories, they range from ±2% to ±5%.

Source: Students' school records in their most recent year in school.

A third measure of achievement examined in the NLTS is whether students with disabilities met minimum competency requirements. Table 9 shows that, overall, 38% of youth who were in schools and at grade levels for which minimum competencies were usually tested were exempted from those tests. Exemption rates are significantly higher for youth with multiple disabilities, including those who are deaf/bind (83% and 80%, respectively), and for youth with mental retardation (73%) than for youth in any other disability categories. Exemption rates are between 20% and 25% for most other disability categories, with youth who have speech impairments being exempted least often (13%).

Table 9: MINIMUM COMPETENCY TEST REQUIREMENTS AND OUTCOMES OF SECONDARY STUDENTS WITH DISABILITIES

		Primary Disability Category:										
	Total	Learning Disabled	Emotionally Disturbed	Mentally Retarded	Speech Impaired	Visually Impaired	Hard of Hearing	Deaf	Deaf/ Blind	Orthoped-ically Impaired	Health Impaired	Multiply Handi-capped
Percentage of youth in schools and at grade levels for which minimum competency tests are required who were exempted from the test	38.0	25.0	22.2	72.9	12.6	21.9	20.1	29.0	80.0	42.0	23.6	82.7
(Number of respondents)	3325		273	510	237	456	328	357	28	303	190	288
Percentage of youth who were required to take minimum competency tests who:												
Passed all of the test	44.0	47.9	36.4	21.0	50.5	72.1	51.9	61.8	--	60.0	40.6	42.5
Passed part of the test	32.3	31.7	40.6	27.7	32.2	20.8	37.4	29.0	--	31.3	37.8	29.5
Did not pass any part of the test	23.6	20.4	22.9	51.4	17.3	7.2	10.8	9.2	--	8.8	21.6	28.0
(Number of respondents)	1923	314	190	131	187	268	258	240	4	157	123	51

Using a 2-tailed test, the sampling error at the 95% confidence level of the estimate of youth exempted from minimum competency testing is ±2%. Confidence intervals for disability categories range from ±4% for the mentally retarded category to ±6% for the deaf/blind category. Confidence intervals for estimates of results of competency testing for the full sample are ±2%. For disability categories, they range from ±4% for youth in the LD category to ±9% for youth in the other health impaired category.

Source: students' school records.

Of the students who were required to take minimum competency tests, 44% passed the entire test and 32% passed some of the test. Fewer than half of youth with learning disabilities, emotional disturbances, mental retardation, or health or multiple impairments fully met the minimum competency requirements to which they were subject. Almost 1 in 4 students failed to pass any part of the minimum competency tests they were required to take.

Finally, Table 10 presents data on school completion as the culmination of school achievement. Overall, in a two-year period 56% of special education exiters left secondary school by graduating. This figure is significantly lower than the graduation rate found in studies of the general student population. For example, the U.S. Department of Education "Wallchart" estimates the graduation rate for the general student population to be 71%, a rate similar to the 75% rate reported by the U.S. Bureau of the Census and the U.S. Center for Education Statistics (CES, 1986a; figures are for 1985). Differences are even more pronounced for youth in some disability

Table 10: SECONDARY SCHOOL COMPLETION STATUS OF SPECIAL EDUCATION EXITERS IN TWO YEARS

Disability Category	Percentage of Exiters in 2 Years Who: Graduated	Dropped Out	Aged Out	Sample Size
All conditions	56.2	36.4	7.5	3045
Learning disabled	61.0	36.1	2.9	533
Emotionally disturbed	41.8	54.7	3.6	334
Mentally retarded	49.9	33.6	16.5	459
Speech impaired	62.7	32.5	4.8	222
Visually impaired	69.5	16.8	13.7	279
Deaf	71.8	11.8	16.4	354
Hard of hearing	72.3	15.5	12.2	249
Orthopedically impaired	76.5	15.6	7.9	246
Other health impaired	65.4	25.9	8.7	142
Multiply handicapped	32.2	17.6	50.2	182
Deaf/blind	43.1	7.8	49.2	45

Using a 2-tailed test, the sampling errors at the 95% confidence level for school completion rates for youth in all conditions is ±2%. For categories of disability, the confidence intervals range from ±5% to ±8% (other health impaired). The confidence interval for the deaf/blind category is ±15% for the graduation and age-out rates, due to the small sample size.

Source: School records and parent reports.

groups. Although the graduation rates for youth with orthopedic, visual, or hearing impairments approach the rate for the general population, the graduation rates for youth with emotional disturbances, mental retardation, or multiple handicaps are below 50% ($p<.005$).

Table 10 further demonstrates that overall, about 8% of special education exiters left school because they exceeded the school age limit. Youth with multiple handicaps, including those who are deaf and blind, were most likely to age out of school (about 50%); about 16% of deaf and mentally retarded youth aged out, and fewer than 5% of youth with learning, speech, or emotional impairments aged out ($p\leq.01$).

More than 1 in 3 exiters from the secondary special education system dropped out of school (36%) in a two-year period, with variation between disability categories. The dropout rate for youth with emotional disturbances, for example, was almost 55%, compared to significantly lower rates for youth with sensory or orthopedic impairments (between 12% and 17%; $p\leq.01$). Youth with learning disabilities, who are the majority of secondary special education students, had a dropout rate of 36%.

Earlier research on dropouts from special education in single states or small samples of districts reports dropout rates in a similar range. For example, state studies have reported dropout rates that range from 31% for mildly impaired youth in several districts in Florida (Fardig, et al., 1985) and 34% in Vermont (Hasazi, Gordon, and Roe, 1985), to 40% for special education students overall in New Hampshire (Lichtenstein, 1988). In urban districts, the rates appear to be higher. Prior research has reported dropout rates for youth with learning disabilities in urban areas that are as high as 42% (Cobb and Crump, 1984), 47% (Levin, Zigmond, and Birch, 1985), 50% (Edgar, 1987), and 53% (Zigmond and Thornton, 1985).

Relating Student Characteristics to School Achievement

Thus far, we have described several aspects of the educational programs and school achievement of secondary students with disabilities. One intent of multivariate analyses for the National Longitudinal Transition Study is to relate programs to achievement. However, before we can fully understand what helps or hinders youth in achieving in school, it is important to understand what kinds of youth have difficulty achieving. What student characteristics relate to school achievement?

Analysis Procedures

To answer this question, we have performed multivariate analyses of one aspect of secondary school achievement: the extent to which youth receive failing grades in school. (Multivariate analyses relating student characteristics to drop out behavior using NLTS data are reported in Butler-Nalin and Padilla, 1989). The dependent variable is a dichotomous variable with a value of 1 if youth were reported by their schools to have received a failing grade in any class in their most recent year in secondary school and a value of 0 if they received passing grades in all courses for which grades were given. Logistic regression analyses were performed using this dichotomous measure as a dependent variable.

Analyses include all youth for whom grades were available and who received grades in at least one class. Youth in completely ungraded programs are eliminated from the analysis because the nature of their program prohibits them from varying on the dependent measure.

Because educational programs and school achievement vary so much based on the disability of the youth, as the descriptive analyses have demonstrated, multivariate analyses are reported separately for youth in 5 major disability groupings. Analyses are reported for these larger groups, rather than for each of the 11 individual disability categories, because the sample size for many categories is too small for the complex explanatory

models developed. Groups are defined to maximize the homogeneity of disabilities of youth within the groups.

Group 1 includes youth that have learning disabilities, emotional disturbances or speech impairments (referred to as LESI), who are not institutionalized and not also mentally retarded. Group 2 includes youth with mild mental retardation (EMR) who may or may not also have other impairments; youth with moderate mental retardation are largely eliminated from these analyses because very few are in graded programs. Group 3 involves youth with health or orthopedic impairments who are not also mentally retarded (referred to as physically impaired). Group 4 includes youth who are deaf or hard of hearing and not also mentally retarded. Group 5 is youth who are visually impaired and not also mentally retarded. Severely impaired youth are not included in the analyses because of the requirement that they be in a graded educational program, an uncommon occurrence for this group.

Logistic regression results are unweighted, unlike the descriptive findings reported in the paper thus far. Sampling weights are based on the primary disability category of the youth and enhance the generalizability of descriptive findings (see appendix). However, when youth from different disability categories are combined into larger groupings for the multivariate analyses, youth with vastly different weights are combined. Results are skewed and generalizable primarily to youth with larger weights. For example, in the LESI group, youth with learning disabilities have much larger weights than youth with speech impairments or emotional disturbances because youth with learning disabilities comprise about half of special education students at the secondary level. Weighted analyses of the LESI group, therefore, would be dominated by youth from the LD category and would not illuminate factors affecting school achievement of youth with speech impairments or emotional disturbances. Unweighted analyses better represent the mixture of disability types within the disability groups.

Independent Variables

Three kinds of independent variables related to student characteristics are used to help explain variations in youths' receipt of failing grades: demographic characteristics of the youth, factors related to their abilities and disabilities, and measures of selected behaviors and experiences. The independent variables are described below. Descriptive statistics for the independent variables are included in the appendix.

Characteristics of the Youth

Research on nondisabled youth has demonstrated the effects of several demographic characteristics on school achievement. Analyses of High School and Beyond data, for example, indicate that males, minorities, youth with lower cognitive ability, and those from households with lower socioeconomic status have lower school achievement, as measured by grade point average (Fetters, Brown, and Owings, 1984). In earlier research, similar relationships between test scores and SES and cognitive ability were found by Bachman for 10th grade boys (Bachman, 1970). Do similar relationships hold for youth with disabilities when receipt of failing grades is the focus? To test the effects of demographics on receipt of failing grades for youth with disabilities, the following variables were included in the analyses. Most background characteristics are based on parent reports.

- The youth's age.
- The youth's gender (1=male; 0=female).
- Ethnic background (1=minority excluding Asian, 0=white or Asian).
- Socioeconomic status, measured by the educational level of the head of household (1=no high school diploma, 2=high school graduate, 3=some college education, 4=college degree or more) and whether the head of household is employed.
- Urbanicity, measured by 2 dichotomous variables indicating if the youth attends school in an urban area or a rural area. The comparison condition is attending school in a suburban area.

Although the analyses are conducted separately for youth in different disability groupings, within ups there is still considerable variation in the combination and severity o disabilities, which could affect receipt of failing grades. Therefore, several variables related to variations in disability within disability groupings are included in the analyses:

- The youth's IQ, as reported by his/her school.
- The youth's functional ability, measured by a scale based on parents' reports of how well youth perform 4 functional tasks on his/her own, without help: counting change, telling time on a clock with hands, reading common signs, and looking up names in the telephone book and using the telephone. Youth were scored from 1 (does the task "not at all well") to 4 (does the task "very well") on each task. Summing these scores on the 4 tasks creates a scale ranging from 4 to 16.
- For youth in the LESI group, 2 dichotomous variables are used to designate whether schools reported youth to have a speech impairment or an emotional disturbance among their disabilities. The comparison group is youth with learning disabilities alone.
- For the EMR group, 3 dichotomous variables distinguish youth whose schools reported they have a speech disability, an emotional disturbance, or a physical or sensory disability, in addition to their mental retardation. One might expect that having any of these disabilities, in addition to the mental retardation that qualified the youth for this group, might further challenge the youth's ability to earn passing grades.
- For the physically impaired group, a dichotomous variable distinguishes youth whose parents reported they used a physical aid, such as a wheel chair, crutches, cane, walker, prosthetic, or orthotic, from those who do not. Physical functioning is measured using a scale based on parents' reports of how well the youth could perform 3 basic self-care tasks on his/her own, without help: dress oneself, feed oneself, and get around to places outside the home, such as a nearby park or neighbor's house. Youth were scored from 1 (does the task "not at all well") to 4 (does the task "very well") on each task. Summing these scores on the tasks creates a scale ranging from 3 to 12.
- For the hearing impaired group, a dichotomous variable distinguishes youth who were categorized by their school or district as deaf from those who were labeled hard of hearing. A second dichotomous variable distinguishes youth who were reported by parents as having trouble with their disability before the age of three from those who reportedly began having trouble at a later age. This variable controls primarily for the effects of variations in speech acquisition.

- For the visually impaired group, a dichotomous variable distinguishes youth who were categorized by their school or parent as completely blind from those who were labeled partially sighted.

In addition to their demographic and disability-related characteristics, youth exhibited particular behaviors and had some experiences that are expected to influence their grades. These variables include:

- Whether the youth had disciplinary problems. A dichotomous variable distinguishes youth whose parents reported they had one or more of a specific set of disciplinary problems from those who reportedly had none of them. These disciplinary problems include: ever being fired from a job, leaving school because of suspension or expulsion, or ever being arrested or incarcerated. We hypothesize that youth who experienced disciplinary problems are more likely also to have received failing grades in school.

- Absenteeism from school is a continuous variable measuring the number of days absent from school, as reported in school records, truncated at 60 days. High absenteeism is expected to increase the likelihood of receiving failing grades.

- Prior school achievement is measured by a dichotomous variable indicating if the youth is older than the typical age-for-grade, suggesting that he/she repeated an earlier grade. We expect youth who repeated an earlier grade to be more likely to have received failing grades in school in their most recent year.

- The degree of social integration of the youth is measured by a dichotomous variable indicating whether parents reported that the youth belonged to any school or community group in the past year. Youth who do not belong to any such groups are expected to be disproportionately represented among those who received failing grades.

- Whether the youth had a job in the past year is indicated by a dichotomous variable distinguishing youth whose parents reported they had a workstudy job (either paid or unpaid) or other work for pay (whether sheltered or competitive) in the past year from youth whose parents reported they had neither kind of job. Research is mixed on the effects of employment on school achievement (Greenberger and Steinberg, 1986) and the direction of its effect in these analyses is not hypothesized.

Findings

Table 11 presents findings of logistic regression analyses explaining variations in whether youth received any failing grades in their most recent year in secondary school.

Across the disability groups, the unweighted percentage of youth who received a failing grade in the most recent year ranges from 36% of youth with learning disabilities, emotional disturbances, or speech impairments (LESI) to 13% of youth in the hearing impaired group.

The independent variables together are significant predictors of receipt of failing grades for all groups of youth ($p \leq .001$). However, not all variables have a consistent affect across all the disability groups, i.e., what significantly relates to receipt of failing grades for youth with one kind of disability may not be related significantly to the dependent measure for youth with other kinds of disabilities. This underscores the need for individualized approaches to special education programs. Variations in findings across groups of youth are noted below.

Demographic Characteristics

- Younger students were more likely to receive failing grades than were older students. The relationship between age and receiving failing grades is negative for all disability groups, is statistically significant for youth in the LESI, EMR, and visually impaired groups ($p \leq .001$ to .05), and approaches significance* for the physically and hearing impaired groups. This finding is consistent with the descriptive results discussed earlier, and may result either from the preponderance of more successful students among those who remain in school until the upper grades or from variations in grading policies and practices across grade levels.

---

* Relationships are considered to approach statistical significance if $p > .05$ but $< .10$.

Table 11: FACTORS ASSOCIATED WITH RECEIPT OF FAILING GRADES

	Disability Group				
	LESI	EMR	Physical	Hearing	Visual
Percent of youth failing	36.4	20.2	22.0	14.6	16.2
Youth Demographics					
Age	-.14***	-.15*	-.19	-.12	-.32**
Youth is male	.56***	.30	.87**	-.06	.86*
Youth is minority	.51**	.78**	.36	-.12	.12
Head of household education	-.08	-.25	-.02	.10	-.13
Youth is in a single parent household	.06	-.12	-.75*	.03	.15
Head of household is employed	-.05	.14	-.52	-.26	-.14
Youth lives in an urban area	.10	-.58	-.16	.42	.34
Youth lives in a rural area	-.05	.01	-.47	.05	-.40
Abilities/disabilities					
IQ	-.00	.02	.02	-.01	-.02
Youth's functional ability	-.07	.08	.05	.10	.14*
Has a speech disability	.16	-.24			
Has an emotional disturbance	.43**	.69			
Has sensory/physical disability		-.23			
Youth began having hearing difficulty before age 3				-.34	
Youth is deaf				-.73**	
Youth is blind					-.62
Youth's self-care ability			.30**		
Youth uses physical device			-.27		
Youth behaviors/experiences					
Number of days absent from school	.05***	.04***	.05***	.06***	.04**
Youth belongs to school/community group	-.28	-.35	-.61	-.62**	-.51
Youth has had disciplinary problems	.56**	.40	--	.38	--
Youth had a job in the past year	-.15	-.73**	-.10	.03	-.17
Youth was held back 1 or more grades	.07	.10	.78*	.17	.53
Number of classes for which grades were received	.20***	.08	.49***	.38***	.34**
N	1109	559	341	773	322
χ^2	214.2	103.7	91.6	119.1	56.6
d.f.	18	19	17	18	16
p<	.001	.001	.001	.001	.001

*=p≤.05; **=p≤.01; ***=p≤.001; ---=Too few cases to include in the model.

- Male students were generally more likely than females to receive failing grades in school. For 4 of the 5 disability groups, being male is associated with a higher likelihood of receiving failing grades; the relationship is significant for youth in the LESI, physically impaired, and visually impaired groups ($p \leq .001$ to .05). This is consistent with findings from High School and Beyond that males in secondary school had generally lower grade point averages than females (CES, 1984).

- Minority youth in the LESI and EMR groups received failing grades at a significantly higher rate than other youth in those groups ($p \leq .01$), controlling for selected measures of socioeconomic status, IQ, and other factors in the models.

- For youth in the physically impaired group, being in a 2-parent household appears to increase the likelihood the youth with receive failing grades. This finding is counterintuitive and calls for additional investigation.

Factors Related to Youths' Abilities/Disabilities

- Among youth in the LESI group, students with an emotional disturbance were significantly more likely than youth with learning disabilities alone to receive failing grades ($p \leq .01$).

- In general, for most groups of youth, less severely impaired youth were more likely to receive failing grades. For example, among youth with visual impairments, youth with higher functional abilities were more likely to receive failing grades ($p \leq .05$). A similar relationship approaches significance for youth in the EMR group. For youth with physical impairments, those who were reported by parents to function better in terms of self-care skills were significantly more likely to receive failing grades ($p \leq .05$). Similarly, among those with hearing impairments, youth who are hard of hearing were significantly more likely than those who are deaf to receive a failing grade ($p \leq .01$). These findings are independent of the number of courses taken for which grades were received. These relationships may be due to the fact that less severely impaired youth are generally more likely to be enrolled in mainstreamed classes, for which grading standards are often stiffer than in special education placements. Or, perhaps even within a given placement, it may be that different grading policies or standards are applied to youth with varying levels of disability; i.e., perhaps teachers expect more of and, therefore, grade more stringently, youth with milder disabilities.

Youths' Behaviors and Experiences

- Youth who were absent frequently from school were significantly more likely to receive failing grades. This relationship is consistent and significant for all groups ($p \leq .001$ or .01). Caution should be exercised in interpreting this finding, however. Although some absenteeism from school for special education students relates to the their disability, much absenteeism at the secondary school level is voluntary. It is not clear whether voluntary absenteeism is a causal factor in receiving failing grades or an outgrowth of it; we do not know if absence from school results in students missing lessons and, therefore, receiving poor grades or whether, knowing they are doing poorly in school, students avoid the school environment and exhibit high absenteeism.

- Youth who do not belong to a school or community group tended to receive failing grades at a higher rate than youth who were involved in such groups. Group membership is associated with a reduced likelihood of receiving a failing grade for all groups, is significant for youth with hearing impairments ($p \leq .01$), and approaches significance for youth in the LESI and physically impaired groups. Again, alternative explanations of this finding are possible. Perhaps group membership increases the bonds between special education students, other students, and school, helping youth with disabilities to meet the expectations of the school environment and avoid receiving failing grades. However, it is also possible that unmeasured aspects of the students explain this relationship. Students with a greater degree of confidence and competence may be more likely to take the social risks inherent in group membership; these students may also be prone to do better in school. The absence from the model of measures of these dimensions of the youth may lead to the apparent relationship between group membership and a reduced likelihood of receiving failing grades.

- Youth who have had disciplinary problems were generally more likely to receive failing grades; this relationship is statistically significant for youth in the LESI group ($p \leq .01$). The effect of having behavior problems is independent of having an emotional disturbance, which is controlled for separately in the model.

- Youth who took more graded classes and, therefore, had more opportunities to receive a failing grade, were significantly more likely to receive such grades than youth who took more courses for which grades were not given. This relationship is consistent in direction across all groups and is significant for all but youth with mild mental retardation ($p \leq .01$ or .001)

Beyond these findings regarding significant effects of individual characteristics on receipt of failing grades, we should also comment on the absence of statistically significant relationships for some variables. Conventional wisdom and prior research have suggested that, for non-handicapped youth, several characteristics of youth have relationships to school achievement. For example, analyses of High School and Beyond data (NCES, 1984) suggest that youth from households with lower socioeconomic status have lower grade point averages in secondary school.

Although we have found no consistent or significant direct relationship between SES and school achievement as measured by receipt of failing grades, we should not conclude that socioeconomic status has no effect on the dependent measure. Other variables entered in the model may more directly measure factors for which SES variables often proxy. For example, being absent frequently from school is positively and significantly correlated with low SES ($p \leq .001$), as are other behavioral factors included in the models. When we omitted from the models variables related to disciplinary problems, being older than the typical age-for-grade (suggesting earlier grade level retention) grade, and absenteeism from school, one measure of SES, head of household education, had significant effects on receipt of failing grades in the expected direction. Hence, behavioral variables are apparently absorbing variation that would be attributed to SES if behavioral factors were not measured directly. With behavioral factors included in the model, SES has a relatively small independent direct effect on receipt of failing grades, but an additional indirect effect through its behavioral manifestations.

The absence of apparent relationship between IQ and receipt of failing grades also deserves mention. The fact that IQ does not have a significant effect on receipt of failing grades in these models is not completely surprising. Eliminating from the analyses youth in ungraded programs reduces the variation in IQ within each group. The limited variation remaining may be insufficient to distinguish youth who receive failing grades.

Summary and Next Steps

The findings reported here offer much new information regarding the school programs of secondary youth with disabilities:

- A majority of secondary students with disabilities attended comprehensive secondary schools with nondisabled students (89%).
- Attending special schools for youth with disabilities was most common for youth with sensory or multiple impairments; 35% of youth with visual impairments, 63% of youth who are deaf, 41% of youth with multiple impairments and 94% of youth who are deaf/blind attended such schools, compared to 8% of special education students overall.
- About 5% of students were declassified from special education during their most recent year in secondary school; youth in the speech impaired category were most likely to be declassified (18%).
- Most special education students (83%) were enrolled in some regular education courses; enrollment in regular education courses ranged from 90% for youth with learning disabilities to 68% of youth with mental retardation to about 30% of youth with multiple disabilities, including those who are deaf/blind.
- More than half of special education students (54%) were enrolled in one or more vocational education courses in their most recent year in school; participation in vocational education exceeded 80% of youth in 11th and 12th grades.
- Schools were the primary provider of services such as speech therapy, personal counseling, and occupational therapy for secondary special education students. More than half of the students received none of the services we investigated as adjuncts to their special education instructional program.

New insights are also provided on the school achievement of secondary special education students:

- Almost 1 in 3 students who received grades received a failing grade in 1 or more courses in their most recent year in school; receipt of failing grades ranged from 45% of youth with emotional disturbances to 6% of youth with multiple impairments.
- About 3 of 4 students in graded programs in grades 7-11 were promoted to the next grade level at the end of the year (74%), 6% were held back.

- Almost two-thirds of students (62%) were required to take minimum competency tests. More than 3 of 4 students tested (76%) passed all or part of the requirements.

- More than 1 in 3 special education students who left school in a 2-year period dropped out of school without graduating (36%). Dropout rates are lowest for youth with sensory and multiple disabilities (from 8% to 17%) and highest for youth with emotional disturbances (55%).

When we examine factors associated with receipt of failing grades, several relationships are suggested. Many of the factors related to receipt of failing grades are characteristics of the youth that are not affected by school experiences (e.g., ethnicity, gender). These analyses demonstrate relationships that are largely consistent with findings for nonhandicapped students. Other factors affecting receipt of failing grades are behavioral, such as absenteeism from school, disciplinary problems, and lack of membership in school or community groups. Alternative interpretations of these relationships have been pointed out; it is unclear whether these factors contribute to poor grade performance or whether they are simply associated symptoms. In either case, educators can consider them warning signs of students who are at risk of failing in school.

Continuing NLTS analyses will give further attention to the relationships suggested here. A primary focus will be to add to these models variables related to educational programs, services, and schools to determine what factors that can be influenced by schools and other service providers relate to improved school performance. In addition, we will be examining the wide variation in receipt of particular programs and services within disability categories and identifying individual, school, and environmental factors that help explain variations in service patterns. As the study moves into its later years and longitudinal data are available on more youth as they leave school, analyses will focus on associations between school experiences and later transition outcomes.

REFERENCES

Bachman, J. G., (1970). Youth in Transition, Vol. II: The Impact of Family Background and Intelligence on Tenth-Grade Boys. Ann Arbor, MI: Institute for Social Research.

Butler-Nalin, P. and Padilla, C. (1989). Factors Affecting Special Education Dropouts: Findings from the National Transition Study. Presented at meetings of the American Educational Research Association, San Francisco, CA, March, 1989.

Center for Education Statistics (1986). The Condition of Education. Washington, D.C.

Cobb, R. and Crump, W. (1984). Postschool Status of Young Adults Identified as Learning Disabled While Enrolled in Learning Disabilities Programs. Final report, USDE Grant No. G008302185. University, AL: University of Alabama.

Council of the Great City Schools (1986). Special Education: Views from America's Cities. Philadelphia, PA: Research for Better Schools.

Edgar, E. (1987). Secondary programs in special education: Are many of them justifiable? Exceptional Children, 53, 555-561.

Fardig, D.B., Algozzine, R.F., Schwartz, S.E., Hensel, J.W., and Westling, D.L. (1985). Postsecondary vocational adjustment of rural, mildly handicapped students. Exceptional Children, 52, 111-121.

Fetters, W.B., Brown, G.H., and Owings, J.A. (1984). High School Seniors: A Comparative Study of the Classes of 1972 and 1980. Washington, D.C.: National Center for Education Statistics.

Hayward, B. (1989). Access and Equity: Participation of Handicapped High School Students in Vocational Education. Washington, DC: Policy Studies Associates, Inc.

Hasazi, S.B., Gordon, L.R., and Roe, C.A. (1985). Factors associated with the employment status of handicapped youth exiting high school from 1979-1983. Exceptional Children, 51, 455-469.

Greenberger, E. and Steinberg, L. (1986). When Teenagers Work. New York, NY: Basic Books.

Levin, E., Zigmond, N. and Birch, J. (1985). A followup study of 52 learning disabled students. Journal of Learning Disabilities, 18 (1), 2-7.

Lichtenstein, S. (1988). Droupouts: A secondary special education perspective. Counterpoint, 8 (3), 13.

Moore, M., et al. (1988). Patterns in Special Education Service Delivery and Cost. Washington, DC: Decision Resources Corp.

National Center for Education Statistics (1984). Two Years in High School: The Status of 1980 Sophomores in 1982. Washington, D.C. U.S. Government Printing Office.

Singer, J., et.al. (1988). Quoted in Viandero, D., Special education is not a "one-way street" for all, study finds. Education Week, May 18, 1988.

U.S. Department of Education (1987). State Education Statistics Wallchart. Washington, D.C.: U.S. Government Printing Office.

U. S. Department of Education (1988). "To Assure the Free Appropriate Public Education of All Handicapped Children": Tenth Annual Report to Congress on the Implementation of the Education of the Handicapped Act. Washington, D.C.: U.S. Department of Education.

Wagner, M. (1989). Influences on the Transition Experiences of Youth with Disabilities: A Report from the National Transition Study. Presented at meetings of the Council for Exceptional Children, San Francisco, CA, April 1989.

Zigmond, N. and Thornton. H. (1985). Learning disabled graduates and dropouts. Learning Disabilities Research, 1 (1), 50-55.

Appendix

OVERVIEW OF THE NATIONAL LONGITUDINAL TRANSITION STUDY OF SPECIAL EDUCATION STUDENTS

As part of the 1983 amendments to the Education of All Handicapped Children Act (EHA), the Congress requested that the U.S. Department of Education conduct a national longitudinal study of the transition of secondary special education students to determine how they fare in terms of education, employment, and independent living. A 5-year study was mandated, which was to include youth from ages 13 to 21 who were in special education at the time they were selected and who represented all 11 federal disability categories.

In 1984, the Office of Special Education Programs (OSEP) of the U.S. Department of Education contracted with SRI International to determine a design, develop and field test data collection instruments, and select a sample for the National Transition Study. In April 1987, under a separate contract, SRI began the actual study.

Study Components

The National Transition Study has four major components:

- *The Parent/Youth Survey*. In the first year of the study, parents were interviewed by telephone to determine information on family background and expectations for the youth in the sample, characteristics of the youth, experiences with special services, the youth's educational attainment (including postsecondary education), employment experiences, and measures of social integration. This survey is expected to be repeated in 1989, when the youth will be interviewed if he/she is able to respond.

- *School Record Abstracts*. Information has been abstracted from the school records of sample youth for the previous year or for the last year they were in secondary school (either the 1985-86 or 1986-87 school years). Information abstracted from school records relates to courses taken, grades achieved (if in a graded program), placement, related services received from the school, status at the end of the year, attendance, IQ, and experiences with minimum competency testing. Records will be abstracted again in 1989 for youth still in secondary school in the 1988-89 school year.

- *School Program Survey*. Schools attended by sample youth in the 1986-87 school year were surveyed for information on student enrollment, staffing, programs and related services offered secondary special education students, policies affecting special education programs and students, and community resources for the disabled.

- *Explanatory Substudies*. More in-depth studies involving subsamples of the main sample will examine the pattern of transition outcomes achieved by youth who are out of secondary school and the relationship between school experiences and transition outcomes.

Sampling

Youth were selected for the sample through a two-stage sampling procedure. A sample of 450 school districts was randomly selected from the universe of approximately 14,000 school districts serving secondary (grade 7 or above) special education students, which had been stratified by region of the country, a measure of district wealth involving the proportion of students in poverty (Orshansky percentile), and district size (student enrollment).* Because of a low rate of agreement to participate from these districts, a replacement sample of 176 additional districts was selected. In addition, participation in the study was invited from the approximately 80 special schools serving secondary-age deaf, blind, and deaf-blind students. A total of approximately 300 school districts and 25 special schools agreed to have youth selected for the study.

Analysis of the potential bias of the district sample indicates no systematic bias that is likely to have an impact on study results when responding districts were compared to nonrespondents on the types of disabilities served, special education enrollment, participations in Vocational Rehabilitations agency programs, the extent of school-based resources for special education, community resources for the disabled, the configuration of other education agencies serving district students, metropolitan status, percent minority enrollment, grades served, and the age limit for service (see Javitz, 1987 for more information on the LEA bias analysis).

The sample of students was selected from rosters of all special education students ages 13 to 21 who were in grades 7 through 12 or whose birthdays were in 1972 or before. The roster of such students was stratified into 3 age groups (13 to 15, 16 to 18, over 18) for each of the 11 federal handicap categories and youth were randomly selected from each age/condition group so that at least 1,000 students would be selected in each handicap category (with the exception of deaf-blind, a low-incidence condition).

Exhibit A-1 indicates the number of youth sampled in each condition, the proportion for which different combinations of data were obtained, and the reasons for nonresponse for youth for whom data could not be obtained. A study of potential nonresponse bias is now being conducted to determine the representativeness of the youth sample.

Weighting Procedures and Population to Which Data Generalize

Youth with disabilities for whom data could be gathered were weighted to represent the U.S. population of such youth. In performing this weighting, three mutually exclusive groups of sample members were distinguished:

---

* The 1983 Quality Education Data, Inc. (QED) database was used to construct the sampling frame. QED is a private nonprofit firm located in Denver, Colorado.

Exhibit A-1

Student Sample by Handicapping Condition

Status	LD	SED	MR	Speech	Ortho	Deaf	H of H	Blind	D/B	Health	Multi	Total
Number of contacts	1650	1321	1642	933	1060	1050	1372	1318	165	1005	1132	12648
No Further Contact Possible												
Unable to locate	59	59	84	50	49	41	70	63	5	33	45	558
Names not provided by LEA	206	271	55	92	18	99	197	120	0	362	212	1632
Deceased	2	0	4	0	11	0	3	2	3	5	2	32
Language barrier/non-Spanish	5	4	5	9	6	12	13	3	0	5	2	64
No respondent exists	23	21	28	18	9	20	11	20	2	9	16	177
Other	3	3	7	5	1	14	6	2	3	5	6	55
Nonworking number	233	178	341	157	146	149	180	193	29	115	94	1815
TOTAL	531	536	524	331	240	335	480	403	42	534	377	4333
(Percentage of total contacts)	32	41	32	35	23	32	35	31	25	53	33	34
Responses												
Completed interview-have consent form	506	326	533	232	388	402	470	475	73	246	362	4013
Completed interview-no consent form	385	258	314	217	216	259	231	255	35	131	159	2460
Total completed interviews	891	584	847	449	604	661	701	730	108	377	521	6473
(% of total contacts)	54	44	52	48	57	63	51	55	65	38	46	51
(% of those to be interviewed)	64	59	57	57	62	73	64	64	69	62	60	62
Have partial data (other sources)	37	43	42	18	35	15	15	20	2	11	24	262
Have partial interview (phone)	39	25	27	25	16	26	17	17	4	19	22	237
Have partial interview (mail)	20	21	49	15	25	23	17	20	4	10	30	234
Total participation	987	673	955	507	680	725	750	787	118	417	597	7206
(% of total contacts)	60	51	59	34	64	69	55	60	72	41	53	57
(% of those to be interviewed)	71	68	64	64	69	80	68	69	75	69	68	69
Refused interview	56	41	40	11	30	19	24	22	3	18	18	282
Refused in earlier contacts	11	3	6	2	20	0	1	3	1	3	9	59
Total refusals	67	44	46	13	50	19	25	25	4	21	27	341
(% of total contacts)	4	3	3	1	5	2	2	2	2	2	2	3
(% of those to be interviewed)	5	4	3	2	5	2	2	2	3	3	3	3
	29	20	19	22	8	34	18	18	4	14	22	238

A. Youth whose parents responded to the telephone-administered Parent Interview.

B. Youth whose parents did not respond to the telephone-administered Parent Interview, but were interviewed in the in-person nonrespondent study.

C. Youth whose parents did not respond to either the telephone or in-person Parent Interview, but for whom the school provided a record abstract.

All sample members belong to one of these three groups.

A primary concern in performing the weighting was to determine whether there was a nonresponse bias and to calculate the weights in such a way as to minimize that bias. Nonresponse bias was primarily of three types:*

1. Bias attributable to the inability to locate respondents because they had moved or had nonworking telephone numbers.

2. Bias attributable to refusal to complete a parent interview.

3. Bias attributable to circumstances that made it infeasible for the record abstractors to locate or process a student's record.

Of these three types of nonresponse, the first was believed to be the most important, both in terms of frequency and influence on the descriptive and explanatory analysis. Type 1 bias was also the only type of nonresponse that we could estimate and correct.

We estimated the magnitude of type 1 nonresponse bias by comparing responses on identical (or very similar) items in the three groups of respondents (after adjusting for differences in the frequency with which different handicaps were selected and differences in the size of the LEAs selected). Group A respondents were wealthier, more highly educated, and more likely to be Caucasian than group B respondents. In addition, group A respondents were much more likely to have youth who graduate from high school than group B or C respondents (who had similar dropout rates). On all other measurable items, the youth described by the three groups were similar, including sex, employment status, pay, self-care skills scale, household-care activities scale, functional mental skills scale, association with a social group, and length of time since leaving school. SRI determined that

---

* In addition, there was a large group of nonrespondents who could not be located because their LEAs would not provide student names. Presumably, had these student names been available, many of those nonrespondents would have chosen to participate at about the same rate as parents in districts in which youth could be identified. The remaining nonrespondents would presumably have been distributed between the three types of nonresponse mentioned above.

adjusting the weights to eliminate bias in the income distribution would effectively eliminate bias in parental educational attainment and racial composition, but would have a negligible effect on dropout rates. It was also determined that group B and C respondents were present in sufficient numbers that if they were treated as no different from the group A respondents in the weighting process, the resultant dropout distribution would be approximately correct.

Weighting was accomplished using the following sequence of steps:

(1) Data from all three groups were used to estimate the income distribution for each handicapping condition that would have been obtained in the absence of type 1 nonresponse bias.

(2) Respondents from all three groups were combined and weighted up to the universe by handicapping condition. Weights were computed within strata used to select the sample (i.e., LEA size and wealth, and student age).

(3) Weights from four rare handicapping conditions (deaf/blind, deaf, orthopedically impaired, and visually impaired) were adjusted to increase the effective sample size. These adjustments primarily consisted of slightly increasing the weights of students in larger LEAs and decreasing the weights of students in smaller LEAs. Responses before and after these weighting adjustments were nearly identical, except for the deaf/blind. The adjustment for the deaf/blind consisted of removing a single respondent from a medium-sized LEA, who was being weighted up to represent two-thirds of all deaf/blind students. Hence, survey results do not represent deaf/blind students in medium or smaller-sized LEAs.

(4) The resultant weights were adjusted so that each handicapping condition exhibited the appropriate income distribution estimated in step 1 above. These adjustments were of modest magnitude (relative to the range of weights within handicapping condition)--the weights of the poorest respondents were multiplied by a factor of approximately 1.6 and the weights of the wealthiest respondents were multiplied by a factor of approximately 0.7.

Statistical Tests

A statistical procedure was used to compute the approximate standard errors of proportions and to test the difference between two proportions. We first computed the weighted percent of "yes" respondents to a survey item and then computed the effective sample size (i.e., the sum of the weights squared, divided by the sum of the squared weights). These two quantities were then used in the usual formula for the variance of a binomially distributed variable (i.e., pq/n where p is the weighted proportion of "yes" responses, q is the complement of p, and n is the effective sample size). To test the difference of two weighted proportions, we computed the difference between the weighted proportions and divided this quantity by the square root of the sum of the variances of the two proportions.

This procedure is only approximately correct because it adjusts only for the difference in weights, but not for cluster-sampling induced covariance among respondents. We are currently in the process of using pseudo-replication to compute more accurate variance estimates. We expect that the true variances are larger than calculated by the effective sample size method, and therefore that stated significance levels (e.g., $p < .01$) will be somewhat too small. Consequently, we have tended to be very conservative, and for the most part, highlight results that are significant at the .005 level.

Analysis

The first stage of the analysis study involves producing descriptive findings related to individual and family characteristics of youth, their experiences with services, their secondary school program, and their outcomes in terms of education, employment, and independent living. Descriptive questions include the following:

- What are the individual and family characteristics of handicapped youth served under EHA?
- What educational experiences and related services are handicapped youth provided under EHA? How do these vary for youth with different handicapping conditions and of different ages? What is the content, duration, intensity, coordination, and provider of these services?
- What are the characteristics of the schools serving youth with disabilities (e.g., with respect to grade levels served, programs and staff available, policies and practices regarding students with disabilities)?
- What are the achievements of youth with disabilities related to their education (secondary school and postsecondary), employment, and independence? How do these vary for youth with different kinds of disabilities?
- What combinations of services, experiences, and outcomes form transitional life paths for youth with different kinds of disabilities?

The second analysis stage will involve multivariate analyses to determine the relationships among the variables depicted in the conceptual model. Explanatory questions include:

- What factors combine to explain the patterns of services that youth receive?
- What factors explain the educational, employment, and independence outcomes of handicapped youth?
- What explains the paths youth take through secondary school and beyond with respect to services, experiences, and outcomes?

Reporting

Findings of the study will be presented in several forms through several channels. Statistical almanacs will present all the descriptive information available from the study for the total handicapped youth population and for each individual handicapping condition. Dissemination activities will entail conference presentations, journal articles, and mailings of key findings to participants in the study and others interested in its findings. A series of special topic reports will present findings from analyses addressing specific policy or research questions. Four methodology reports will detail the sampling, data collection, and analysis procedures used for the project and the reliability/validity of findings. A final report to OSEP will provide comprehensive documentation of findings.

MEANS AND STANDARD DEVIATIONS OF INDEPENDENT VARIABLES IN MULTIVARIATE MODELS EXPLAINING RECEIPT OF FAILING GRADES BY SECONDARY STUDENTS WITH DISABILITIES

	Disability Group									
	LESI		EMR		Physical		Hearing		Visual	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Youth Demographics										
Age	17	1.8	18	1.9	17	1.8	18	1.9	18	1.7
Youth is male	.72	.45	.54	.50	.58	.50	.51	.50	.57	.50
Youth is minority	.26	.44	.40	.49	.37	.48	.35	.48	.33	.47
Head of household education	2.1	1.1	2.0	1.2	2.5	1.3	2.3	1.2	2.3	1.2
Youth is in a single parent household	.31	.46	.34	.47	.30	.46	.34	.47	.30	.46
Head of household is employed	.80	.40	.69	.46	.79	.41	.81	.40	.77	.42
Youth lives in an urban area	.32	.47	.35	.48	.60	.49	.48	.50	.36	.48
Youth lives in a rural area	.30	.46	.32	.47	.09	.29	.15	.36	.25	.43
Abilities/disabilities										
IQ	92	13	64	10	92	15	97	13	100	14
Youth's functional ability	14.8	1.7	12.7	3.0	14.4	2.2	14.3	1.9	12.9	2.9
Has any speech disability	.24	.43	.35	.48						
Emotional disturbance is primary disability	.29	.45								
Has emotional disturbance			.10	.30						
Has sensory/physical disability			.43	.50						
Youth began having hearing difficulty before age 3							.76	.43		
Youth is deaf							.66	.47		
Youth is blind									.36	.48
Youth's self-care ability					10.2	2.5				
Youth uses physical device					.49	.50				
Youth behaviors/experiences										
Number of days absent from school	14	13	13	14	15	16	10	10	10	11
Youth belongs to school/community group	.40	.49	.36	.48	.44	.50	.52	.50	.60	.49
Youth has had disciplinary problems	.17	.38	.08	.27	--	--	.05	.23	--	--
Youth had a job in the past year	.74	.44	.57	.50	.45	.50	.64	.48	.58	.50
Youth is older than average age-for-grade	.75	.43	.86	.35	.67	.47	.79	.41	.66	.48
Number of classes for which grades were received	6.6	1.7	6.3	1.8	7.0	1.6	6.9	1.6	7.0	1.5

DEMOGRAPHIC CHARACTERISTICS OF YOUTH WITH DISABILITIES

Characteristics	Total	Primary Disability Category: Learning Disabled	Emotionally Disturbed	Mentally Retarded	Speech Impaired	Visually Impaired	Hard of Hearing	Deaf	Deaf/ Blind	Orthoped-ically Impaired	Health Impaired	Multiply Handi-capped
Age												
15-16	33.0	34.7	36.9	26.5	48.7	29.3	30.9	21.9	9.9	25.2	29.2	30.5
17-18	38.1	40.6	38.9	23.7	33.0	37.2	35.8	29.4	20.5	35.0	40.5	27.5
19-20	22.9	21.7	20.3	27.4	16.1	24.3	22.2	27.8	14.3	30.9	23.5	20.7
≥21	5.9	2.9	3.8	12.3	2.1	9.1	11.1	20.8	55.2	3.9	6.8	21.2
Youth is:												
Male	68.5	73.4	76.4	58.0	59.5	55.6	52.0	54.5	49.5	54.2	56.0	65.4
From 1 parent household	36.8	34.3	44.3	38.6	42.2	36.8	32.0	38.9	30.8	38.5	43.0	36.9
From household with 1986 income <$25,000 per year	67.7	64.9	69.4	73.9	70.2	65.7	64.1	65.7	66.3	66.7	68.5	71.9
Attends school in area that is:												
Urban	31.6	29.2	42.5	29.0	32.4	39.2	44.5	37.9	42.8	40.8	59.7	35.4
Suburban	33.7	36.5	32.8	28.1	35.8	33.0	32.5	40.4	15.5	34.1	16.7	33.6
Rural	34.7	34.3	24.7	43.0	31.8	27.8	22.9	21.7	41.8	25.1	23.5	31.0
Ethnicity												
Black	33.0	21.6	25.1	31.0	28.0	25.9	18.7	24.5	25.0	19.0	20.3	19.1
White	38.1	67.2	67.1	61.0	54.2	63.6	63.4	62.7	67.0	63.1	54.2	65.6
Hispanic	22.9	8.4	6.0	5.6	14.2	8.1	13.6	9.6	5.8	15.1	22.5	12.2
Other	5.9	2.8	1.7	2.4	3.5	2.4	4.3	3.2	2.2	2.8	3.0	3.2
Head of household education												
Less than high school	41.0	37.8	43.7	49.4	46.0	36.6	36.1	33.6	38.5	32.5	35.6	32.4
High school graduate	36.0	39.1	29.1	33.1	28.3	33.0	36.1	36.9	38.2	32.9	28.7	38.4
Some college/2-year degree	14.0	14.5	18.0	10.2	13.0	15.7	14.8	18.7	11.5	17.6	19.1	16.4
College graduate	4.7	4.5	5.1	4.2	5.0	8.5	6.8	5.3	7.0	6.0	8.9	6.1
Graduate studies or degree	4.2	4.1	4.1	3.1	7.6	6.1	6.2	5.4	4.8	11.0	7.8	6.7

DOCUMENT RESUME

ED 306 293 TM 013 163

AUTHOR Nitko, Anthony J.; Pettie, Allan
TITLE The Sixteen Quality Indicators: Standards for Evaluating Criterion-Referenced Tests.
PUB DATE Mar 89
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989). Text contains some small print.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Armed Forces; *Content Analysis; *Criterion Referenced Tests; Evaluation Methods; Formative Evaluation; Military Personnel; Quality Control; *Rating Scales; *Standards; Test Construction; *Test Reliability
IDENTIFIERS *Quality Indicators; *Sixteen Quality Indicators; Skill Qualification Test

ABSTRACT

The development, formative evaluation, and potential uses of the "Sixteen Quality Indicators" (16 QI) rating scale are described. The scale was developed as a systematic way to rate the quality of Skill Qualifications Tests (SQTs) in the United States Army. An SQT measures a soldier's knowledge of a military occupational specialty. It is a criterion-referenced test that samples the tasks in a specific specialty area. Several hundred SQTs are developed annually. The 16 QI is a list of critical criterion-referenced test characteristics. Scale drafts were reviewed by army job specialists and civilian testing experts to form a five-point scale. The 16 QI are grouped into characteristics of the total test, the task-measuring part of the test, and the item. The 16 QI rating scale has not yet been evaluated thoroughly, but would appear to have potential for monitoring SQT quality and diagnosing what needs to be done to improve the quality of a test. As an organized and systematic procedure, the 16 QI may be useful in other applications to evaluate criterion-referenced tests. Four tables present the elements of the 16 QI and the regulations and policies that support its use. (SLD)

The Sixteen Quality Indicators:

Standards for Evaluating Criterion-Referenced Tests

by

Anthony J. Nitko
University of Pittsburgh

and

Allan Pettie
U.S. Army Training Support Center

A paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, California, March, 1989.

Nitko
Pettie

The Sixteen Quality Indicators:

Standards for Evaluating Criterion-Referenced Tests

This paper describes the development, formative evaluation, and potential uses of the Sixteen Quality Indicators (16 QI) rating scale. The scale was developed as a systematic way to rate the quality of Skill Qualifications Tests (SQTs) in the U.S. Army. The concepts used in developing this rating scale may be useful to developing similar instruments for assessing the quality of criterion-referenced test development in other contexts.

Background

SQTs are one part of the U.S. Army's Individual Training Evaluation Program and its Enlisted Personnel Management System. An SQT measures a soldier's knowledge of a military occupational specialty (MOS). An MOS is a job classification (e.g., M48/M60 Armor Crewman). Soldiers must pass the SQT covering their MOS to maintain their certification. The domain of knowledge and abilities for an MOS is defined in detail by a Soldier's Manual which lists the tasks and performances (i.e., objectives) which comprise the MOS. An SQT is a criterion-referenced test that samples the tasks in a specific MOS domain. New forms of an SQT are developed for each MOS each year.

The Army Training Support Center (ATSC) provides guidance for the development of each SQT, but the actual development is the responsibility of one of the 21 proponent Army Training schools. Several hundred SQTs are developed each year and because of cost factors, the quality of only selected SQT is monitored. Guidance to the training schools' development staff is provided through test development regulations, specifically Regulation

351-2, Skill Qualification Test and Common Task Test Development Policy and Procedures.

In spite of the regulations, however, SQT quality is generally uneven. Implementation of SQT development principles varies widely from school-to-school, from MOS-to-MOS, and from one year's SQT to the next year's SQT. The regulations provide policy and guidance, but do not articulate a specific set of quality standards for systematically monitoring, in a relatively objective way, the quality of SQT scores. Without systematic monitoring it is difficult to (a) identify MOSs having better SQTs, (b) target special help to schools most in need of it, and (c) identify test development practices highly related to high quality SQTs.

One approach to this problem is to identify a small set of criterion-referenced test quality indicators and to organize these indicators into a standardized scale that can be used to systematically monitor the SQTs developed by each school. The set of indicators should meet psychometric validity and reliability criteria and be practical to use. Each quality indicator should be (a) related directly to the technical quality of the SQT scores which decision-makers use, (b) linked closely to existing policy, regulations, and accepted test development practices, and (c) of considerable diagnostic value for test developers who are charged with improving the tests.

Method of Developing the 16 QI

The authors have several years' experience in reviewing SQTs and working with SQT developers. Using this experience and suggestions for criterion-referenced test development in the psychometric literature, a list of critical criterion-referenced test characteristics were developed.

This list was refined by examining Regulation 351-2 and assuring that the quality indicators in the list were explicitly or implicitly implied by the official policy on SQT development. Scale drafts were circulated among ATSC staff members associated with SQT development and among civilian testing experts whom the Army had hired to review SQT quality in the recent past. The result was a list of 16 critical SQT characteristics which needed to be evaluated if the quality of an SQT was to be measured. Table 1 summarizes these characteristics.

The characteristics can be organized in several ways. One reviewer suggested organizing them by the categories: content adequacy, item-writing quality, and technical quality. This organization focuses on the nature of the expertise needed by a person to use the characteristics in evaluating an SQT. However, Table 1 shows the way chosen to organize them: quality of the total test scores, quality of the task (subtest) scores, and quality of the test items. This focuses evaluations of SQTs on the nature of the decisions which tend to be made from them. For example, a soldier must pass the total SQT with a minimum passing score of 60 on a standardized scale. Failing to pass places a soldier's MOS certification in jeopardy. Similarly, the regulations encourage individual soldier and group remediation of those who fail specific tasks' tests within an SQT. Task test scores are used for this purpose. Finally, since the subtest and total test scores are linked directly to the quality of the test items, it was deemed important to focus a good part of the evaluation on them.

The quality characteristics selected for the rating scale need to be justified not only on psychometric grounds but on policy grounds as well. A testing program is driven by the policy and decision context in which it will be used. Each of the characteristics selected for inclusion in Table 1 was

supported by some portion of the regulations pertaining to the SQT program. As an example, consider the second characteristic listed in Table 1, "decision consistency of the total score". The policy statements and regulations pay considerable attention to the minimum passing score and the use of SQT results to make pass-fail decisions. Table 2 illustrates how the regulations support the use of decision-consistency as one quality indicator of SQTs. Details of how each quality indicator is supported by Army policy and regulations are given elsewhere (Nitko, 1988).

Each quality indicator then needed to be oeprationalized before a scale could be formed. This required revieiwing the psychometric literature to identify recommended ways to measure or rate each quality characteristic. A number of indices for measuring decision-consistency , for example, have been presented in the literature (e.g., see Beck, 1984 and Subkoviak, 1984 for reviews). In this instance, Subkoviak's (1988) procedure for estimating Kappa coefficient, which uses coefficient alpha and a spcial table, was used because (a) coefficient alpha is a reasonably accurate indicator of an SQT's reliability, (b) this coefficient is calculated already by ATSC in connection with its item analysis report for each SQT, (c) the special tables provided by Subkoviak are relatively short and easy to use, and (d) only one administration of the test is needed. It should be noted that other investigators may have selected a different way to operationalize this quality indicator.

A third step was to translate the measure or index of a quality to a 5-point scale. This was needed in order to identify the quality levels of each indicator and to place each indicator on a similar quality scale. The quality scale, in turn, could communicate to test developers where each

SQT stood in relation to its quality rating on each indicator. Table 3 shows an example of this translation for the decision-consistency indicator. In this table, the "Excellent" or "4" category reflects Subkoviack's (1988) rule of thumb for judging the goodness of the Kappa coefficient. An alternate possibility for making this translation from measure to rating scale is to obtain distributions of the measure (e.g., Kappa coefficient) and use the quintiles of these distributions as break points for defining interval boundaries. This was not done for this version of the 16 QI.

Figure 1 shows the current version of the 16 QI rating scale. To the right of each verbal statement of the quality indicator is a horizontal bar marked in segments numbered 0 through 4. These numbered segments represent the quality ratings for that indicator. Below each bar are numbers which represent the interval boundaries of the quantified measure of that quality indicator. For example, for Quality Indicator 2, Decision-consistency of total score, the numbers below the bar represent values of Kappa coefficient. Thus a value of Kappa greater than or equal to .60 is given a rating of 4, .40 to .59 a 3, and so on.

The boundaries shown in this version of the 16 QI were set rationally using judgment and any guidance provided by Army SQT policy and suggestions from the psychometric literature. Both the index used for each quality indicator and the boundary for translating to quality ratings should be subject to further validation research.

Who Completes the 16 QI Rating Form

Although it is possible for one person to complete the 16 QI rating form, this is not necessary and may be undesirable. Different parts of the rating form require different kinds of competence to complete. Some parts

of the 16 QI are based on statistical analyses which already exist in or can be appended to the ATSC item analysis program (Indicators 2, 3, 6, 7, 8, and 14). The other quality indicators require reviewing and judging the quality of various aspects of an SQT. Subject-matter experts would be needed to judge the item-task congruence, whether items measure MOS-specific knowledge, and whether the keyed answer is correct. Testing specialists could judge the quality of the item-writing. Perhaps a team of persons could review several SQTs.

Possible Diagnostic Value of the 16 QI

One of the potential uses of the 16 QI is to point to specific ways in which an SQT could be improved. Since each quality indicator is operationally defined, a low rating implies that a specific test development action is needed to raise the rating. For example, to continue with Quality Indicator 2, decision-consistency, a low value of Kappa could be obtained because the test was too short (thus, lowering KR20) or because the minimum passing score needs to be adjusted. Table 4 lists each of the 16 QIs and gives suggestions as to how to raise a low rating on it.

Formative Evaluation and Current Status of the 16 QI

Because the 16 QI has not been evaluated thoroughly, it has no official status in the U.S. Army. It is currently undergoing formative evaluation so it may be improved. Empirical studies are under way to ascertain the extent to which the statistical indices for Indicators 2, 3, 6, 7, 8, and 14 are functioning to distinguish SQTs of various quality. Preliminary results indicate that the speededness index used for QI Number 3 is not distinguishing among different SQTs, even those which appear to be somewhat speeded. Also

the decision-consistency indices (Kappa coefficient) for task test scores (subtest scores) are quite low probably because many of the task tests are comprised of 4 to 7 items. Given that an SQT must cover 15 to 20 tasks, it may not be reasonable to insist that these subtests be made longer or, it may require that the Army not use these subtests to make individual training decisions at the task level. Also, Indicator 13, related to the distribution of answer patterns, seems not to distinguish SQTs. Apparently almost all current SQTs do not have a fixed or set pattern of correct answer choice positions. This raises the question of whether to keep 13 as a QI, even though it reflects the current regulations. If it were withdrawn from a quality monitoring instrument such as the 16 QI, violations of this rule might creep into the testing program (as it had in years past).

Some civilian testing specialists who are reviewing SQTs and who are using the 16 QI are uncomfortable judging Indicators 4 (item-task congruence) and Indicator 10 (whether items measure MOS specific knowledge), believing that a subject-matter expert should judge these qualities. Other civilian testing specialists seem not to mind doing this judging. A problem that arises here has to do with the nature of the SQT development effort. Subject-matter experts are usually noncommissioned officers who are assigned the job of writing and reviewing test items as a temporary assignment. They are not trained for the job and are often transferred after a short while. Thus, they frequently have no motivation to carefully review a test item to assure it exactly matches the task or that it cannot be answered by common sense, general knowledge, or other non-MOS specific means.

Another problem arose in connection with Quality Indicator 1, the extent to which an SQT represents the domain of tasks written in a Soldier's

Manual (SM). A SM covers all essential aspects of an MOS job. Previous Army regulations required that an SQT sample the entire domain implied by the SQT, preferably through stratified random sampling. Recently the regulation was changed so that SQT are to reflect only those tasks from the MOS which are considered necessary to make a soldier battle-ready. That is, each SQT is to be a purposive sample of tasks (perhaps all tasks) that will give it a "battle focus." Thus, the current QI on domain coverage is no longer valid.

Other studies which should be done before making the 16 QI operational include reliability and validity investigations. For example, several persons should independently rate the same SQTs using the 16 QI and the same data-base. The consistency among ratings should be studied. Further, several SQTs should be rated wholistically (perhaps by a team) and ranked according to perceived quality. Then, these same SQTs should be rated using the 16 QI. The two sets of ratings may be correlated to see if the 16 QT has some degree of predictive validity.

Summary

The 16 QI is a set of quality standards for systematically evaluating criterion-referenced tests developed in a decentralized testing program. The specific application discussed in this paper is the U.S. Army SQT testing program. The 16 QI has potential for monitoring sQT quality in this program. If specific SQTs consistently receive high ratings, this would indicate that the development process is probably working well. Consistently low ratings would indicate a breadwon in the developmental process and would signal the need to target technical assistance to specific SQT development units.

An important use of the 16 QI is in diagnosing what needs to be done

to improve the quality of a criterion-referenced test. Each of the 16 scales diagnoses a particular flaw in a test. Each flaw can be corrected by specific test development actions which will raise SQT quality. Table 4 described the actions a test developer should take to remediate a low rating on each quality indicator. Further, because the 16 QI is an organized and systematic rating procedure, one may easily monitor whether the remedial action has been taken and the impact it has had on test quality.

Although the 16 QI is presented in the context of the U.S. Army's SQT program, it has practical utility in other contexts. Many criterion-referenced programs are organized similarly to SQTs: domains are defined, domains are sampled, tests are designed to measure each sampled objective, and decisions about mastery are made for each objective and for the domain as a whole. With only slight modification, the 16 QI could be used to evaluate such criterion-referenced tests in other branches of the military, in occupational testing programs, and in public schools.

Finally, from a systems analysis perspective, the 16 QI could help identify criterion-referenced test development practices which consistently yield quality tests. Test quality may be measured by the 16 QI. An analysis of the test development process at a particular site can identify specific procedures which can be correlated with test quality indicators. Those procedures which consistently distinguish better tests from poor ones can be fostered at other test development sites.

References

Berk, R. A. (1984). Selecting the index of reliability. In R. A. Berk (Ed.). A guide to criterion-referenced test construction. Baltimore: Johns Hopkins University Press, pp. 231-266.

Brittain, C. V. (1987). Minimum passing score (MPS) on skill qualification tests (SQTs). (Memo dated 141402 July).

Nitko, A. J. (1988). The Sixteen Quality Indicators: A Rating Form for Evaluating Skill Qualification Tests. (Final Report). Fort Eustis, VA: Clay V. Brittain, U.S. Army Training Support Center (Contract DAAL03-86-D-001, Delivery Order 0534, Scientific Services Program).

Subkoviak, M. J. (1984). Estimating the reliability of mastery-nonmastery classifications. In R. A. Berk (Ed.), A guide to criterion-referenced test construction. Baltimore: Johns Hopkins University Press, pp. 267-291.

Subkoviak, M. J. (1988). Skill qualification test (SQT) and common task test (CTT) development policy and procedures. (RCS ATTG-17(R1). TRADOC Reg 351-2. Fort Monroe, VA: Headquarters, U.S. Army Training and Doctrine Command.

Table 1. Organization of the critical SQT characteristics which need to be assessed.

A. TOTAL TEST CHARACTERISTICS

1. SQT tasks as a representative sample of the SM domain.
2. Decision-consistency of the total score
3. Sufficiency of testing time limits

B. TASK TEST CHARACTERISTICS

4. Congruence of items to task specifications
5. Inclusion of conditions of task performance on the test
6. Decision-consistency of task test scores
7. Length of task tests

C. ITEM CHARACTERISTICS

(a). Characteristics of items as functioning units

8. Easiness and difficulty of items
9. Performance-orientation of items
10. Items as measures of MOS-specific knowledge

(b). Characteristics of item stems

11. Freedom from flaws in phrasing the stem

(c). Characteristics of correct answers

12. Correctness of and freedom from ambiguity in the correct answer
13. Distribution of the correct answer position

(d). Characteristics of distractors

14. Plausibility of the distractors
15. Freedom from flaws in phrasing the distractors

(e). Other item characteristics

16. Freedom from other design flaws

Sixteen Quality Indicators for MOS Skill Qualification Tests

Evaluator_____ Date_____ SQT Test No _____

Quality Indicators — Ratings

1. Representativeness of SM domain

0	3	4
other	random	stratified

sampling plan used

2. Decision-consistency of total score

0	1	2	3	4
.60	.10	.20	.40	.60	1.00

Kappa coefficient

3. Sufficiency of testing time limits

0	1	2	3	4
1.0	.9	.2	.1	0

Speededness Index

4. Task-item congruence

0	1	2	3	4
1.00	10	5	1	0

Percent items not matching tasks

5. Conditions of task performance

0	1	2	3	4
+4	3	2	0	

Number of tasks missing conditions

6. Decision-consistency of task test scores

0	1	2	3	4
.60	.10	.20	.40	.60	1.00

Average Kappa for task tests in SQT

7. Length of task tests

0	1	2	3	4
0.0	3.0	4.0	5.0	6.0+

Average number of items per task test

8. Easiness and difficulty of items

0	1	2	3	4
100	15	10	5	3	0

Percent of items that are too easy or too difficult

9. Performance-orientation of the test items

0	1	2	3	4
0	90	93	95	97	100

Percent of performance-oriented items

10. Items measuring MOS-specific knowledge

0	1	2	3	4
100	5	2	1	0

Percent of items not requiring MOS-specific knowledge

11. Phrasing the stems of items

0	1	2	3	4
100	15	10	5	3	0

Percent of items having flaws in the stems

12 Keyed answer correct and free from ambiguity

0	1	2	3	4
+6	5	3	1	0

Number of items miskeyed or have ambiguous answers

13. Distribution of correct answer positions

3	4
Discernable (set)	Not discernable (set)

Pattern of correct answers

14. Plausibility of distractors

0	1	2	3	4
100	15	10	5	3	0

Percent of items with fewer than 1% of lower group choosing a distractor

15. Phrasing the distractors of items

0	1	2	3	4
100	15	10	5	3	0

Percent of items with flaws in distractors

16. Other design characteristics of items which are not rated above

0	1	2	3	4
100	15	10	5	3	0

Percent of items having other design flaws

Summary of Quality Indicator Ratings

I. Total test score characteristics: Average of 1, 2, and 3 = _____

II. Task test score characteristics: Average of 4, 5, 6, and 7 = _____

III. Item characteristics: Average of 8, 9, 10, 11, 12, 13, 14, 15, and 16 = _____

IV. Overall SQT rating: Average of 1 through 16 = _____

Table 2. Examples of regulations and policy statements that support the need to use decision-consistency of the total score as a quality indicator for an SQT.

Statement/doctrine	Reference
a. SQT results indicate MOS proficiency for training and personnel management decisions	Reg. 351-2, Par 2-2b
b. SQTs are standardized so that decisions are consistent from one place and time to the next	ATSC, Bulletin 86-1, pg. 5
c. Minimum passing scores are to be set carefully and fairly	Brittain (1987)
d. Task test standards are set to maximize decision conssitency	Reg. 351-2, Par F-12g

Table 3. Example of the translation of a measure of a quality indicator to a quality rating. (In this case, translating the estimated Kappa coefficient for an SQT to a quality rating on a 5-point scale.)

Numerical value of Kappa for the SQT total test score	Rating Assigned	Possible interpretation
0.60 - 1.00	4	Excellent
0.40 - 0.59	3	Good
0.20 - 0.39	2	Mediocre
0.10 - 0.19	1	Poor
0.00 - 0.09	0	Very Poor

Table 4. What to do to raise a low rating on each area of the 16 QI Rating Form.

Quality Indicator	How to remediate a low rating
1. Representativesness of SM domain	1. Create and use a stratified random sampling plan for selecting tasks for the SQT
2. Decision-consistency of total score	2. (a) Increase the number of questions on the SQT (b) Adjust the MPS
3. Sufficiency of testing time limits	3. (a) Increase the SQT's time limits (b) Reduce the number of questions on the SQT (c) Make the SQT items less complicated
4. Task-item congruence	4. (a) Review each item carefully to be sure it matches the SM, TM, or FM task specifications (b) Use the murder board review process more effectively
5. Conditions of task performance	5. (a) Review and analyze more carefully the task descriptions found in the SM, TM, or FM (b) Create "situation" statements that capture the important task conditions
6. Decision-consistency of task test scores	6. (a) Increase the number of questions on these task tests with low decision-consistency coefficients (b) Eliminate from task test items that are too hard, too easy, or too complicated (c) Adjust the "go/no go" score
7. Length of task tests	7. Increase the average number of questions per task test
8. Easiness and difficulty of of items	8. (a) Rewrite difficult items to eliminate ambiguity, unnecessary complexity, and item-writing flaws (b) Replace "give-away", common sense, and copying items with performance-oriented items

Table 4 (continued)

9.	Performance-orientation of items	9.	(a) Be sure items require an actual performance of tasks where possible (b) Eliminate items asking for definitions of terms (c) Be sure items focus on who, what, where, when, how often, etc.
10.	Items measuring MOS specific knowledge	10.	(a) Eliminate items testing general knowledge, common sense, copy skills, simple reading skills (b) Write items that only those who can perform well on an MOS can answer correctly (c) Increase the ratio of "key" performances tested relative to the "essential" performances tested
11.	Phrasing the stems	11.	(a) Use standard testing and measurement guidelines and checklists to review and revise the item stems (b) Be sure the item stem is focused on a single performance and asks a direct question
12.	Keyed answer correct	12.	(a) Check the answer key before submitting to ATSC (b) Make more effective use of the murder board reviewers by asking them to actually take the SQT without seeing the answer key (c) Use the ATSC Expanded Item Analysis Report to identify items exhibiting ambiguous answers, then revise these items before using them again
13.	Distribution of correct answer positions	13.	(a) Review the SQT answer key to be sure there is no set pattern of keyed answers (b) When writing each item, put the response choices in a logical order
14.	Plausibility of distractors	14.	(a) Use the ATSC expanded Analysis Report to identify items exhibiting this flaw before using the item again (b) Eliminate non-functioning distractors (c) Replace nonfunctioning distractors with distractors based on

Table 4 (continued)

		errors or misconceptions of who are known to be among the poorest performers of that MOS (d) Administer stems without distractors to MOS holders: Use their responses as a basis for writing distractors
15.	Phrasing the distractors of items	15. Use standard testing and measurement sources and checklists to review each distractor set and correct the flaws identified
16.	Other design characteristics of items	16. (a) Follow the suggestions found in Regulation 351-2 for writing items and using pictorial material (b) Ask the murder board to review the items in light of the item-writing suggestions found in Reg 351-2