

A Paper-and-Pencil Inventory for the Assessment of Piaget's Tasks

Henry O. Patterson and Louis Milakofsky
The Pennsylvania State University

Although science educators conversant with Piaget's work have recognized the importance of adapting instruction and curricula to the cognitive level of their students, such attempts have been difficult because of a lack of appropriate cognitive assessment instruments. To meet such a need, a comprehensive, objective paper-and-pencil inventory was investigated using 542 subjects, 8 years through adulthood, in order to determine its usefulness for normal and retarded students. The results showed that the inventory was acceptably reliable and valid and had advantages over other Piaget tests. With some suggested improvements, it was concluded that the instrument had potential as an educational and theoretical research tool.

The need for an easily administered standardized instrument to measure cognitive development in children and adults has been widely discussed in the educational and psychological literature. Although many Piagetians see attempts at psychometrizing the tasks as invalid and useless (e.g., Elkind, 1971; Herron, 1978), others have advocated the development of objective standardized instruments (e.g., Bart, 1978; Kaufman, 1971; Tuddenham, 1971).

Several possible research and applied uses of such a standardized instrument have been enumerated well by Tuddenham (1971), but al-

ready researchers in the area of science education are working to develop a streamlined assessment tool to use in the identification of cognitive stages of students. A growing number of chemists, physicists, and mathematicians are involved in this applied Piagetian research; but, unfortunately, some of these researchers—not well versed in test construction, standardization, or Piaget's theories and desiring a quick assessment instrument suitable for use with groups of students—have reacted to the dearth of suitable tests by devising their own instruments. The danger with some of these applied research studies by “naive” but well-intentioned science teachers, as pointed out by Herron (1976, 1977, 1978), is that curricula will be prematurely altered and students unfairly categorized on the basis of self-styled Piaget assessment devices having questionable reliability or validity.

Notwithstanding the theoretical discussions of the usefulness and validity of paper-and-pencil Piaget tests, and the necessity of their proper standardization, a literature review showed that some 17 paper-and-pencil tests have already been reported (Patterson & Milakofsky, 1978). Almost all of these tests were designed to measure formal thought; most require a high level of verbal and reading ability; and all can be administered in groups. Although most are based on tasks identified by Inhelder and Piaget (1958), different tests use vastly disparate num-

bers of items and scoring techniques. Most importantly, and with few exceptions (e.g., Lawson, 1977; Sheehan, 1970), very little has been published about both the reliability and validity of these tests.

If a conclusion can be drawn from the studies that have taken a psychometric approach to Piaget's tasks, it would be that it is possible to construct assessment instruments that (1) faithfully measure cognitive development and yield results similar to the classic Piaget individual interviews (with the possible exception of the timing for formal thought) and that (2) greatly facilitate Piagetian research. Of all the varied instruments reported, however, there appear to be some notable gaps. No test clearly meets all of the following criteria:

1. Comprehensively covers concrete *and* formal tasks;
2. Requires minimal language and reading skills;
3. Can be administered to a wide range of ages;
4. Can be quickly and easily administered to virtually any size group;
5. Can be quickly and objectively scored; and
6. Has been standardized and adequately studied for reliability and validity.

A published but little known instrument that seems to meet all but the last criterion does exist. An Inventory of Piaget's Developmental Tasks (IPDT), published by Furth (1970), is an experimental instrument to be used in the study of cognitive development with subjects 8 years of age and older. Although 10 years have passed since its publication, the existence of this inventory has not been widely reported; and despite limited studies by Ross (1968), Filer (1972), and Patterson (1975), there still are no major published reports on reliability or validity.

Because of the need for such a test, and because preliminary investigations with this inventory suggested that it had merit, a major research project was conducted to evaluate the reliability, validity, and usefulness of the IPDT.

The initial results of this project have recently been reported (Milakofsky & Patterson, 1977, 1979; Patterson & Milakofsky, 1978). The purpose of this paper is to expand on these previous reports by presenting additional data on the reliability, validity, and usefulness of the IPDT and by making suggestions for its improvement as an educational assessment instrument.

Method

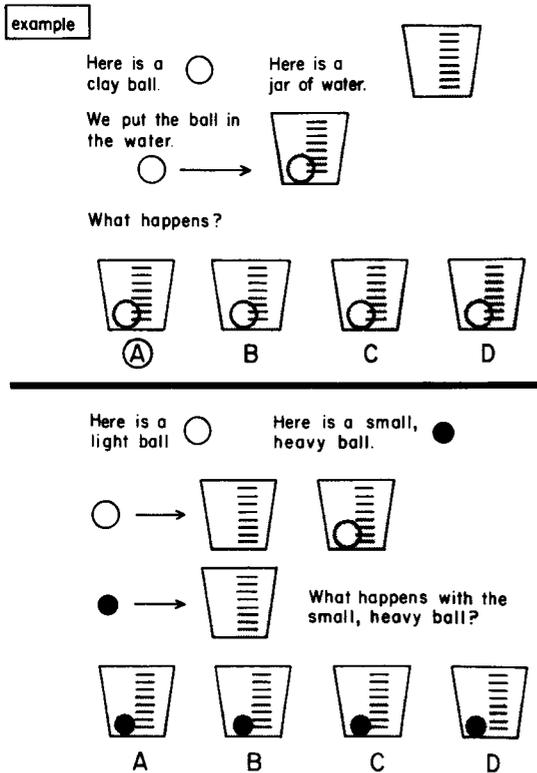
The IPDT

The IPDT is a 72-item untimed multiple-choice paper-and-pencil inventory designed originally as part of a project involving the testing of cognitive development in Navaho Indian children. A sample page showing an item in the Conservation of Volume Subtest is included as Figure 1. It was designed to translate some of Piaget's concrete and formal operational tasks into an objective, quick, standardized paper-and-pencil format requiring minimal reading ability. Presented in illustrations are 18 sets of tasks, divided into subtests of 4 items each. Each set is introduced by an example that is correctly answered in the test booklet. Five problem areas are covered in the test: conservation (4 subtests), images (4 subtests), relations (3 subtests), classification (4 subtests), and laws (3 subtests). The 18 subtests are identified in Table 1 with the approximate ages for their mastery according to the Piaget literature.

In this study, the IPDT was administered in four forms:

1. *Group Test.* Groups of students were given answer sheets with the IPDT test booklet and were read instructions on how to take the inventory. Subjects were allowed as much time as necessary to complete all items. Approximately 45 minutes were required for this administration.
2. *Group Test, Part 1.* The same procedure was followed as with the Group Test except that only half the test was given, i.e., subjects only answered Items 2 and 3 in each

Figure 1
Sample Page from the IPDT Showing the
Example and the Last Item on the
Conservation of Volume
Subtest (Courtesy Furth, 1970)



subtest. Items 1 and 4 in each subtest were covered in the booklet and on the answer sheet. This administration usually required 25 minutes.

3. *Group Test, Part 2.* The same as Group Test, Part 1, except that subjects answered only Items 1 and 4 in each subtest, and Items 2 and 3 were covered. Approximately 25 minutes was required.
4. *Individual Test.* To determine if the paper-and-pencil format obscured subjects' understanding of the concepts tested, the second and third items in each test (same items as in Group Test, Part 1) were chosen to present individually to subjects. Objects

similar or identical to those illustrated in the IPDT booklet were collected or constructed and spread out on tables. Subjects were brought in individually, shown the objects, and asked questions similar to those in the booklet. Testers were trained not to ask leading questions, but to make sure subjects understood what was being asked. Testers recorded responses on an answer sheet and wrote comments when appropriate. This administration required about 30 minutes per subject.

Subjects

A total of 542 subjects were tested in all during 1976 and 1977: 250 third-, sixth-, and ninth-grade students in a suburban school district; 210 freshmen and sophomores at The Berks Campus of Penn State University; 60 adults in college continuing education classes at Berks; 8 educably retarded children at a special education center; and 14 educably retarded adults at a sheltered workshop. Informed consent was obtained for all participants in the study.

For the major reliability-validity part of the study, 4 groups of approximately 60 students each were used in the design shown in Table 2. In the public schools, classes were selected by the respective school principals to participate in the study, and individuals were assigned to a group randomly. College students in psychology courses were randomly assigned to one of the three groups. Subjects not in the 12 groups used for the reliability-validity study were adults in selected continuing education psychology courses, additional college freshmen and sophomores in selected chemistry and logic courses, and additional sixth graders from a class in the same school as Groups 4 to 6. The retarded subjects were chosen on availability by the special education teacher in the center and an administrator in the sheltered workshop.

The following background information, where available, was obtained on subjects: age, sex, number of brothers and sisters, birth order, home neighborhood, father's occupation, latest

Table 1
Content Summary of IPDT

Subtest No.	Subtest Name	Problem Area	Concept Assessed	Approx. Mastery Age
1	Quantity	Conservation	Conservation of Quantity	7-8
2	Levels	Images	Transformational Imagery	9-10
3	Sequence	Relations	Ordinal Relations	7-8
4	Weight	Conservation	Conservation of Weight	9-10
5	Matrix	Classification	Classification	7
6	Symbols	Classification	Combinativity	7-8
7	Perspective	Images	Perspective	9-10
8	Movement	Images	Kinetic Imagery	8
9	Volume	Conservation	Conservation of Volume	11-12
10	Seriation	Relations	Ordinal Relations	7-8
11	Rotation	Laws	Kinetic Imagery	8-9
12	Angles	Laws	Reciprocal Implication	12
13	Shadows	Images	Perspective	9-10
14	Classes	Classification	Classification	12-13
15	Distance	Conservation	Conservation of Length	10
16	Inclusion	Classification	Verbal Class Inclusion	11-12
17	Inference	Relations	Verbal Transitivity	11-12
18	Probability	Laws	Probability	10-11

Note. From Milakofsky and Patterson (1979). Copyright 1979 by Journal of Chemical Education. Reprinted by permission.

IQ and achievement test scores (SAT scores for college students), and grade averages.

Testing Procedure

As Table 2 indicates, for the reliability-validity study there were three phases of testing with at least a month elapsing between each phase. In Phase 1 students in each group took tests in one of three orders: (1) one third took an Individual Test first, followed by the Group Test a few days later (Groups 1, 4, 7, and 10); (2) a second third had the Group Test first followed by the Individual

Test (Groups 2, 5, 8, and 11); and (3) a final third had only Group Test, Part 1 (Groups 3, 6, 9, and 12).

In the second phase, the first two groups in each grade took retests of the Group Test, and the third group took the Group Test Part 2. The third phase consisted of tape-recorded interviews with 30 randomly chosen students in Groups 3, 6, and 9; they were asked to explain their answers to selected questions on which their group had performed poorly. The omission of Group 12 interviews was simply due to lack of time before the subjects finished school.

Table 2
Reliability-Validity Research Design

Subjects and Groups	Phase I Testing		Phase II Testing	Phase III Testing
Grade 3(8-9 Years Old) $N=58$				
Group 1	Individual	Group	Retest Group	
Group 2	Group	Individual	Retest Group	
Group 3		Group Pt. 1	Group Pt. 2	Individual Interviews
Grade 6(11-12 Years Old) $N=57$				
Group 4	Individual	Group	Retest Group	
Group 5	Group	Individual	Retest Group	
Group 6		Group Pt. 1	Group Pt. 2	Individual Interviews
Grade 9 (14-15 Years Old) $N=62$				
Group 7	Individual	Group	Retest Group	
Group 8	Group	Individual	Retest Group	
Group 9		Group Pt. 1	Group Pt. 2	Individual Interviews
College (18-19 Years Old) $N=62$				
Group 10	Individual	Group	Retest Group	
Group 11	Group	Individual	Retest Group	
Group 12		Group Pt. 1	Group Pt. 2	

Note, From Milakofsky and Patterson (1979). Copyright 1979 by Journal of Chemical Education. Reprinted by permission.

The above counterbalanced design was used so that reliability and validity information could be assessed without contamination by prior exposure to the inventory. The testing sequence within each grade permitted uncontaminated group scores to be compared with uncontaminated individual scores, and test-retest correlations could be computed without administration order contamination. Only Groups 3, 6, and 9 were interviewed because all other groups had been contaminated by the individual test.

The subjects who were not part of the above design were administered the IPDT Group Test; the retarded subjects had to be given special individual administrations where items were read to them out of the booklet.

Results and Discussion

Reliability

Three measures of reliability were computed:

1. Test-retest reliability (coefficient of stability) was determined by correlating total IPDT Group Test scores for 8 groups on 2 administrations of the Group Test, 1 month apart. The results are shown in Table 3.
2. Internal consistency (split-half reliability) was assessed by correlating IPDT Group Test, Part 1, scores with Group Test, Part 2, scores (corrected by the Spearman-Brown Formula) to yield reliability coefficients for the test as a whole (Table 4).

Table 3
Pearson Product-Moment Correlation
Coefficients for IPDT Group Test-Retest Scores

Group	Grade								
	3		6		9		College		
	N	r	N	r	N	r	N	r	
1	20	.87**							
2	20	.75**							
4			20	.40					
5			19	.62*					
7					20	.84**			
8					20	.86**			
10							21	.95**	
11							21	.67**	

* $p < .01$; ** $p < .001$

3. Total IPDT Group Test and retest scores were compared by an analysis of variance. A significant phase effect was found ($F(1,120) = 6.242, p < .02$); but a *post hoc* analysis indicated that the only significant difference in test-retest scores was for Group 2 ($F(1,120) = 16.138, p < .001$). Figure 2 clearly shows the similarity of means for all groups except two.

These data on reliability correspond closely to reliability studies on other Piaget tests (Lawson, 1977; Sheehan, 1970). The lower test-retest correlations in Grade 6 can be explained and indeed are actually predicted by Piaget's theory. Students 11 to 12 years of age are in the transitional stage between concrete and formal opera-

Figure 2
Mean Scores of IPDT Group Tests and Retests 1 Month Later
(Courtesy Milakofsky & Patterson, 1979)

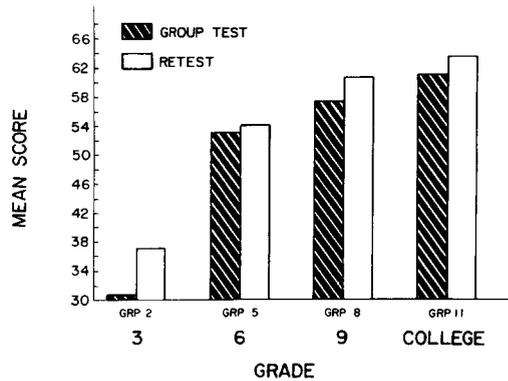


Table 4
Split-Half Reliability Coefficients for IPDT

Grade	Group	N	r^a
3	3	17	.71
6	6	18	.63
9	9	17	.84
College	12	20	.71

^aCorrected by Spearman-Brown formula.

tions, and their reasoning is not consistent. This is shown by the larger variances in scores for Grade 6 than for any other group (see Table 5).

The significant test-retest difference found in the scores of Group 2 is more difficult to explain. Since Phase 2 retest scores for Grades 1 and 2 were virtually identical, and since Group 1

did not change significantly from test to retest, the most likely explanation for the Group 2 change is that it is an artifact of the first Group Test administration. A review of that particular testing session revealed that there were indeed administration abnormalities that could have depressed the Group 2 scores.

Table 5
Mean Scores of Five Most Difficult Subtests
of IPDT by Grade

Subtest Name and No.	\bar{X}^a	SD
Grade 3(8-9 Years Old; $N=20$)		
Volume (9)	.60	.68
Perspective (7)	.75	.91
Classes (14)	.90	.72
Rotation (11)	1.00	.65
Shadows (13)	1.45	1.19
Probability (18)	1.45	1.15
TOTAL ^b	30.05	5.75
Grade 6(11-12 Years Old; $N=94^c$)		
Rotation (11)	1.70	1.21
Classes (14)	1.73	1.13
Probability (18)	1.91	1.22
Volume (9)	1.98	1.20
Distance (15)	2.18	.81
TOTAL ^b	47.31	10.22
Grade 9(14-15 Years Old; $N=20$)		
Classes (14)	1.95	1.39
Distance (15)	2.35	.81
Shadows (13)	2.55	1.10
Volume (9)	2.65	1.18
Probability (18)	2.80	1.24
TOTAL ^b	57.00	6.84
College (17+ Years Old; $N=226^c$)		
Classes (14)	2.27	1.37
Distance (15)	2.74	.76
Rotation (11)	2.96	1.07
Shadows (13)	3.06	.84
Volume (9)	3.25	1.03
TOTAL ^b	62.27	5.92

^aMaximum score = 4.

^bMaximum score = 72

^cIncludes subjects who were not part of the reliability study.

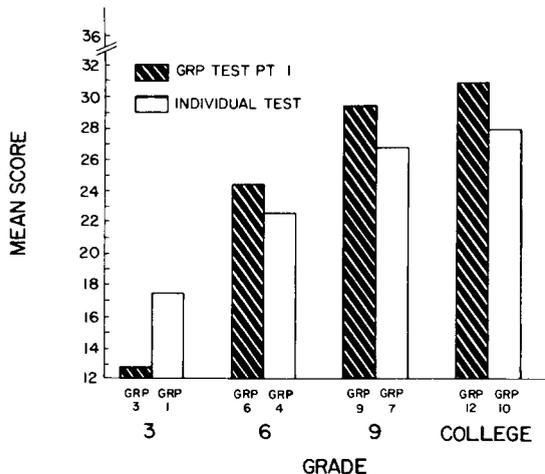
The general conclusions concerning reliability, therefore, are that the IPDT scores are reasonably stable over a short period for a wide range of age groups and that scores are not seriously affected by situational testing variables. Furthermore, the inventory has considerable internal consistency and appears to be measuring a general lasting trait rather than a specific temporary trait.

Validity

To evaluate concurrent validity, an analysis of variance was performed to determine if there were any significant differences between the Individual Test and the identical items in the Group Test. Figure 3 illustrates the finding that there was no significant difference in Group Test scores and Individual Test scores at any grade level ($F(1,148) = .848, p > .20$). Although the difference for Grade 3 is large, it was not a significant difference.

Construct validity was investigated by comparing mean scores for each group on the IPDT.

Figure 3
Mean Scores of Identical Items on Group Test, Part 1, and Individual Test for Each Grade
(Courtesy Milakofsky & Patterson, 1979)



Additional analyses of variance showed a significant grade effect in both Group ($F(3,120) = 70.36, p < .001$), and Individual ($F(3,152) = 72.518, p < .001$) IPDT scores; however, a post hoc analysis revealed no significant difference in group or individual scores between ninth grade and college ($F(1,120) = .83, p > .20$; and ($F(1,152) = 1.24, p > .20$). These results seem consistent with Piaget's stage theory.

The same pattern of increasing mean scores on two different kinds of administration of the IPDT shown in Figures 2 and 3 is also evident in a breakdown of the 18 subtests by problem area, as seen in Table 6. If the criterion of 75% or more correct responses on items in a subtest (i.e., at least three out of the four items) achieved by 75% or more of a group is used to define mastery of a subtest concept (such a criterion has been used by Elkind, 1962), the data show 15 out of 18 concepts (83%) were mastered by college students, 12 out of 18 (67%) by ninth graders, 5 out of 18 (28%) by sixth graders, and none by third graders. Mastery clearly increases with age; and the big increments between Grades 3, 6, and 9 are evident.

When IPDT group scores were compared to other standardized instruments, the results showed significant correlations with achievement tests and, with one exception, IQ tests (Table 7). For college students, the highest correlation with an achievement measure was with high school grade-point average, e.g., for Group 11, $r = .71, N = 21, p < .001$.

The final and least objective indicator of validity was the Phase 3 informal interviews with the public school students. Ten subjects from each grade (Groups 3, 6, and 9) were randomly selected for interviews. They were shown and asked to explain their answer to the most frequently missed items on the IPDT. The unequivocal finding of both investigators interviewing separately was that the students who missed items had not achieved the requisite cognitive structures; neither the paper-and-pencil multiple-choice format of the inventory nor the way any individual items were presented ap-

Table 6
 Percentage of Subjects by Grade Attaining
 at Least Three Out of Four Items Correct
 on Each Subtest

Subtest Number	Subtest Name	Problem Area	Percentage of Subjects			
			Grade 3 (N = 20)	Grade 6 (N = 94)	Grade 9 (N = 20)	College (N=226)
1	Quantity	Conservation	55.0	81.9	90.0	92.9
2	Levels	Images	70.0	89.4	85.0	96.0
3	Sequence	Relations	10.0	55.9	90.0	96.0
4	Weight	Conservation	30.0	86.1	85.0	98.7
5	Matrix	Classification	45.0	78.1	100.0	97.0
6	Symbols	Classification	35.0	74.5	100.0	96.9
7	Perspective	Images	5.0	54.2	80.0	95.6
8	Movement	Images	15.0	63.8	95.0	92.9
9	Volume	Conservation	0.0	35.1	55.0	77.5
10	Seriation	Relations	40.0	89.3	95.0	99.5
11	Rotation	Laws	0.0	24.5	65.0	69.9
12	Angles	Laws	10.0	54.3	80.0	84.1
13	Shadows	Images	20.0	64.9	65.0	81.8
14	Classes	Classification	5.0	23.4	35.0	46.0
15	Distance	Conservation	15.0	37.2	45.0	66.0
16	Inclusion	Classification	20.0	53.4	85.0	92.5
17	Inference	Relations	35.0	59.6	95.0	93.8
18	Probability	Laws	20.0	33.0	60.0	82.3

peared to significantly mislead the students into an incorrect answer.

The general conclusion concerning validity is that the IPDT basically shows the developmental progression of reasoning found by Piaget and other researchers in the five major areas included on the inventory and that it yields a result similar to the traditional individually administered tasks even though using a paper-and-pencil format. These indicators, closely corresponding to similar validity studies of other tests (Bart, 1972; Lawson, 1977; Renner, 1977), offer some evidence that the IPDT has concurrent and construct validity.

Other Results

A few other findings are noteworthy. Table 5 summarizes the results of scores obtained from

all students given the Group Test (excluding Groups 1, 4, 7, and 10) and lists the five most difficult subtests for each group. The data in Table 6 show that no group mastered Classes, Distance, or Rotations and that only college students mastered Conservation (volume), Images (shadows), and Probability—and these were barely mastered. These results clearly support recent studies showing that Piaget was incorrect about the timing of the attainment of formal thought (e.g., Elkind, 1962; McKinnon & Renner, 1971; Towler & Wheatley, 1971); a large number of adolescents (including college students) are not fully formal thinkers.

An analysis of scores among full-time college students (Table 8) revealed a slightly higher score for science majors than for non-science majors, but the difference was not statistically significant. The administration of the IPDT to

Table 7
Pearson Product-Moment Correlation
Coefficients for Group IPDT
Scores and Standardized Test Scores

Test Score	Grade 3 Group 2 (<i>N</i> =20)	Grade 6 Group 5 (<i>N</i> =19)	Grade 9 Group 8 (<i>N</i> =20)	College Group 11 (<i>N</i> =21)	Adults ≥ 22 yrs. (<i>N</i> =53)
Full Scale IQ ^a	.42	.49*	.68***		
Total SAT Composite Achievement ^b				.44*	.63***
	.58**	.87***	.46*		

^aLarge-Thorndike Intelligence Test.

^bIowa Tests of Basic Skills.

p*<.05; *p*<.01; ****p*<.001.

the two mentally retarded groups resulted in means of 21.5 for the 7- to 9-year-old group (*N* = 8) and means of 25 for the 19- to 45-year-old group (*N* = 14). Because of the small sample size, and because individual IQs and background information could not be obtained for these subjects, it is difficult to make any interpretations of the results. It would appear, however, that the IPDT could be of use in identifying particular strengths for mildly or moderately retarded adults.

Concerning correlates of scores with subject background information, the comparisons of total IPDT group scores with sex, number of siblings, birth order, and neighborhood showed no significant trends. There clearly is a relationship between age and score, but not within a grade. A correlation of $-.50$ (*N* = 53, *p* < .001) between IPDT group score and age for subjects over 22 years is consistent with other findings on the decline of formal thought through adulthood (Coleman, 1973).

Usefulness

The IPDT proved to be a quick, easily administered test that can be given to normal subjects 8 years and older in groups and to mentally retarded subjects on an individual basis. Since minimal reading proficiency is required, minor-

ity and culturally deprived students with reading problems would not appear to be seriously disadvantaged by the test, although this study does not deal with such samples. Most normal subjects from age 8 through adulthood immediately understood how to take the inventory, and most 8-year-olds had little difficulty reading the items. Most normal subjects required about 45 minutes to complete the inventory; the mentally retarded subjects generally needed more time. In addition, all ages seemed to find the inventory intrinsically interesting and challenging, and most found it enjoyable.

Several uses are foreseen. Piaget researchers could use this inventory to further study cognitive development over a wide age range and with different socioeconomic or minority groups. Educators could use the IPDT to aid in curriculum assessment, especially in determining readiness and subsequent ordering for certain science and math courses. In addition, educators who have developed Piaget-based courses and either need an instrument to identify students with reasoning difficulties or would like to assess pretest or posttest performance might find the IPDT useful.

Although there appears to be wide potential for this instrument, prospective users should be aware of the limitations to this inventory. As most Piagetians contend, a single score on a pa-

Table 8
Mean Subtest Scores for College Science-Engineering
and Non-Science Students

Subtest ^a		Science (N=89)		Non-Science (N=21)		F*
		\bar{X}	SD	\bar{X}	SD	
Number	Name					
1	Quantity	3.42	.64	3.29	.64	.707
2	Levels	3.89	.51	3.67	.97	2.157
3	Sequence	3.85	.49	4.00	.00	1.855
4	Weight	3.97	.24	3.86	.48	2.308
5	Matrix	3.89	.46	3.95	.22	.388
6	Symbols	3.70	.51	3.81	.51	.834
7	Perspective	3.82	.39	3.57	1.08	3.132
8	Movement	3.70	.61	3.62	.74	.252
9	Volume	3.36	.98	3.10	1.09	1.184
10	Seriation	3.89	.41	3.95	.22	.486
11	Rotation	3.08	.96	2.90	1.09	.532
12	Angles	3.62	.72	3.33	.91	2.411
13	Shadows	3.17	.64	2.95	.92	1.605
14	Classes	2.44	1.19	2.19	1.25	.726
15	Distance	2.78	.88	2.52	.75	1.474
16	Inclusion	3.71	.81	3.48	.75	1.414
17	Inference	3.82	.58	3.86	.36	.079
18	Probability	3.38	.98	3.19	.81	.685
TOTAL ^b		63.46	4.98	61.24	6.24	3.056

^aMaximum Score = 4.

^bMaximum Score = 72.

*p n.s. with df of 1,108

per-and-pencil inventory cannot yield the same kind of qualitative information that an individual interview can. The IPDT therefore would be best used in situations where individual interviews are impractical or impossible and where a quick gross measure of functioning is required. Even then, significant decisions on placement of a student certainly should not be made entirely on the strength of the IPDT score. Interpretation of scores should only be made in light of Piaget's theory; the IPDT is not designed to be used as an IQ test or an achievement test.

Users should also keep in mind that the IPDT is directed more at concrete thought than at formal thought. As would be expected, therefore,

for junior high age and older there is a ceiling effect on parts of the test, since 90% or more of ninth graders answered the same 35 items correctly and 90% of college students answered the same 40 items correctly. It appears that roughly half of the test is too easy for subjects 14 to 15 years and older; however, with recent evidence of regression of thought in old age (Coleman, 1973; Papalia, Kennedy, & Sheehan, 1973), the concrete operational emphasis of the test would make this inventory quite useful for life-span studies.

For use with normal junior and senior high school and college students, a new form of the inventory could be printed to omit the following

eight subtests: Quantity (Subtest 1), Levels (Subtest 2), Sequence (Subtest 3), Matrix (Subtest 5), Symbols (Subtest 6), Movement (Subtest 8), Seriation (Subtest 10), and Inferences (Subtest 17). Possibly Weight (Subtest 4), Inclusion (Subtest 16), and Perspective (Subtest 7) could also be eliminated. The resulting inventory, while half as long, should be as reliable and valid as the complete IPDT.

Conclusion

In light of the widespread search for standardized instruments to assess Piaget's tasks, the IPDT is clearly a promising theoretical and applied research instrument. The data presented in this study on reliability and validity certainly are a beginning at standardization of the instrument and provide the necessary initial background information for judicious use of the IPDT by researchers.

Much more research needs to be done with the instrument before it might be recommended for widespread applied use. Norms for each age level and various socioeconomic groups need to be determined, and additional reliability and validity studies using a more representative sample need to be done. A more comprehensive analysis of the items also might lead to a revision of the inventory beyond the suggestion of eliminating certain subtests for older subjects.

This study has only been the beginning of a long path of research to develop a highly effective instrument for the assessment of cognitive development, a needed tool in the empirical investigation of Piaget's theories and in the improvement of education.

References

- Bart, W. M. Construction and validation of formal reasoning instruments. *Psychological Reports*, 1972, 30, 663-670.
- Bart, W. M. Issues in measuring formal operational reasoning. *Genetic Epistemologist*, 1978, 8, 3-4.
- Coleman, S. B. The effect of aging on Piaget's developmental stages: A study of cognitive decline (Doctoral dissertation, Temple University, 1973). *Dissertation Abstracts International*, 1973, 34 (3A), 1122A-1123A. (University Microfilms No. 73-18, 712).
- Elkind, D. Quantity conceptions in college students. *Journal of Social Psychology*, 1962, 57, 459-465.
- Elkind, D. Two approaches to intelligence: Piagetian and psychometric. In D. R. Green, M. P. Ford, & G. B. Flamer (Eds.), *Measurement and Piaget*. New York: McGraw-Hill, 1971.
- Filer, A. A. *Piagetian cognitive development in normal and in emotionally disturbed children*. (Doctoral dissertation, University of Rochester, 1972). *Dissertation Abstracts International*, 1972, 33 (5-B), 2342.
- Furth, H. *An Inventory of Piaget's Developmental Tasks*. Washington, DC: Catholic University, Department of Psychology, Center for Research in Thinking and Language, 1970.
- Herron, J. D. *More Piaget for chemists: Things that I wish I had told you*. Paper presented at the 2YC3 Conference, Kansas City, October 1976.
- Herron, J. D. *Piaget applied: Suggestions for inaction*. Paper presented at the American Chemical Society Conference, New Orleans, March 1977.
- Herron, J. D. Piaget in the classroom: Guidelines for applications. *Journal of Chemical Education*, 1978, 55, 165-170.
- Inhelder, B., & Piaget, J. *The growth of logical thinking from childhood to adolescence*. New York: Basic Books, 1958.
- Kaufman, A. S. Piaget and Gesell: A psychometric analysis of tests built from their tasks. *Child Development*, 1971, 42, 1341-1360.
- Lawson, A. E. *The development and validation of a classroom test of formal reasoning*. Paper presented at the annual convention of the National Association for Research in Science Teaching, Cincinnati, March 1977.
- McKinnon, J. W., & Renner, J. W. Are colleges concerned with intellectual development? *American Journal of Physics*, 1971, 39, 1047-1052.
- Milakofsky, L., & Patterson, H. O. *The reliability and validity of an instrument to assess Piaget's tasks*. Paper presented at the meeting of the Division of Chemical Education, 174th American Chemical Society National Meeting, Chicago, August 1977.
- Milakofsky, L., & Patterson, H. O. Chemical education and Piaget: A new paper-pencil inventory to assess cognitive functioning. *Journal of Chemical Education*, 1979, 56, 87-90.
- Papalia, D. E., Kennedy, E., & Sheehan, N. Conservation of space in noninstitutionalized old people. *The Journal of Psychology*, 1973, 84, 75-79.

- Patterson, H. O. *Cognitive development in the manifest content of children's reported dreams*. Unpublished master's thesis, Bucknell University, 1975.
- Patterson, H. O., & Milakofsky, L. *The reliability and validity of a paper-pencil Piaget inventory*. Paper presented at the Eighth Symposium of the Jean Piaget Society, Philadelphia, May 1978.
- Renner, J. W. *Evaluating intellectual development using written responses to selected science problems* (NSF Report, Grant No. EPP 75-19596, Analysis of cognitive processes). University of Oklahoma, Norman, OK, 1977.
- Ross, B. M. *Report on An Inventory of Piaget's Developmental Tasks in testing Navaho Indian children*. Unpublished manuscript, Catholic University, 1968.
- Sheehan, D. J. The effectiveness of concrete and formal instructional procedures with concrete and formal-operational students (Doctoral dissertation, State University of New York at Albany, 1970). *Dissertation Abstracts International*, 1970, 31 (6A), 2748A. (University Microfilms No. 70-25, 479).
- Towler, J. O., & Wheatley, G. Conservation concepts in college students: A replication and critique. *Journal of Genetic Psychology*, 1971, 118, 265-270.
- Tuddenham, R. D. Theoretical regularities and individual idiosyncrasies. In D. R. Green, M. P. Ford, & G. B. Flamer (Eds.), *Measurement and Piaget*. New York: McGraw-Hill, 1971.

Acknowledgments

Parts of this study were funded under a grant from The Pennsylvania State University Commonwealth Campus Scholarly Activity Fund (Phase III), The Berks Campus Scholarly Activity Fund, and the College of Science and College of Liberal Arts Computer Funds.

The authors gratefully acknowledge the generous cooperation and assistance of the following: Stanley T. Dubelle, Christ J. Bucolo, Stephen Gancar, and Calvin Bossler of the Governor Mifflin School District; Robert E. Brown, Berks Campus; Scott Schaefer, Cathy Plushanski, William Hockley, James Henry, Joel Cleveland, Robin Juckem, Thomas Gerhard, and Nancy Miller, student assistants; and Nancy Patterson, Barbara Stutzman, Sandy Kreis, Kathy Lora, and Judy Donahue.

We also thank Carol H. Ammons, University of Montana, for her review of an earlier draft of this manuscript. An earlier version of this paper was presented to the Division of Educational Psychology at the 86th annual convention of the American Psychological Association, Toronto, Ontario, Canada, August 28, 1978.

Authors' Address

Send requests for reprints or copies of the authors' other papers on this topic to Henry O. Patterson and Louis Milakofsky, The Pennsylvania State University, Berks Campus, R.D. #5, Tulpehocken Road, P.O. Box 2150, Reading, PA 19608.