On the Test-Retest Reliability of the Scale of the Difficulty of Numbers Grasped as Connotative Meaning Experienced by Fourth and Sixth Grade Pupils

by
Saburoh MINATO
The Faculty of Education, Akita University

Collaborators:
Hisao Yatsuyanagi, Noriko Yatsuyanagi,
Takao Saitoh

INTRODUCTION

In this short report, the author intend to examine the test-retest reliability of the scale of the difficulty of numbers, grasped as connotative meaning, for fourth and sixth graders of elementary schools.

The author and his collaborators introduced a psychological concept: difficulty of numbers (or simply, DN) and specified the scaling procedures (1). The scales obtained were called the difficulty of numbers grasped as connotative meaning (or simply, DNCM).

In the previous report (2) , we calculated the reliabilities of the scale obtained for college students. In spite of the fact that the relation between the difficulty of multiplication combinations and DNCM has been studied, the reliability of the scale of DNCM for elementary school children has not been examined and reported.

So far there were two types of the scaling procedures of DNCM, the scales obtained being named φ and ψ . It seems that the correlation coefficients between the scale φ and the scale ψ were 0.96 or more (1),(2),(3) and the scale ψ was not applicable to the calculation of item reliabilities (the meaning of item reliability was in a previous report (2) and will be cited

after.). In this paper, therefore, only the scale φ is treated.

PROCEDURE

Instrument

For each number 0,1,...,9, a bipolar adjectives easy - difficult was combined. Each subject judged the degree of his feeling or impression of easy or difficult of number a on a given five point continuum: easy ___:___:___:___:___difficult, which, it would be expected, was more workable for elementary school children than seven point continuum used in standard SD (4), or the previous study with the subjects of college students.

Following the judgment of subject j on a five point continuum, number 1,2,..,5 were alloted, where number 1 was indicated easiest, and number 5 extremely difficult. We obtained in this procedure the value $e_j(\underline{a})$ which denoted the difficulty of number a grasped as connotative meaning, experienced by the subject j. We calculated $\varphi(\underline{a})$ which was the group mean score of $e_j(\underline{a})$ over a group of the subjects. (note: each $e_j(\underline{a})$ and $\varphi(\underline{a})$ used in the previous study (2) were a score by seven point continuum.)

Administration of the instrument

The test and retest (after a week) of the above instrument were administered during June and July, 1978. One hundred and ninety two subjects who attended to three elementary schools denoted as N, I and T , located at Akita Prefecture, Japan were involved in the study. They were fourth and sixth graders. The distribution of the subjects by grade and school is shown in Table 1.

Calculation of the test-retest reliabilities

In this paper, Pearson product-moment correlation coefficients are calculated for examining the test-retest relia-

TABLE 1. DISTRIBUTION OF SUBJECTS BY GRADE AND SCHOOL

grade	school		total
4	N	73	105
	I	32	
6	T	59	87
	I	28	
total	192		192

bilities. Then two kinds of test-retest reliabilities are calculated, which is as follows:

(1) The test-retest reliability of the group mean scores

In the calculation, the group mean score $\bar{y}(\underline{a})$ ($\underline{a} = 0,1,..,9$) is a unit of the study.

(2) The test-retest item reliabilities

In the calculation, a person's score $e_j(\underline{a})$ judged by each subjects j is a unit of the study (2).

RESULTS

(1) The test-retest reliability of the group mean scores is $r = 0.997^{**}$ (** significant at 0.01 level, through the paper)

(2) The test-retest item reliability on each number is as follows:

TABLE 2. TEST-RETEST ITEM RELIABILITIES

number	0	1	2	3	4	5	6	7	8	9
r	.715**	.725**	.371**	.670**	.542**	.488**	.646**	.776**	.731**	.743**

(3) The test-retest reliability of the group mean score for the fourth graders is 0.988^{**} and for the sixth graders is 0.998^{**}.

(4) The test-retest item reliabilities on each number for the fourth and the sixth graders are as follows

TABLE 3. TEST-RETEST ITEM RELIABILITIES FOR THE FOURTH AND THE SIXTH GRADERS

number	0	1	2	3	4	5	6	7	8	9
4th graders	.675**	.564**	.350**	.412**	.467**	.462**	.539**	.674**	.692**	.684**
6th graders	.733**	.850**	.417**	.805**	.548**	.507**	.652**	.761**	.687**	.686**

ADDITIONAL RESULTS

In this section, additional results are stated.

(1) The correlation between the scale with the subjects of fifth and sixth graders (3) (In that study, φ denoted the scale.) and the scale of this study is $r = 0.931^{**}$

(2) The correlation between the scale with subjects of college students (2) (In that study, φ_4 denoted the scale.) and the scale of this study is $r = 0.440$.

(3) The group mean scores of the subjects, the group of the fourth graders and the group of the sixth graders are as follows:

TABLE 4. GROUP MEAN SCORES

number	0	1	2	3	4	5	6	7	8	9
the Ss	1.41	1.22	1.29	1.51	1.55	1.38	1.83	2.29	2.16	2.36
4th graders	1.29	1.19	1.31	1.31	1.37	1.32	1.53	1.81	1.77	1.87
6th graders	1.55	1.26	1.26	1.74	1.76	1.45	2.20	2.87	2.62	2.95

CONCLUSIONS AND IMPLICATIONS

The test-retest reliability of group mean scores is very high and almost as same as the reliability of the previous study (2) involving subjects of college students. This result is secured by those of the fourth and the sixth graders.

It is interesting to note that the correlation between the scale with the subjects of the fifth or the sixth graders of

the previous study (3) and the scale of this study is high, but the correlation between the scale with the subjects of college students and the scale of this study is not so. Further study is needed for clarifying the reason.

The test-retest item reliability coefficients of the subjects are distributed from 0.371 to 0.776. These values seem to be not less than those of the previous study (2) with subjects of college students.

As long as the group mean scores of DNCM is used for any study, for example, the study of ours (3), the results of the study are expected to be sufficiently reliable.

REFERENCES

(1) Minato,S., Yatsuyanagi,H. and Saitoh,T.,(1977); On the Scaling of the Difficulty in Numbers 1,2,...,9, Grasped as Connotative Meaning, Reports of Mathematical Education (Japanese, English Summary), 32, pp.1 - 17.

(2) Minato,S.,(1978); A Report on the Reliability of Scales of the Difficulty of Numbers Grasped as Connotative Meaning (English, Japanese Summary), Bulletin of Japanese Curriculum Research and Development, 3, pp.101 - 107.

(3) Minato,S., Saitoh,T. and Yatsuyanagi,H. ; Relation between the Difficulty of Numbers Grasped as Connotative Meaning and the Relative Difficulty of Multiplication Combinations, (in preparation).

(4) Osgood,C.E., Suci,G.J. and Tannenbaum,P.H.,(1957); The Measurement of Meaning, Urbana, University of Illinois Press.

小学4，6年生の感じる内包的意味における数の困難度の尺度の再テスト信頼性について

湊　三郎（秋田大学）

研究協力者：八柳久夫，八柳敦子，斎藤芳雄

この報告は，内包的意味における数の困難度の信頼性を，小学校4，6年生が被験者の場合についてもめることを目的としている。ここでは，ふみで尺度構成が論じられ，使用された尺度水準までは扱わない。したがって測定用具はSDの変種と考えられる，やさしい－むずかしいを二極形容詞対とするような尺度であり，やさしい側から1，2，…，5が与えられて数量化される。

検査は1975年5月から6月にかけて，1週間をおいて2度実施され，比較された。被験者は秋田市とその周辺の小学校3校で，被験者数は172名である。このうち4年生は105名，6年生は67名であった。

再テスト信頼性係数は群平均値に関するもの－即ち尺度Φの信頼性と，数a（a＝0，1，…，9）における被験者jの尺度値 $e_j(a)$ をこの検査についてもめ，172個の対を作ったときの信頼性係数－項（item）の信頼性をもとめた。

群平均値に関する再テスト信頼性係数は全被験者については 0.997，4年生では 0.988，6年生では 0.998 であった。各数毎の項の信頼性係数はちらばっており，全被験者については 0.371（数2）から 0.776（数7），4年生では 0.350（数2）から 0.692（数8），6年生では 0.417（数2）から 0.761（数7）であった。

過去の検査において得られている資料と今回得られた資料を対にして群平均値に関する信頼性をもめると，秋田県内小学校5，6年生との相関は 0.931，大学生についての一つの検査との比較では相関は 0.440 であった。

以上のことから，内包的意味における数の困難度を群平均値として一括して扱う場合（｛φ(0)，φ(1)，…，φ(9)｝）には高い信頼性をもって研究が行われ得ることがわかった。

（秋田数学教育会会報　no.10）