First publ. in: Studies in Educational Evaluation 1 (1975), 2, pp. 65-76

EVALUATION REFLECTIONS

SUBJECT-MATTER DIRECTED MOTIVATION AND ITS EVALUATION BY MEANS OF QUESTIONNAIRES

WILHELM F. KEMPF and MANFRED LEHRKE
Institute for Science Education (IPN) at the University of Kiel

1. INTRODUCTION

Learning ought to be motivated by its subject matter and by the material to be learned, not by pressure, force, or reward from the teacher or parents. Institutionalised learning differs from the acquisition of information outside of school: in the former, learning is demanded more or less independently of the needs and wishes of the individuals concerned, whereas the latter is "naturally" motivated. However, if the student has no motivation directed at dealing with the subject matter of instruction, his actions will usually be guided by the anticipation of punishment. He is not motivated to deal with the material but to avoid punishment. The subject matter thus becomes merely a means to that end.

Teaching situations which are characterized by the absence of subject-matter directed motivation necessarily entail a number of undesirable consequences. This statement applies to both the results of cognitive learning and some educational goals of a higher order. For instance, when learning behavior occurs only under outside pressure and is maintained only by it, one cannot count upon the development of interests which extend beyond the particular class or the school years. One cannot even expect an open mind with respect to the areas covered at school.

Subject-matter directed motivation (smd-motivation) has been defined as a turning towards the subject-matter, a readiness to occupy oneself also voluntarily with that material, to acquire further information, and to digest it (cf. Lehrke and Kempf, 1975a). Hence, subject-matter directed motivation is opposed to any motivation of learning which is alien to and independent of the subject-matter and thus at any rate extrinsic. Lehrke and Kempf (1975b) also distinguish subject-matter directed motivation from such concepts as achievement motivation and intrisic motivation in that these are not limited to observed or observable modes of behavior but are used for their classification on the basis of their genesis:

- In order to be classified as achievement motivated, behavior must be determined by competition using a standard of exellence. (Mc Clelland, Atkinson, Clark, and Lowell, 1953; cf. Heckhausen, 1965). Such an evaluation is not implied for smd-motivated behavior.
- One speaks of intrinsic motivation only if a form of behavior does not occur as a result of its possible consequences, but because the activity itself is experienced to be pleasant. Subject-matter directed motivation is a more general concept.

Apart from the difficulty of isolating intrinsic motivation in a "pure" form, it seems neither realistic nor desirable to want to motivate students only intrinsically, namely,

through the learning activity independent of its consequences. Extrinsic sources of motivation differ quantitatively, in intensity, and according to their importance for the learning process. By no means do they always inhibit learning.

One cannot immediately determine what initiated some behavior which is considered to be governed by subject-matter directed motivation. Moreover, one need not assume that the same behavior came about in the same way in different students which were studied. It becomes a question of educational design to determine through what kind of instruction one may succeed in changing students' approach to a particular subject-matter.

For the evaluation of these aspects of curriculum material, a measuring instrument is required which can be easily used, which must not interfere with teaching, and which must be repetitively usable within the same class. In addition, it must be applicable to various educational levels, school subjects, contents, methods, and teachers. In line with the above stated criteria, not long-term but evoked subject-matter directed motivation is to be measured.

2. THEORETICAL FOUNDATIONS

The above definition of subject-matter directed motivation comprises a multitude of modes of behavior. Each of these modes of behavior is considered to be an indicator for subject-matter directed motivation: In the same way as a student who continues voluntarily (e.g., outside of school) to occupy himself with the subject matter, whoever shows interest in the content of what is being taught in a class, whoever enjoys it, may be considered smd-motivated.

By combining such different modes of behavior within the single common concept of smd-motivation it is implicitly assumed that students can be ranked according to the degree to which they are motivated and that, taking two students, for example, the one who enjoys the content of the course is also the one who more strongly tends to pursue it on his own outside of class. The problems involved are the same as those that arise in the use of psychological trait concepts (cf. Kempf 1974, 1976). The introduction and application of the concept of subject-matter directed motivation presuppose the presence of an *empirical* lawfulness: statements about subject-matter directed motivation which are derived from some mode of behavior must be transferable to other modes of behavior. If *more than one* mode of behavior is considered to be an indicator of smd-motivation, and if, of any two students v and v', the one (v) who demonstrates that mode of behavior (i) *with greater probability* ($P_{vi} > P_{v',i}$) is considered to be more strongly motivated, then all those probabilities of behavior (P_{vi}; i = 1, 2, ...) must be transferable by means of strictly increasing transformations. If a more powerful scale of motivation (e.g. an interval- or rational-scale) is desired, higher order relations among the students must also be defined. For instance, one might postulate some specific assumptions concerning the structural form of the functional relationship between the above-mentioned probabilities of behavior (P_{vi}) and the subject-matter directed motivation of the students. Today, it is possible to demonstrate that such assumption must satisfy certain requirements of philosophy of science and of mathematical statistics. Thus, the number of possible functions is reduced to a rather small one, such that it can still be easily grasped.

One could, of course, dispense with further specifications concerning this functional relationship and estimate the students' rank order by means of the number of relevant modes of behavior demonstrated. As Kempf (1976) shows, this method leads to an unbiased estimate of the rank order, but it involves an uncontrollable loss of information when this number is not a sufficient statistic for the student's motivation. However, the requirement of sufficient statistics of the form

(1) a_{vo} = number of relevant modes of behavior demonstrated by student v

necessarily leads to very specific assumptions concerning the functional relationship between the student's motivation and his or her probability of demonstrating modes of behavior indicating motivation. When the different modes of behavior ($i = 1, 2, \ldots$) are "locally independent" (so that all correlations between them can be explained by differences in motivation among students), this functional relationship has the form of the Rasch model (1960)

$$(2) \qquad P_{vi} = e^{\xi_v - \sigma_i} / (1 + e^{\xi_v - \sigma_i}), \quad \begin{matrix} i = 1, 2, \ldots \\ v = 1, 2, \ldots \end{matrix}$$

It expresses the probabilities P_{vi} as a logistic function of the students' motivation (ξ_v) and the difficulties of the modes of behavior (ε_i). Thus, it is not even possible to rank the students with respect to their subject-matter directed motivation without making reference to a specific test model. Otherwise, it may happen that students are considered as equally motivated although a more exhaustive analysis *of the same data* could give empirical evidence that this is not the case.

An inherent characteristic of the Rasch model is that it allows for *specific objectivity* in comparisons of individuals and items: the comparison of any two subjects can be carried out in such a way that no parameters are involved other than those of the two subjects - neither the parameters of any other subjects nor any of the item parameters. In particular, the parameters of any two subjects (or of the same subject at different points in time) can be compared on their own virtues alone, irrespective of the group or population to which for some reasons they may be referred. This opens up the possibility of measuring a student's subject-matter directed motivation with respect to different course contents and obtaining comparable results. Similarly, the difficulties of the items (i. e., of the various modes of behavior considered as indicators of subject-matter directed motivation) can be measured quite irrespective of the sample of examinees.

Specific objectivity can be accomplished only, however, if the empirical data conform with the model, i. e., if all items of a test actually do measure the same trait. The testing of this aspect of "homogeneity" of a test thus becomes the crucial point in item analysis.

The conformity of the data with the model can be tested both in form of a likelihood ratio test (cf. Andersen 1973) which provides an asymptotically χ^2-distributed test statistic and in form of a graphical plot. As stated above, the Rasch model allows for comparisons of item difficulties which are independent of the examinees involved. This opens up the possibility of estimating the item difficulty parameters from systematically chosen sub-samples of individuals and comparing those estimates. If all items measure the same latent variable, then the estimates must differ only slightly

among themselves and from the overall estimates. In the graphical control of the model the restricted estimates are plotted against the overall estimates. If the model holds, the thus plotted points will lie approximately on a straight line which goes through the origin of the system of coordinates and has a slope of 45^0.

The statistical goodness of fit test for the model follows the same rationale – comparison of restricted and over all estimates of item difficulties. It is preferable to the graphical test because the latter does not offer any strict criterion as to the tolerable degree of deviation of an item. In the present study, therefore the graphical test will be only used for the purpose of illustration.

3. THE QUESTIONNAIRE

In current evaluation projects at the Institute for Science Education (IPN), an attempt was made to measure subject-matter directed motivation by means of questionnaires. To this end an item pool of various statements was compiled.[1] The items deal with modes of behavior like "showing interest in course content", "occupying oneself with the course content on one's own", etc. The statements may or may not apply to the individual student and the students are asked to check "agree" or "disagree" depending on what applies to them.

Some of the items are stated "positive", i. e. agreement is considered to indicate subject-matter directed motivation. Other items are formulated "negative". All items refer to the same course content and can be brought into relation with specific sections of the course. The content does not appear in the individual items, however, but the relationship is established through the instructions. Thus, the questionnaire can be applied to different course contents, and the obtained results can be compared.

Of course, such a comparison, is possible only if all items actually do measure the same latent variable – namely the students' subject-matter directed motivation. Moreover, it must be demonstrated that the difficulty of the items does not change when the questionnaire is applied to different course contents and teaching techniques. Otherwise, differences in test scores cannot be traced back to differences in motivation.

In a study by Lehrke and Kempf (1975b), 22 items were grouped into four partially overlapping sub-tests. These were administered to 356 9th grade students in physics classes. The first three questionnaires were carried out during a two-month course on the concepts of models, the fourth after almost two months of teaching in small groups on aspects of electronics. Various sub-areas within the course-work pertaining to the concept of models were covered in different sequences and with varying details in different classes. During subsequent instruction, different groups dealt with completely different issues: Experiments with transistor circuits, logical circuits, voltage amplifiers, purchase of phono and TV appliances.

At the beginning of each new instructional section a questionnaire concerning the previous section was administered. All students received the sub-tests in the same order. Thus, each of the sub-tests was applied to learning experiences which differed at least with respect to length, content, and teacher.

Item analysis with the Rasch model showed that not all items met the requirements to the same extent: Two items had to be eliminated because they were ambiguously

[1] The authors wish to thank Miss Verena Schmid for her contributions.

formulated. Two further items "*Bestimmte Dinge sind meinem Interesse nach zu kurz behandelt worden*"[2] and "*Es gab keine Fragen mit denen ich meinem Interesse nach zu lange beschäftigen mußte*"[3] concerned a weighted relationship between the students' motivation and the actual information presented in the course and thus showed a dependence of item difficulty on the variability of the teaching.

The answers to three items. "*Ich habe weder in Lexika noch woanders nachgeschlagen, um mehr Informationen über das behandelte Gebiet zu bekommen*". "*Wenn ich jetzt etwas über die Dinge, die wir im Unterricht behandelt haben, in der Zeitung finden würde, würde ich es wahrscheinlich nicht lesen*" and "*Wenn mir jemand ein Buch über die zuletzt behandelten Themen schenken würde, würde ich mich darüber nicht besonders freuen*" were codetermined by certain marginal conditions of the behavior to which they referred, namely the "reading habits" of the students.

These results point out that the concept of subject-matter directed motivation can only be applied under certain restrictions. Since it covers and emphasizes only one aspect of the "motivated" behavior and since this behavior can be codetermined by other conditions as well, it must be assured that inference is based only on such modes of behavior for which these marginal conditions are relatively stable within the group of students under consideration.

As the results by Lehrke and Kempf (1975b) suggest, the course of instruction itself can effect such marginal conditions. After group teaching, during which the students acquired a large part of the learning material from the literature by themselves, "reading habits" no longer appeared as an interfering variable.

4. CROSS VALIDITY OF THE RESULTS

Based on these findings, Lehrke and Kempf (1975a) called in question whether it is possible at all to construct *universally applicable* tests of subject-matter directed motivation. The authors suggested, however, that one may attempt to produce item pools for which the marginal conditions remain sufficiently stable within well defined student populations and under given teaching/learning conditions. In order to test whether the item pool compiled provides an useful basis for such an effort, cross validation of the results of the item analyses is required.

To this end a further study was carried out. After omitting those items which were either ambiguously formulated or showed a dependence of item difficulty on the variability of the teaching, and after including two additional items (statements 8 and 17 in Table 1) an item pool of 20 statements resulted. These are summarized in Table 1. Nine of the items stated are "positive" (statements 1, 2, 4, 9, 11, 14, 17, 19 and 20), the remaining items are formulated as "negative" (statements 3, 5, 6, 7, 8, 10, 12, 13, 15, 16 and 18). Since mere translation cannot guarantee equivalents of the original items, the statements are also cited in German. Transfer of the results to other cultures will require further research.

[2] Certain items were dealt with too briefly for me.

[3] There were no questions that I had to deal with longer than I was interested.

Table 1 The Item Pool

1 Es gab Fragen bzw. Themen, die mich besonders interessiert haben.	1. There were some questions or topics that were of particular interest to me.
2. Ich habe auch außerhalb des Unterrichts über manche Dinge nachgedacht, die wir zuletzt gelernt haben.	2. I have been thinking about some of the things that we have learned recently even ontside of class.
3. Ich habe weder in Lexika noch woanders nachgeschlagen, um mehr Informationen über das behandelte Gebiet zu bekommen.	3. I did not resort to either encyclopedias or other reference books in order to obtain more information about the subject covered.
4. Ich habe mich hauptsächlich deshalb am Unterricht beteiligt, weil mir das Thema Spaß gemacht hat.	4. I participated in class primarily because I enjoyed the subject.
5. Ich möchte mich in den nächsten Stunden lieber nicht mehr mit ähnlichen Fragen beschäftigen.	5. In the next few periods I would prefer not to deal with these kinds of questions any more.
6. Während des Unterrichts habe ich nicht gemerkt, daß es Spaß machen kann, so etwas zu lernen.	6. I did not notice during class that it can be fun to learn something like that.
7. Wenn ich jetzt etwas über die Dinge, die wir im Unterricht behandelt haben, in der Zeitung finden würde, würde ich es wahrscheinlich nicht lesen.	7. If I found something in the paper about the material that we have been dealing with in class, I probably would not read it.
8. Manchmal fand ich es schwierig, mich auf die Sache zu konzentrieren.	8. Sometimes it was difficult for me to concentrate on the topic.
9. Ich habe das Gefühl, für mich selbst etwas dazugelernt zu haben.	9. I have the feeling of having learned something new for myself.
10. Noch mehr möchte ich nicht über dieses Thema erfahren	10. I do not want to learn anything else about this topic.
11. Die Schule würde mir mehr Spaß machen, wenn wir öfter solche Dinge behandeln würden.	11. School would be more fun if we dealt with more subjects like this one.
12. Wenn mir jemand ein Buch über die zuletzt behandelten Themen schenken würde, würde ich mich darüber nicht besonders freuen.	12. If, as a gift, someone were to give me a book on the topics that we just covered, I would not be particularly thrilled about it.
13. Ich hätte mich nie mit diesem Gebiet beschäftigt, wenn ich nicht gemußt hätte.	13. I would never have occupied myself with this subject if I had not been forced to do so.
14. Ich wünschte, es gäbe bald eine Fernsehsendung über dieses Thema.	14. I wished there would be a program on TV soon about this subject.
15. Einiges hat mich ausgesprochen gelangweilt.	15. Some things absolutely bored me.
16. Mein Interesse an Naturwissenschaften ist nicht größer geworden, seit wir diesen Stoff durchgenommen haben.	16. My interest in natural science has not increased since we covered this material.
17 Es hat Spaß gemacht, mein Verständnis für dieses Thema zu vertiefen.	17. It was fun to deepen my understanding of this subject.
18. So interessant war es nie, daß ich es schade fand, als es klingelte und die Stunde vorbei war.	18. It was never so interesting that I regretted it when the bell rang and the period was over.
19. Ich habe mit Freunden, Eltern oder Geschwistern über Dinge aus diesem Gebiet gesprochen.	19. I discussed some of the things from this subject with my friends, parents, or brothers and sisters.

20. Mit solchen Themen hätte ich mich auch freiwillig gerne beschäftigt, z. B. in einer Arbeitsgruppe ohne Zensuren oder zu Hause.

20. I would have gladly occupied myself with these kinds of topics on my own, for example, in a work-group without grades or at home.

The questionnaire was administered to 228 10th grade students after an introductory course on kybernetic models and temperature regulation. Since possible effects of response sets cannot be excluded (cf. Lehrke & Kempf 1975b), "negative" and "positive" items were analysed separately.

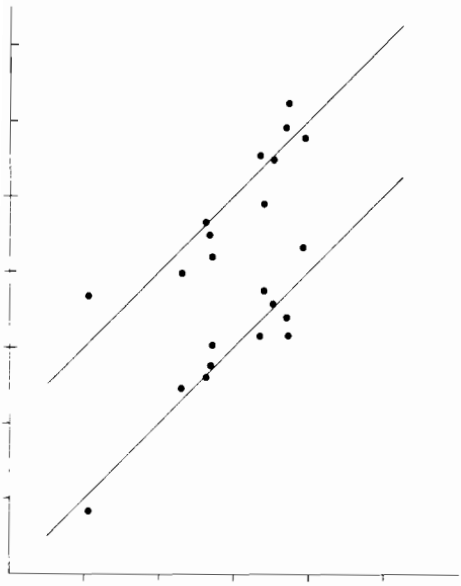

Figure 1. Graphical control of the model, "negative" items 3, 5, 6, 7, 8, 10, 12, 13, 15, 16, 18.

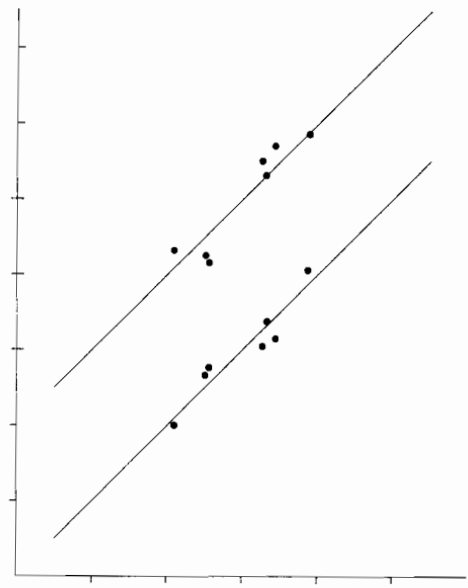

Figure 2. Graphical control of the model model, "negative" items 5, 6, 10, 13, 15, 16, 18.

The results for the "negative" items are summarized in Figure 1. The corresponding test statistic is 37.4823. It lies beyond the 99.9% percentile of the limiting χ^2-distribution with 10 degrees of freedom. The assumption that the marginal conditions are sufficiently stable for all of the "negative" items can thus be rejected with high statistical significance. Since the statements codetermined by "reading habits" of the students (statements 3, 7, 12) were still included among the "negative" items, this result is in line with the prior findings by Lehrke and Kempf (1975b). Moreover, it turned out that additional statement No. 8 did not prove useful, either. Not surprisingly, difficulties in concentrating on the subject-matter can also be caused by a multitude of factors other than the lack of subject-matter directed motivation.

After omitting these statements, the remaining items (statements 5, 6, 10, 13, 15, 16, 18) showed reasonably good homogeneity: the test statistic 10.5254 is between the 80% and the 90% percentile of the asymptotic χ^2-distribution with 6 degrees of freedom. Figure 2 shows the corresponding graphic.

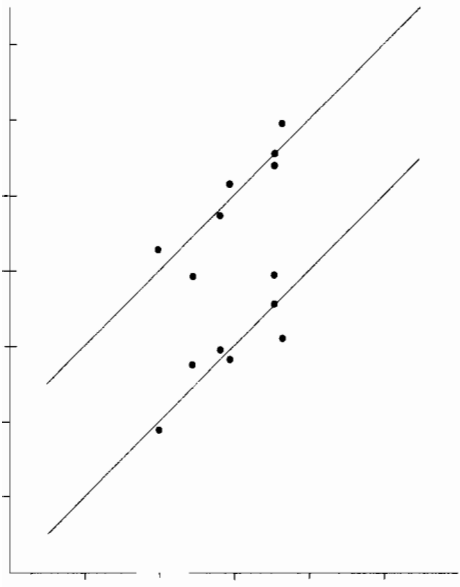

Figure 3. Graphical control of the model. "positive" items (1), 2, 4, (9), 11, 14, 17, 19, 20.

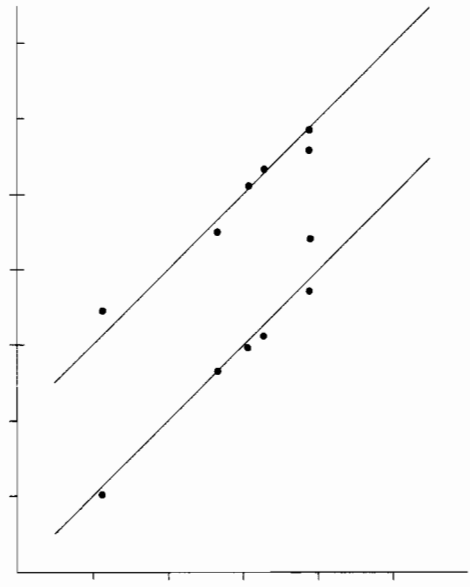

Figure 4. Graphical control of the model. "positive" items 1, 4, 11, 14, 17, 20.

The analysis of the "positive" items leads to results summarized in Figure 3. Items No. 1 and 9 were not included in this analysis. In view of the generally high level of subject-matter directed motivation among the students involved, these item proved to be too easy. Almost all of the students agreed. The remaining items plotted in Figure 3 do not yet form a homogenious set. The corresponding test statistic 48.1005 lies beyond the 99.9% percentile of the asymptotic χ^2-distribution with 6 degrees of freedom. Two items (statements 2 and 19) had to be omitted in order to obtain the results shown in Figure 4. The test statistic pertaining to the items 1, 4, 11, 14, 17 and 20 is 6.0663 and indicates a satisfactory fit of the model. With 5 degrees of freedom the test statistic lies approximately at the 70% percentile.

A slight misformulation in item No. 2, "*Dinge die wir zuletzt gelernt haben*", caused the students to construe this item as a reference to the very last lesson, not the overall course. In that it dealt with temperature regulation in engineering and not biology, this lesson was untypical for the course, however. The evoked subject-matter directed motivation is not independent of long-term motivation and there are considerable deviations in students' long term motivation towards biology and physics. Sex differences may also play a role (cf. Todt et al., 1974).

The reactions to item No. 19 were possibly influenced by the social background of the students, which codetermines the students' chance of communicating about school subjects with his or her parents and peers. While the present study involved both municipal and rural schools, the social background of the students was more homogenious in the preceding study by Lehrke and Kempf (1975b). This may have caused that the short-comings of item No. 19 could not be detected then.

5. APPLICATION

After omitting the weak items, two sets of items remained for which the marginal conditions were sufficiently stable. This result concurs with that of the previous study by Lehrke & Kempf (1975 b) and underlines the possibility of producing homogenious item pools for measuring subject-matter directed motivation.

However, even if the marginal conditions are stable: evaluating teaching on the basis of the amount of subject-matter directed motivation evoked in the students remains problematic. While it is possible that certain forms of teaching specially favour particular modes of behavior (so that the difficulties of the modes of behavior change with the teaching), a comparison of different kinds of teaching with respect to their degree of subject-matter directed motivation evoked in the students can only be justified then, when the different modes of behavior which are regarded as motivated are aroused to the same extent. Otherwise, a differential evaluation is needed: which modes of behavior are favoured more strongly by which type of teaching?

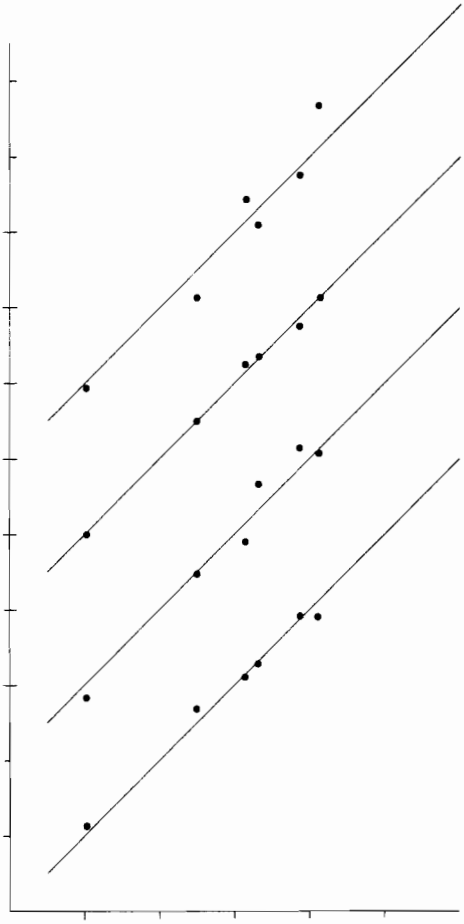

Figure 5. Relative stability of the difficulties of the "positive" items.

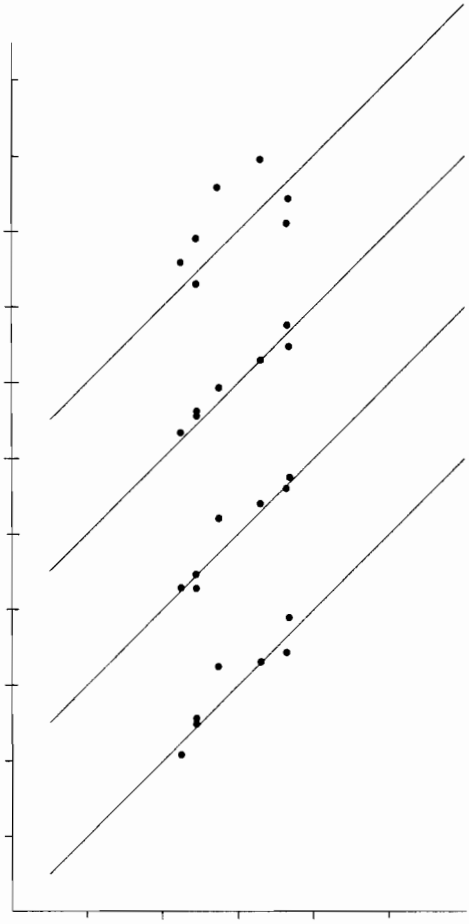

Figure 6. Relative (in-)stability of the difficulties of the "negative" items.

Such differential evaluation would require new item analyses with respect to each application. however. Although differential evaluation of teaching with respect to the subject-matter directed motivation evoked may be an important goal in educational research. in most cases the teacher will lack both theoretical training and technical facilities to accomplish it in his everyday work. For routine application in school one would therefore ask for a questionnaire that refers only to such modes of behavior for which the item-difficulties are stable under varying teaching conditions and course contents.

In order to investigate whether the retained item sets meet this requirement. the questionnaire was administered to the same students in other biology courses. also. These courses dealt with ecological problems. blood circulation and blood pressure. and hormones. In different classes different amounts of kybernetic models were used for the explanation of the phenomena taught.

Figures 5 and 6 show the relative stability of the difficulties for the "positive" and "negative" items. The corresponding test statistics are 16.4804 (df = 15) and 96.6014 (df = 18) for the "positive" and "negative" items. respectively.

Evidently. the item difficulties are sufficiently stable for the "positive" items. Table 2 containes the overall difficulties of these items estimated from the total set of data pertaining to all courses to which the questionnaire was applied.

Table 2. Overall estimates of item difficulties for the "positive" items, 1, 4, 11, 14, 17, 20.

i	ε_i
1	1.9718
4	.3227
11	.8707
14	1480
17	-.4940
20	1.1245

The relative instability of the "negative" items is highly significant. This relative instability points out that the responses to the "negative" items are in fact influenced by the teaching. Hence. they do measure the evoked (and not the longterm) subject-matter directed motivation. However. the modes of behavior to which they refer are not evoked uniformly. Which of them are favoured more strongly and which are not. may depend on the teaching and the course content.

On the "positive" items there is – if any – a uniform effect of course content. Specifying the average motivation in a class during instruction on content c as $\bar{\xi}_c = \bar{\xi} + \nu_c$ where $\bar{\xi}$ denotes the average initial motivation in class and where ν_c denotes the average effect of instruction. testing the null hypothesis $\nu_c = 0$ can be used as a device for assessing whether the "positive" items actually do measure the *evoked* subject-matter directed motivation. This hypothesis can be tested by means of a likelihood ratio test. For the present data the resulting test statistic is 106.5866. which lies beyond the 99.9% percentile of the limiting χ^2-distribution with 11 degrees of freedom. The null-hypothesis according to which only long-term motivation is measured can thus be rejected.

In the everyday school situation the teacher may not be in a position to make use of parameter estimates. Moreover, he may lack the teachnical facilities for computing likelihood ratio tests. Nonetheless, it is easy for him to assess how many questions a student has checked in the direction of motivation. For the routine application of the questionnaire, therefore, it is important that these raw scores be still sensitive enough to detect variations in subject-matter directed motivation.

In order to test whether this is the case, two groups of students were compared with respect to the mean scores on the retained "positive" items. Both groups of students had been instructed on the subject of hormones. While the instruction of group 1 ($n = 47$) involved kybernetic explanations, this was not the case in the instruction of group 2 ($n = 54$). The hypothesis was that the use of kybernetic models would make instruction less dull, thus increasing the students' subject-matter directed motivation. The mean raw scores for the two groups are 4.36 and 3.67. The respective standard deviations are 1.67 and 1.82. Consequently, $t = 1.96$ with 99 degrees of freedom, and the hypothesis is thus confirmed by the empirical findings.

As stated above, subject-matter directed motivation is supposed to have an impact both on the results of cognitive learning and certain educational goals of a higher order. In particular, there should be a dependency between subject-matter directed motivation and achievement among students exposed to the same instruction. In the present study such homogeneity of instruction was only given in the introductory course on kybernetic models and temperature regulation. Although the investigation of correlations between motivation and achievement were not part of the focus of our study, the data pertaining to this introductory course can be used for a rough assessment.

Since a deficit in subject-matter directed motivation can be compensated by other forms of motivation, and since achievement in class is co-influenced by a multitude of social factors (e.g., sociometric structure, hierarchy of achievement levels, roles) and general ability or "talent" factors, correlations between achievement and subject-matter directed motivation *within a given class* may cancel. In fact, these correlations tended to be very small. The mean correlation was $r = -.019$. The standard deviation of the coefficients was $s_r = .22$. Two of the six correlation coefficients even had a negative value.

A lack of subject-matter directed motivation within a whole class cannot be compensated completely, however. Since the achievement of the others in a class form a frame of reference to which the individual students relate their own efforts, subject-matter directed motivation is essential for the over all cognitive outcome of a course: When the mean achievement scores in the classes were correlated with the mean motivation raw scores the result was $r = .58$ for the "positive" items and $r = .97$ for the "negative" ones.

This result indicates that the modes of behavior to which the "negative" items refer are not only more strongly affected by the teaching than those to which the "positive" items refer. They are also more closely related to achievement. In view of the relative instability of the difficulties of the "negative" items, however, routine application of these items is possible only to the evaluation of differences in motivation within a given course; not to the comparison of different courses, as this calls for a differential evaluation in order to determine which of the modes of behavior are favoured and to what extent so.

6. FINDINGS IN BRIEF

The concept of subject-matter directed motivation, measured by means of questionnaires, proved useful for evaluation, though it covers and emphasizes only one aspect of those modes of behavior, which are regarded as motivated. Since the behavior can be codetermined by other conditions as well, the concept has to be applied under certain restrictions. It must be assured that inference is based only on such modes of behavior for which these marginal conditions are relatively stable within the group of students under consideration.

The course of instruction itself may effect such marginal conditions. Furthermore it is possible that different kinds of teaching favour particular modes of behavior. An overall comparison of different kinds of teaching with respect to the degree of subject-matter directed motivation evoked in the students can be justified only, when the different modes of behavior are aroused to the same extent. Otherwise, a differential evaluation is needed: which modes of behavior are favoured more strongly by which type of teaching ? Such differential evaluation would require new item analyses with respect to each application.

The questionnaire was administered to the same students four times in different courses. Separate analyses of the "positive" and the "negative" items led to the results, that the difficulties of the "positive" items were sufficiently stable, while those of the "negative" ones were significantly influenced by the teaching. Hence, routine application of the latter item-set is possible only to the evaluation of differences in motivation within a given course but not in the comparison of different courses.

Both item-sets are able to measure variations in *eroked* subject-matter directed motivation. As indicated by high positive correlations between mean achievement and mean motivation scores of school classes, subject-matter directed motivation is closely related to achievement.

REFERENCES

ANDERSEN, E. B. A goodness of fit test for the Rasch model. *Psychometrika* 1973, **38**, 123.

HECKHAUSEN, H. Leistungsmotivation. In: Thomae, H. (Ed.) *Motivationslehre. Handbuch der Psychologie. Bd.* 2. Göttingen: Hogrefe. 1965.

KEMPF, W. F. Basisprobleme der Diagnostik der Aggressivität. In Eckensberger, L. (Ed.) *Bericht über den* 28. *Kongreß der Deutschen Gesellschaft für Psychologie in Saarbrücken* 1972. Göttingen: Hogrefe, 1974.

KEMPF, W. F. Dynamic models for the measurement of "traits" in social behavior. In Kempf, W. F., Repp, B. (Eds.) *Some Mathematical Models for Social Psychology.* Bern, Huber. 1976 (in print).

LEHRKE, M., KEMPF, W. F. Sachbezogene Motivation im Unterricht und die Messung ihrer Veränderung. In Tack, H. (Ed.) *Bericht über den* 29. *Kongreß der Deutschen Gesellschaft für Psychologie in Salzburg* 1974. Göttingen Hogrefe, 1975 (a).

LEHRKE, M., KEMPF, W. F. Diagnostik der sachbezogenen Motivation im Schulunterricht. *Zeitschrift für Erziehungswissenschaftliche Forschung* 1975(b). (in print).

MC CLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., LOWELL, E. L. *The Achievement Motive.* New York, 1953.

RASCH, G. *Probabilistic Modells for some Intelligence and Attainment Tests.* Copenhagen: Nielson & Lydiche. 1960.

TODT, E., ARBINGER, R., SEITZ, H., WILDGRUBE, W. *Untersuchungen über die Motivation zur Beschäftigung mit naturwissenschaftlichen Problemen.* Bericht über das Projekt Motivation am Fachbereich Psychologie der Justus-Liebig-Universität Gießen, 1974.