



HAL
open science

Thought structures as an instrument to determine the degree of difficulty of modelling tasks

Xenia-Rosemarie Reit, Matthias Ludwig

► **To cite this version:**

Xenia-Rosemarie Reit, Matthias Ludwig. Thought structures as an instrument to determine the degree of difficulty of modelling tasks. CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Charles University in Prague, Faculty of Education; ERME, Feb 2015, Prague, Czech Republic. pp.917-922. hal-01287266

HAL Id: hal-01287266

<https://hal.science/hal-01287266>

Submitted on 12 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thought structures as an instrument to determine the degree of difficulty of modelling tasks

Xenia-Rosemarie Reit and Matthias Ludwig

Goethe-University Frankfurt, Institute for Mathematics Education, Frankfurt, Germany, reit@math.uni-frankfurt.de

Although efforts have been made to integrate the concept of mathematical modelling in school, studies show that it has not arrived yet in everyday school classes. From a teacher's point of view, multiple solution approaches and a varying task difficulty complicate especially the development and assessment of modelling tasks. Taking up this issue, this study aims at developing a method to determine the degree of difficulty of solution approaches of modelling tasks based on so called thought structures of student solutions. Thought structures of student solutions provide information about the task difficulty and can be taken as a basis for a well-founded rating scheme. We want to consider the question of whether the method used to describe the degree of difficulty can be reproduced by empirical results.

Keywords: Mathematical modelling, assessment, degree of difficulty, task space.

INTRODUCTION

Within the community of mathematics education there is a broad consensus that the integration of mathematical modelling and applications must be promoted and increased. Of course this awareness is not new and efforts have been made during the last decade. However, several studies provide evidence that modelling is far away from playing an integral role in daily school teaching in Germany and also elsewhere (Blum, 2007). The proportion of modelling in daily school routine is rather low (Jordan et al., 2006). When researching into that problem it is worthwhile to have a closer look on the teachers' point of view. What prevents teachers from teaching modelling tasks? Schmidt (2010) found out that teachers often mention complexity and lack of predictability as motives for waving modelling tasks. A look in textbooks

shows that modelling tasks are still rare, especially those which can be used in a normal school class setting in contrast to modelling tasks used within larger projects encompassing several lessons. The development of modelling tasks leads, more than ordinary task formats, to difficulties concerning the estimation of the task space, the task difficulty and finally also concerning the assessment.

The study presented here takes up the issues from a teacher's perspective and aims at a better structuring of modelling tasks in terms of a better manageability for teachers. In detail we developed a method to determine the degree of difficulty of modelling tasks and an assessment scheme, both, building on solution approaches with its particular thought structures. Empirical results will show whether the theoretically determined degree of difficulty is verifiable and certifies a good applicability.

THEORETICAL FRAMEWORK AND METHOD

A common instrument to determine the degree of difficulty is the solution rate by applying a dichotomous rating. Since the answer format of modelling tasks is open and compared to others rather extensive, this procedure seems not adequate to reflect the full scope which is provided by modelling tasks. Cohors-Fresenborg, Sjuts and Sommer (2004) applied a method analysing the text of the task. They identified task specific indicators for the difficulty of tasks by investigating PISA-2000 items. The task format of the investigated items was not restricted to modelling tasks. Since the task space of modelling tasks is, compared to other task formats, rather large, this aspect would get lost by focusing on the text of the task. In our study this specificity of modelling tasks is taken into account since the method for determining the

degree of difficulty comes directly out of the solution approaches of the student solutions.

In a first step we analysed the student solutions of every modelling task to identify main solution approaches. On the one hand, this classification was based on the mathematical model used and, on the other hand, on the solution process, which can be different although the same mathematical model is used. Each solution approach has a specific structure which can be revealed by dismantling the solution approach into its single thought steps, which provide the opportunity to consider solution steps from cognitive aspects. This idea builds on structural considerations in the field of word problems. In this context Breidenbach implemented the term “Simplex” as a task consisting of three items and every item can be determined by the two others (Breidenbach, 1963, p. 200). By visualizing the logical structure of the mathematical operations to be done in terms of such Simplexes he built up a kind of arithmetic tree or flow chart. These flowcharts have been refined by Winter and Ziegler (1969) and serve as a basis for the so called thought structure analysis of the study described here. For the present study we refined Breidenbach’s definition by identifying the single cognitive steps, here called thought operations, which have to be carried out in order to arrive at a solution. We define thought operations as follows:

A thought operation is a necessary (intermediate) result which is obtained directly (without intermediate calculations) from one or several (initial) data.

These thought operations can then be arranged as a kind of flow chart which illustrates the incremental proceeding and in addition, also the complexity of the solution process (see Figure 1).

A natural but empirically not yet validated conclusion is that the complexity of a mathematics task is dependent on the number of simplexes and the nesting of them (Graumann, 2002, p. 93). Cohors-Fresenborg and colleagues (2004) emphasize in this context the simultaneity and nesting of thought steps. With the help of theories within the field of cognitive psychology, we can operationalize the effects of nesting and simultaneity on complexity. Fletcher and Bloom (1998) found in their study about text comprehension, where they assumed text comprehension to be a kind of problem solving process, that information being the direct pre-

decessor of another information must be kept actively in the working memory. Under the assumption of a working memory with limited capacity (Sweller, 1988) these findings indicate that several aspects in a task which are related to each other and have to be considered and understood at the same time, may load the working memory. Thought operations are considered to be parallel if they either originate from the same thought operation or determine the same thought operation of the subsequent solution level.

By applying the findings from above to the thought structure considerations, especially the amount of parallel thought operations (as it is the case in level two of Figure 1) of a solution approach appears to be a difficulty generating aspect. Thus, during the process of determining the degree of difficulty, parallel thought operations complicate the solution approach to a greater extent than those being processed consecutively. To describe this circumstance we developed a model which gives more weight to parallel than consecutive thought operations. Each level in the thought structure contributes to the overall degree of difficulty according to its number of parallel thought operations. From the number of parallel thought operations per solution level we calculate the factorial and then finally all levels are added up. In the following we want to reconstruct this procedure by taking modelling task “Potato” (Figure 2) as an example.

One solution approach, which we could identify, is called “Layer” (Figure 3, left). Based on the given length of the potato the student assumes a height and a diameter or depth of the potato. Together with an assumption about the measures of a potato stick the student is then able to calculate the number of sticks

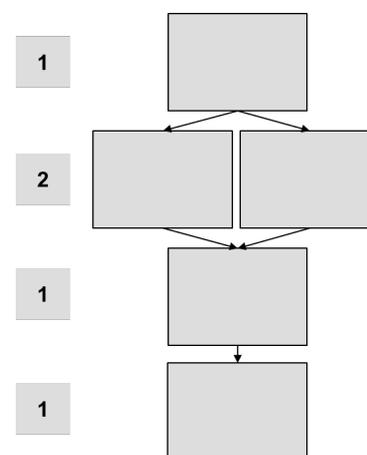


Figure 1: Exemplary thought structure of a solution approach together with the number of thought operations per level

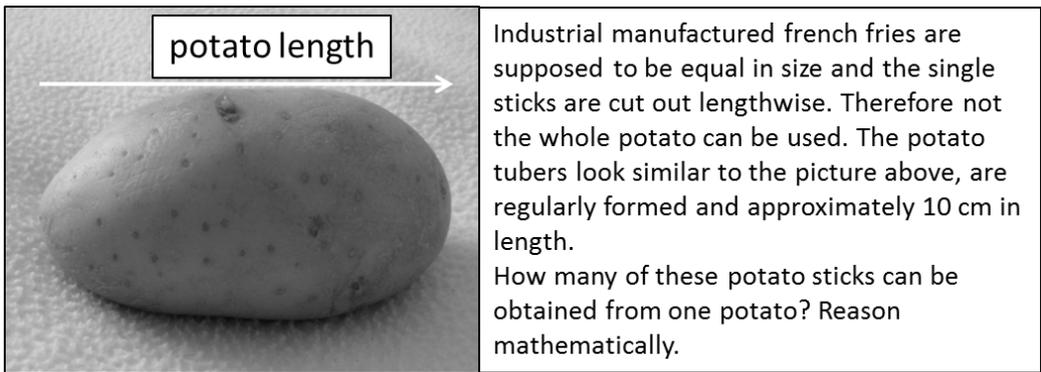


Figure 2: Modelling task "potato"

per layer in height and depth. Multiplication of these two numbers leads to the total number of potato sticks. The corresponding thought structure has been developed which arranges these thought operations as sort of arithmetic tree (Figure 3, right). Besides the chronological order of the calculations to be made, the thought structure also provides information about the difficulty of the solution approach. The numbers right next to each solution level indicate the number of parallel thought operations to be completed per thought structure level. As explained above summing up the factorial per level yields the degree of difficulty of the solution approach. Thus, in case of solution approach "Layer", the degree of difficulty is $1!+2!+1!+1!=5$ (see grey Figure 3, right).

Based on these considerations an assessment scheme has been developed to utilize the full scope of the promising method on the one hand and to verify the theoretically determined degree of difficulty on the other hand. When observing the assessment routine of mathematics teachers it becomes clear, that assessment of tasks is based on a sample solution and its

important partial aspects. The teacher compares each solution with a sample solution and awards important intermediate results. Our assessment scheme is based on this everyday school routine procedure. Intermediate results are represented here by thought operations and are assessed according to whether they have been conducted correctly or wrong. For partly right thought operations half points can be awarded. That means that thought operations are scored 0, 0.5 or 1 dependent on their completion. Thus, the maximum score per solution approach is defined by the number of thought operations. In case of solution approach "layer" the maximum score is 5 (number of thought operations (grey boxes) in Figure 3, right).

Before being able to solve the task, one has to identify the relevant information of the presented linguistic context. According to Cohors-Fresenborg & Sjuts (2001) difficulties are especially evoked by linguistic constructions concerning the logical structure and formulations conditioned by the authenticity of the situation. The textual differences of the tasks are integrated in the process of determining the level of

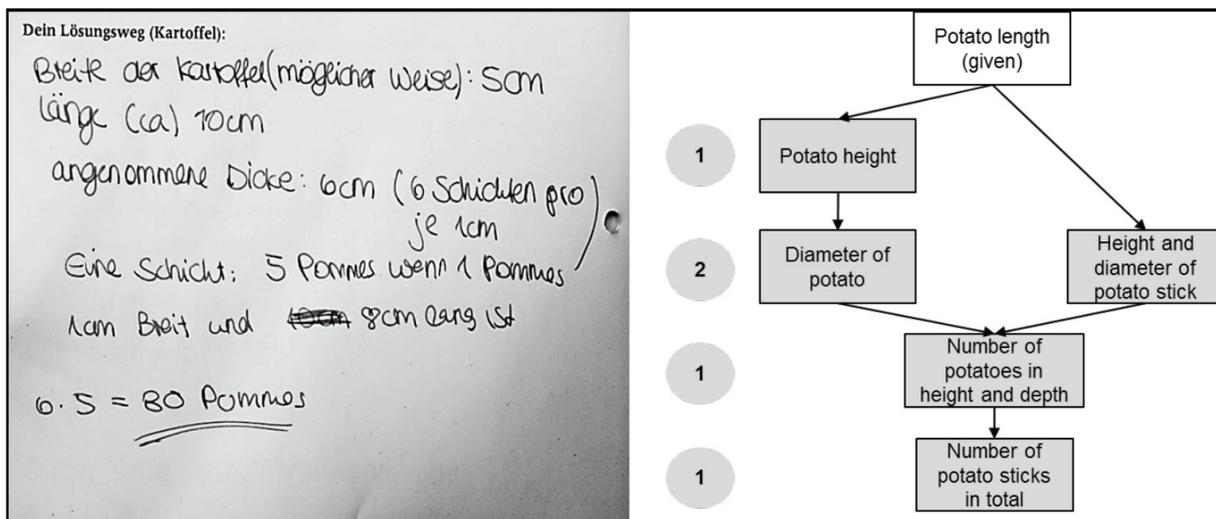


Figure 3: Solution of solution approach "Layer" and corresponding thought structure

difficulty by adding zero, one or two single thought operations according to the linguistic requirement of the task text.

- 0 Picture and/or text contain all relevant data. Only simple main clauses are used.
- 1 Missing data has to be estimated and there are several main and subordinate clauses due to a larger text, containing explanatory but mathematically irrelevant text passages.
- 2 Picture and/or text do not explicitly clarify the dimension of the object of the task, e.g. the two dimensional picture of a potato is out of keeping with the three dimensional reality leading to not include a third dimension in the calculations.

Modelling task potato has been assigned to linguistic level two according to its requirements coming from the text. Therefore we additionally add two single thought operations to the solution approach based degree of difficulty. Thus the total degree of difficulty of the solution approach "Layer" consists of linguistic requirements (2) and solution approach specific thought operations (5) and finally adds up to $2+5=7$.

STUDY DESIGN

Five modelling tasks have been developed according to predefined criteria (see Reit & Ludwig, 2013):

- Authentic context (Maaß, 2007)
- Realistic numeric values (Müller et al., 2007)
- Problem solving character (Maaß, 2007)
- Naturalistic format for questions
- Openness relating to the task space

Authenticity and relation to reality are core elements of modelling tasks. We want to avoid ostensible relations to reality like they are often used in word problems in textbooks. There is no a priori known solution algorithm for the task which can be directly applied by the students. That means that the solution makes itself out to be a problem on students' level. The questioning is supposed to either be close to the

living environment of the students or take up a realistic question which could arise in reality. Openness of tasks is reflected by the task space. There has to be more than one solution approach which leads to a solution. The solution approaches distinguish themselves by their mathematical model. Thereby the students are able to have more options to arrive at a solution. Openness should rather be based on the alternatives of mathematical models to solve the tasks than on approximating sizes. Demanding this we do not deny that making assumptions is an important part of mathematical modelling but we want it to be limited to a degree which ensures an assessable solution interval.

The study splits up into pilot study and main study. The pilot study encompassed

- an a-priori definition of the task space of the tasks,
- the identification of the thought structure for each solution approach
- and the establishment of a rating scheme for each solution approach.

The implementation into a normal 45-minute lesson also required a time limitation of approximately 10–12 minutes per task. During the main study from December 2013 until April 2014 approximately 1800 students of grade 9 (15/16 years of age) of German grammar schools took part and solved three modelling tasks each.

To validate the thought structure method it must be investigated in how far the theoretically determined degree of difficulty is empirically reproducible. In other words, will tasks, which have been rated as difficult, be solved worse than those rated as rather simple? To do so the student solution together with its assessment is associated with the predetermined degree of difficulty of the respective solution approach.

VALIDITY OF THE THOUGHT STRUCTURE ANALYSIS

The modelling tasks, which have been developed especially for the purposes of the study, show a good variability of applied solution approaches. Although this was an important criteria during the process of development it could not be stated with certainty how

the students solve the task. At least three target aimed solution approaches could be identified per modelling task (Figure 4). Hence, from the viewpoint of openness in the sense of an adequate task space the developed modelling tasks seem to meet the requirements.

The main objective of the study is the validation of the thought structure method to determine the degree of difficulty by comparing the student performance with the statement of the theoretical model. Hereafter we consider the question of whether the method makes a valid statement about the degree of difficulty of the modelling tasks.

Within a modelling task the student performance verifies the degree of difficulty in terms of scoring higher using solution approaches determined as easier and vice versa. This situation is illustrated in Figure 5. Taking three of the five modelling tasks into account (modelling task “Cola” and “Bridge” are not yet analysed completely) we can see a distinct decrease of the average score when using a solution approach being determined as more difficult. The value of the power of the regression function is in case of modelling task “Tennis” and “Taj Mahal” nearly similar what indicates that the coherence of score and degree of difficulty reacts almost in the same way. The degree of difficulty of modelling task “Potato” seems to have a firmer influence on the average score than it is the case with the two other modelling tasks. An increase in difficulty results in a steeper deterioration. In general a comparability of solution approaches within a modelling task, based on the theoretical degree of difficulty is reasonable.

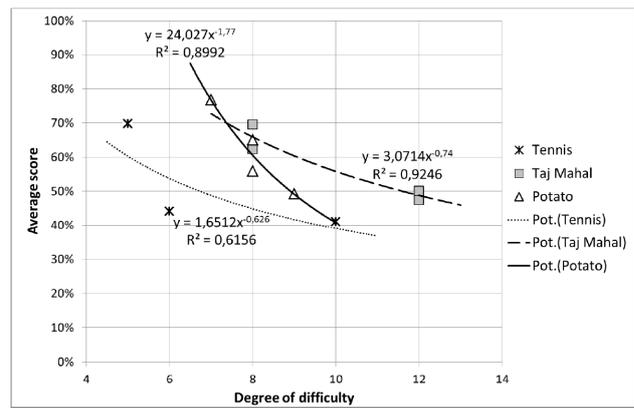


Figure 5: Coherence of score and degree of difficulty concerning the solution approaches per task

Another interesting question is to what extent the determination of the degree of difficulty can be used in a cross-task context. What statements can be made by comparing different modelling tasks? To answer this question we compared the average score of a modelling task dependent on its average degree of difficulty. The decreasing tendency of the average score with increasing degree of difficulty is in evidence (Figure 6). In detail the power of the regression function suggest that if the degree of difficulty doubles, then the score will half.

SUMMARY AND OUTLOOK

Especially when considering modelling tasks it is challenging to estimate the degree of difficulty and to assess their solutions reasonably and satisfying for students. On the part of teachers this brings along uncertainty and might contribute to a restrained attitude towards modelling tasks in everyday teaching. The presented study considers that problem and de-

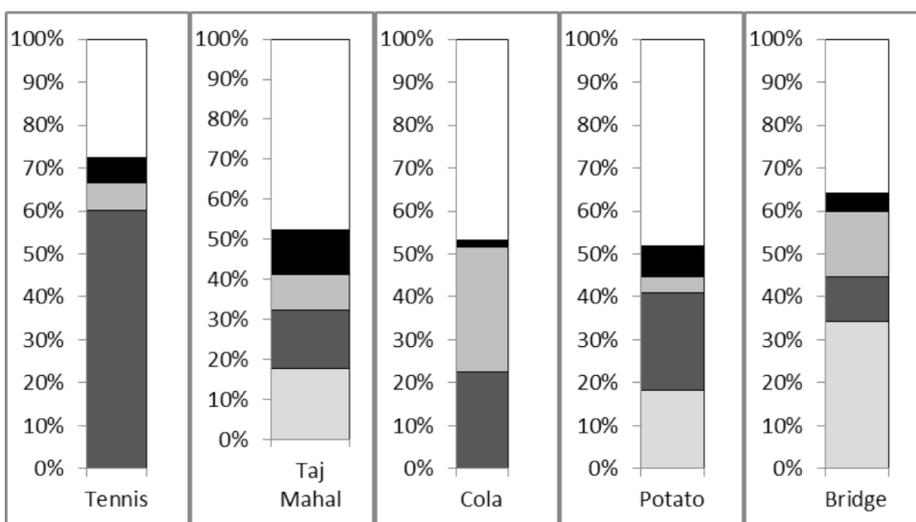


Figure 4: Distribution of solution approaches per modelling task (target aimed solutions in greyscale, others white)

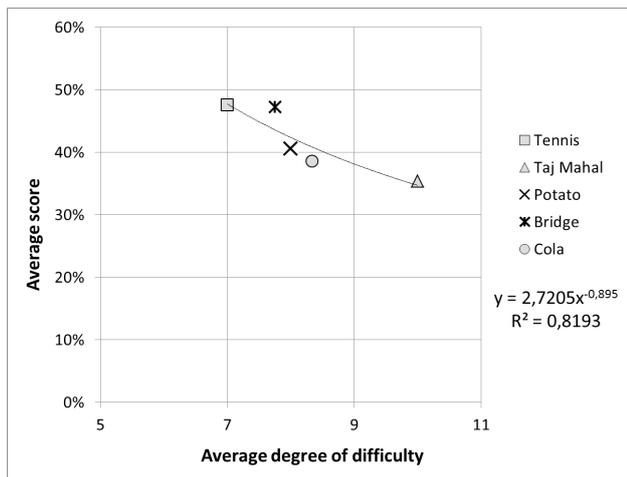


Figure 6: Average score of the modelling tasks subject to the degree of difficulty

veloped a promising method to firstly determine the degree of difficulty and secondly to form the basis for a reasonable and conclusively substantiated assessment scheme. The analysis so far encourages the assumption of a good applicability. The results up to now show that using the thought structure based method to determine the degree of difficulty is in line with the empirical results. Students show better results when using those solution approaches with a rather low degree of difficulty and vice versa. Similar results could be obtained considering the modelling tasks as a whole. The average student performance is better at easier tasks and worse at tasks being rated as more difficult. Additional analysis especially to modelling tasks “Potato” and “Bridge” will show how reliable and convincing the method is. Besides these affirmative outcomes modelling task “Potato” gives rise to questions concerning the limits of the method. The divergent results of this task may be due to the fact that their solution approaches are somehow intertwined in the sense of being difficult to distinguish. This sometimes leads to the problem of false solution approach classification, thus to a falsification of results. This may support the conclusion that a distinct discriminability is a necessary requirement for the thought structure method. Further reflection and the complete analysis of the two remaining modelling task may give deeper insights to that aspect.

REFERENCES

Blum, W. (2007). Mathematisches Modellieren – zu schwer für Schüler und Lehrer? In *Beiträge zum Mathematikunterricht 2007* (pp. 3–12).

Breidenbach, W. (1963). *Rechnen in der Volksschule: Eine Methodik* (11., neubearb. und erw. Aufl). Hannover: Schroedel.

Cohors-Fresenborg, E., & Sjuts, J. (2001). Die Berücksichtigung von kognitiver und metakognitiver Dimension bei zu erbringenden und zu beurteilenden Leistungen im Mathematikunterricht. In C. Solzbacher & C. Freitag (Eds.), *Anpassen, Verändern, Abschaffen? Schulische Leistungsbewertung in der Diskussion* (pp. 147–162). Bad Heilbrunn: Klinkhardt.

Cohors-Fresenborg, E., Sjuts, J., & Sommer, N. (2004). Komplexität von Denkvorgängen und Formalisierung von Wissen. In M. Neubrand (Ed.), *Mathematische Kompetenzen von Schülerinnen und Schülern – Vertiefende Analysen im Rahmen von PISA 2000* (pp. 109–138). Wiesbaden: VS Verlag für Sozialwissenschaften/ GWV Fachverlage GmbH.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language*, 27(3), 235–244. doi: 10.1016/0749-596X(88)90052-6

Graumann, G. (2002). *Mathematikunterricht in der Grundschule*. Bad Heilbrunn/ Obb.: Klinkhardt.

Jordan, A., Ross, N., Krauss, S., Baumert, J. B., Blum, W., Neubrandt, M., Löwen, K., Brunner, M., & Kunter, M. (2006). *Klassifikationsschema für Mathematikaufgaben. Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt. Materialien aus der Bildungsforschung Nr. 81*. Berlin: Max-Planck-Institut für Bildungsforschung.

Maaß, K. (2007). *Mathematisches Modellieren – Aufgaben für die Sekundarstufe*. Berlin: Cornelsen Verlag Scriptor.

Müller, M., Leib, D., Schukajlow, S., Blum, W., & Messner, R. (2007). Auswendiggelernt – Abgehakt – Abgefragt? In *Beiträge zum Mathematikunterricht* (pp. 723–726). Hildesheim: Franzbecker.

Reit, X.-R., & Ludwig, M. (2013). Wege zu theoretisch fundierten Testaufgaben zur Modellierungskompetenz. In G. Greefrath, F. Käpnick, & M. Stein, *Beiträge zum Mathematikunterricht 2013* (pp. 805–808). Münster: WTM-Verlag.

Schmidt, B. (2010). *Modellieren in der Schulpraxis: Beweggründe und Hindernisse aus Lehrersicht*. Hildesheim, Berlin: Franzbecker.

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2), 257–285. doi: 10.1207/s15516709cog1202_4

Van Gerven, P. W. (2003). The efficiency of multimedia learning into old age. *British journal of educational psychology* 73(4), 489–505. doi: 10.1348/000709903322591208

Winter, H., & Ziegler, T. (1969). *Neue Mathematik: Lehrerheft*. Hannover: Schroedel.