ISSN 1512-1801

UDC 37.02

DIDACTIC COMPLEXITY OF GENERAL PHYSICS COURSE SECTIONS: EVALUATION RESULTS

Robert Mayer
Doctor of Pedagogical Sciences, Professor of Physics and Didactic of Physics Chair
The Glazov Korolenko State Pedagogical Institute, Glazov, Russia

Abstract: *The article is devoted to the development of a method for assessing the didactic complexity of various sections of the general physics course, which takes into account the complexity of text and formula information. The essence of the proposed method is as follows: 1) write out ten key concepts with the greatest complexity from each sub-paragraph; 2) collect all keywords in one text file and divide them into several difficulty categories; 3) using the computer program Slozhnost.pas accessing the dictionary Slovar.txt, analyze the files with keywords and determine the total amount of information contained in the key concepts from the various paragraphs; 4) estimate the proportion of mathematical and other formulas for each paragraph; 5) for each sub-paragraph, write out 5 key terms of maximum complexity used in mathematical formulas; 6) for each paragraph, create files containing keywords included in the formulas; 7) assess the complexity of each term and use the program Slozhnost.pas to determine the total amount of information in keywords for each paragraph; 8) from different sections of the manual, randomly select sentences, determine the average words length and the average number of syllables in words, find an indicator of the complexity of the sentences structure; 9) for each paragraph, calculate differential and integral complexity for text and formula information.*
Keywords: *complexity, didactics, educational text, information, knowledge folding, physics, semantics.*

Introduction

The optimization of the educational process requires a systematic assessment of the didactic complexity (DC) of various elements of the educational material. The higher the complexity of the studied issue, the more time it takes to consider it at the lesson, and the more effort a schoolchild or student should spend to study it [1]. Therefore, the development of effective methods for assessing the DC of the text, the identification of difficult-to-understand issues are an important practical tasks; their solution allows to increase the effectiveness of the educational process. Of particular importance is the problem of assessing the DC of educational texts in the general course of physics which is studied in most Russian universities.

The purpose of the article: to develop a methodology for evaluating the DC of large-volume physics training texts taking into account the complexity of textual and formula information, and to test it on one of the textbooks for the general physics course. This work is related to the problems of determining the cognitive complexity of various systems, the use of quantitative methods in didactic research, measuring the overall information content and average information density of educational texts, and is developing the method proposed by the author in [2, 3].

The methodological basis of the research is the ideas of the following scientists: 1) Je. G. Gel'fman and M. A. Holodnaja [4] (the psychology and learning theory); 2) I. P. Kuznecov [5], L. A. Chernjahovskaja [6] (the semantic information theory); 3) Val. A. Lukov, Vl. A. Lukov [7] (the thesaurus approach); 4) Ya. A. Mikk [1] (the educational text theory); 5) M. D. White and E. E. Marsh [8] (the content analysis of texts); 6) T. V. Batura [9], A. S. Kisel'nikov [10], Ju. N. Marchuk [11] (automated assessment of the text complexity). In this case, the methods of analysis and synthesis, separation of the main, pair comparison, expert assessments, counting markers, complex indicators, histograms, elements of statistical and cluster analysis has been used.

ISSN 1512-1801

1. Essence of the didactic complexity determination method

In order to evaluate the DC of the general physics course sections, a content analysis of the textbook [9] is carried out. Its choice is due to the following: 1) the manual covers all sections of physics; 2) the manual is written by a team of 4 authors (Yu. A. Barkov, G. N. Votinov, O. M. Zverev, A. V. Perminov), it is reviewed and approved; 3) there is a pdf version of the textbook; 4) the manual has a small volume (407 pages). It consists of two parts: 1) theoretical foundations of general physics; 2) control and laboratory work. Only theoretical material is analyzed; examples of problem solving, control and laboratory work are not taken into account. The theoretical foundations consist of five chapters corresponding to the main sections of physics: 1) Mechanics (M); 2) Molecular physics and Thermodynamics (MF+T); 3) Electrodynamics (E); 4) Optics (O); 5) Atomic and nuclear physics (ANP). Each chapter contains from 2 to 5 paragraphs; for example, the first chapter «Mechanics» includes the following paragraphs: 1.1. «Kinematics», 1.2. «Dynamics», 1.3. «Oscillations and Waves», 1.4. «Hydro- and Aeromechanics». The total is 18 paragraphs, consisting of several sub-paragraphs of 3 to 5 pages, which can be characterized by textual and formula complexity.

To determine the textual complexity of a subparagraph, 10 key terms with the greatest complexity should be written out from it. For example, sub-paragraph 2.1.3. «Average energy of molecules» is characterized by the following set of key terms (pressure, atom, molecule, absolute temperature, kinetic energy, ..., average square of velocity, energy capacity). They are the ones that present the greatest difficulty to understand; the assimilation of this sub-paragraph depends on their assimilation. If you combine all the key terms included in the paragraph, add up their complexity s_i, and divide the sum by their total number (from 20 to 100), then the resulting information folding coefficient (IFC) can be considered as an indicator of the average complexity of textual information in this paragraph. Similarly, the complexity of the formula information is estimated: for each sub-paragraph, 5 key terms are written out, which are included in the formulas. It turns out like this: 3.2.1. «Characteristics and conditions of the existence of direct current» (derivative, integral, work of external forces, EMF, current density). To calculate the text structural complexity, the formula is used: $C = D_W \ln(D_S + 1)$, where D_W – is the average number of syllables in words (not counting conjunctions, prepositions), D_S – is the average number of words in a sentence [2]. A school graduate can easily read a sentence of 7 words with an average word length of 3 syllables, the structural complexity of which is $C_0 = 3\ln(8) \approx 6{,}24$.

To assess the semantic complexity of concepts, it is necessary to classify them according to the abstractness degree [2, 3] which takes into account the presence of a particular term in the thesaurus of a 6-7-year-old preschooler, 10-11-year-old, 15-16-year-old and 17-18-year-old school graduate. In this case, the method of counting significant words in definitions and the method of paired comparisons are used. The evaluated words are divided into 6-7 categories, the complexity of which takes values 1, 2, 4, 8, 16, 32, 64. Also, the complexity of mathematical concepts s_i is estimated by the method of definition analysis and it is found that s_i varies from 2-3 (add, multiply) to 100-200 (gradient, divergence, rotor). The resulting comparison base can be used to determine the complexity of other mathematical terms [3].

The proposed methodology for evaluating the didactic complexity of textual and formula information is as follows:

1. For each sub-paragraph to write out 10 key terms, the knowledge of which is necessary for understanding the text and is an indicator of the educational material assimilation. Keywords from all the sub-paragraphs of the first paragraph should be placed in a file T1.txt. To repeat all this for the other paragraphs by creating the files T2.txt, T3.txt, …

2. To collect all the selected key terms in a text file and to rank them basing on semantic complexity, dividing all physical terms and ordinary words into 6-7 categories according to the method discussed in [2]. To create a file Slovar.txt, which lists all terms (including double terms)

ISSN 1512-1801

with the indication of the difficulty category in numerical form. To take into account previously obtained estimates of the mathematical terms complexity [3].

3. Using a special program Slozhnost.pas that accesses the dictionary Slovar.txt to analyze files T1.txt, T2.txt, T3.txt ... and to determine the total amount of information contained in the key concepts from the various paragraphs, as well as the average information capacity of the key concepts $IFC_{Ti} = I_{Ti} / N_{Ti}$ for each $i-$th paragraph. The formula is applied: $I_T = n_1 s_1 + n_2 s_2 + ... + n_r s_r + ...$, where n_r is the number of uses of the $r-$th term, and s_r is its complexity.

4. For each paragraph, to estimate the proportion ε_i of mathematical and other formulas in relation to the total volume of the text. To do this, read the formulas, naming the values included in it, for example $\rho = m / V$, – «density is equal to the ratio of mass to volume».

5. For each sub-paragraph, to write out 5 key physical or mathematical terms of maximum complexity used in mathematical formulas. To create files for each paragraph F1.txt, F2.txt, F3.txt ... containing the keywords included in the formulas.

6. To evaluate the complexity of each term, to create a dictionary Slovar1.txt and use the program Slozhnost.pas for each paragraph to determine the total amount of information in keywords and the information folding coefficient (IFC) $IFC_{Fi} = I_{Fi} / N_{Fi}$, where N_{Fi} is the number of key concepts in the formulas of the $i-$th paragraph.

7. From different sections of the manual to randomly select sentences, to determine the average word length and the average number of syllables in words, to find the structural complexity and the complexity index of the sentence structure $S_{str} = C / C_0$.

8. For each section, to calculate the differential and integral didactic complexity

$$DDC_i = S_{str}((1-\varepsilon_i) \cdot IFC_{Ti} + \varepsilon_i \cdot IFC_{Fi}), \quad IDC_i = DDC_i \cdot V_i .$$

9. For each paragraph and each chapter, to calculate the IDC separately for textual and formula information:

$$IDC_{Ti} = S_{str}\, IFC_{Ti}(1-\varepsilon_i) \cdot V_i, \;\; IDC_{Fi} = S_{str}\, IFC_{Fi} \cdot \varepsilon_i \cdot V_i,$$

$$IDC_{Tk} = \sum_{j=1}^{N_k} IDC_{Tj}, \;\; IDC_{Fk} = \sum_{j=1}^{N_k} IDC_{Fj},$$

where j is the paragraph number of the $k-$th chapter ($k = 1$, 2, ..., 5), N_k – the quantity of paragraphs in the $k-$th chapter.

2. Results of the didactic complexity assessment of the physics course sections

The results of the DC assessment are presented in Table 1 which contains the columns: 1) the number of paragraph i; 2) the paragraph title; 3) the volume of the $i-$th paragraph in pages; 4) the semantic complexity of the key concepts I_{Ti}, measured with the help of the program; 5) the number of key concepts N_{Ti}; 6) the information folding coefficient $IFC_{Ti} = I_{Ti} / N_{Ti}$ in keywords selected from the $i-$th paragraph; 7) the semantic complexity of the key physical and mathematical concepts I_{Fi} that make up the formulas; 8) the number of key concepts in the formulas N_{Fi}; 9) the information folding coefficient $IFC_{Fi} = I_{Fi} / N_{Fi}$ in the key concepts included in the formulas; 10) the proportion of formulas in the text ε; 11) differential didactic complexity DDC_i; 12) integral didactic complexity IDC_i. For the analyzed textbook, the average word length in syllables is D_W = 3.17, the average sentence length is D_S =11.9 words, the indicator of the sentence structure complexity is S_{Str} =1,3.

Table 1 shows the DDC and IDC for each section and paragraph; basing on it, the histograms in Fig. 1 and 2 are constructed. From Fig. 1.1 it is seen that the sections «Electrodynamics» (24.7) and «Atomic and Nuclear Physics» (23.5) have the largest DDC (which is proportional to the average IFC of key concepts and shows the difficulty of text understanding), then «Molecular Physics and Thermodynamics» (19.7), «Mechanics» (18.3), «Optics» (17.6). Since Chapter 3 «Electrodynamics» has the largest volume (73 pages), its IDC, which is proportional to the total amount of semantic information, is also maximum (Fig. 1.2). It is followed by: 1. «Mechanics» (1040), 5. «Atomic and Nuclear Physics» (800), 4. «Optics» (800) and 2. «Molecular Physics and Thermodynamics» (690). Chapter 3 «Electrodynamics» is characterized by a high IFC for text (16.8) and formulas (23.9) and a large proportion of formulas; therefore, it has the highest DDC (24.7). In section 5 «Atomic and Nuclear Physics», the IFC of the text is high (18.2), but the IFC for formulas is not large (9.9), the fraction of formulas ε is small, the text volume is small (34 pages), so with a high DDC (23.5), the integral IDC is not high (798).

Table 1. Results of the assessment of the didactic complexity of the physics sections.

i	Paragraph name	V	I_T	N_T	IFC_T	I_F	N_F	IFC_F	ε	DDC	IDC
1	1.1. Kinematics	10	360	30	12,0	360	15	24,0	0,33	20,7	207
2	1.2. Dynamics	18,3	952	70	13,6	525	35	15,0	0,27	18,2	333
3	1.3. Oscillations & waves	25	1090	90	12,1	974	42	23,2	0,2	18,6	466
4	1.4. Hydro & aeromechanics	3,5	148	20	7,4	135	10	13,5	0,09	10,3	36
General (Mechanics)		56,8	2550	210	12,1	1994	102	19,5	--	18,3	1042
5	2.1. Molec.-kinetic theory	18	1094	80	13,7	557	45	12,4	0,16	17,5	315
6	2.2. Thermodynamics	17,2	882	70	12,6	630	22	28,6	0,27	22,0	379
General (Mol. phys. & thermod.)		35,2	1976	150	13,2	1187	67	17,7	--	19,7	694
7	3.1. Electrostatics	21,2	1301	90	14,5	715	40	17,9	0,23	19,8	420
8	3.2. Direct electric current	6,5	650	40	16,3	435	20	21,8	0,26	23,0	149
9	3.3. Magnetism	31,5	2044	100	20,4	1166	43	27,1	0,17	28,0	883
10	3.4. EM oscillations & waves	13,5	721	50	14,4	748	25	29,9	0,34	25,6	346
General (Electrodynamics)		72,7	4716	280	16,8	3064	128	23,9	--	24,7	1799
11	4.1. Geometric optics	6,5	364	30	12,1	178	14	12,7	0,09	15,8	103
12	4.2. Wave optics	25,9	792	69	11,5	445	45	9,9	0,15	14,6	378
13	4.3. Quantum optics	12,8	757	40	18,9	385	20	19,3	0,16	24,7	316
General (Optics)		45,2	1913	139	13,8	1008	79	12,8	--	17,6	797
14	5.1. Atomic structure	7	456	30	15,2	106	15	7,1	0,22	17,4	122
15	5.2. Wave properties of mat.	7,5	280	20	14,0	78	10	7,8	0,03	18,0	135
16	5.3. Atomic nucleus	12,6	752	40	18,8	191	13	14,7	0,08	24,0	303
17	5.4. Elementary particles	4,5	336	11	30,5	0	0	0,0	0	39,7	179
18	5.5. Cosmology	2,4	192	10	19,2	0	0	0,0	0	25,0	60
General (Atomic & nuclear phys.)		34	2016	111	18,2	375	38	9,9	--	23,5	798

ISSN 1512-1801

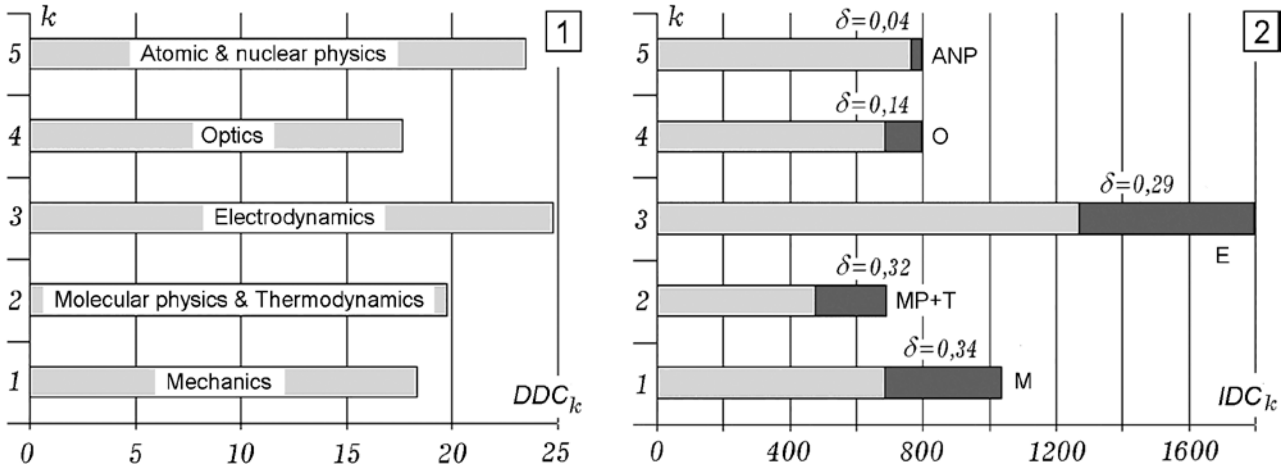

Fig. 1. DDC and IDC indicators for individual sections of physics in [9].

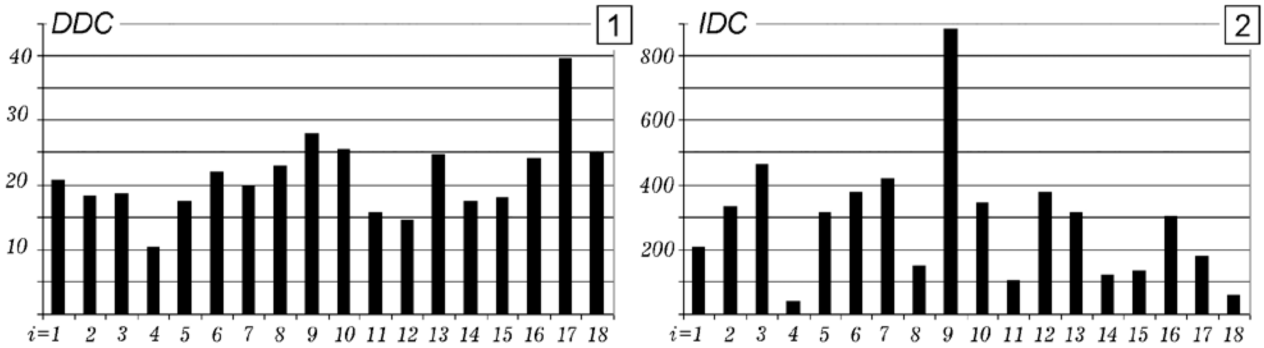

Fig. 2. Indicators of DDC and IDC in individual paragraphs of the textbook [9].

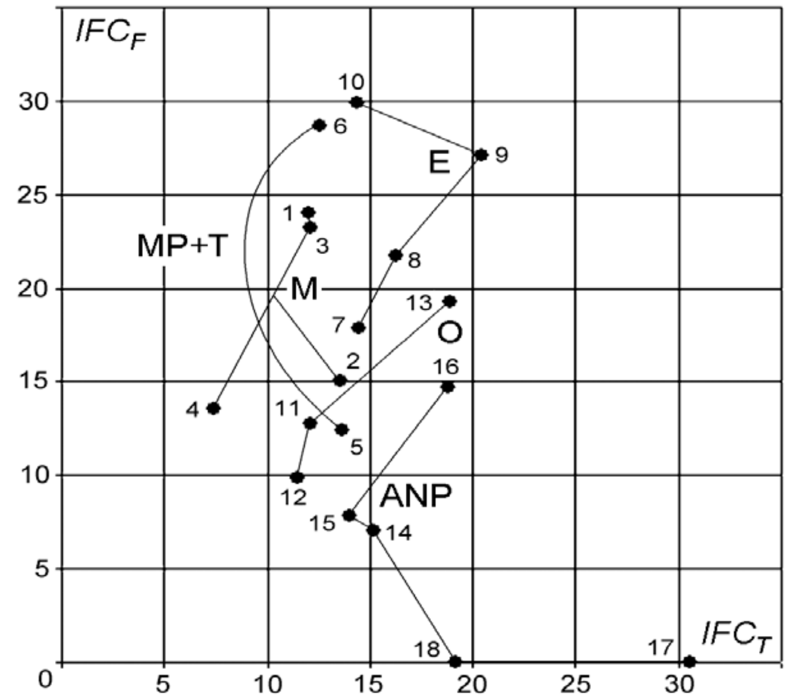

Fig. 3. Paragraphs distribution of the textbook [9] within space «textual IDC_T vs. formula IDC_F ».

ISSN 1512-1801

From table 1 it is possible to calculate the IDC of the textual and formula components for each chapter of the manual [9]; the resulting distribution is shown in Fig.1.2. The share of the formula IDC_{Fi} in the total complexity IDC of each physics section $\delta_k = IDC_{Fk} /(IDC_{Tk} + IDC_{Fk})$ ($k = 1, 2, \ldots, 5$) is shown next to it. It turns out that in the manual [9] the most complex in mathematical terms is Chapter 3 «Electrodynamics» ($IDC_{F3} =$ 528), which uses the concepts «circulation», «divergence», «rotor»; the least complex is Chapter 5 «Atomic and Nuclear Physics» ($IDC_{F5} = 36$).

As follows from Fig. 2.1 and 2.2, the DDC indicators of some paragraphs of the textbook [9] differ 4 times (paragraphs 4 and 17), and the IDC – 24 times (paragraphs 4 and 9). The highest DDC is in paragraph 17 «Elementary particles» (key concepts: neutrino, positron, meson, antiparticle, hadron, lepton, annihilation, quark, baryon, hyperon, gluon), the smallest – paragraph 4 «Hydro- and aeromechanics» (also it corresponds to the smallest IDC). Paragraph 9 «Magnetism» has the maximum volume (32 pages) and the largest IDC (883).

Let us consider the paragraphs distribution for the textbook [9] within the space «textual IDC_T – formula IDC_F» (Fig. 3). The textual IDC_T for most paragraphs is in the range from 7 to 20, and the formula IDC_F is in the range from 7 to 30. Since the points corresponding to the paragraphs are located randomly, and do not line up, it can be argued that the IDC for textual and formula information are weakly correlated with each other and are independent characteristics of educational texts in physics. From Fig. 3 it is seen that paragraph 17 has a high IDC_T, but there are no formulas; paragraphs 1, 3, 6, 8, 9, 10 have the high IDC_F, i.e. the mathematical models discussed in them is difficult to understand.

Summary

The article proposes a method for assessing the differential and integral didactic complexity of educational texts in physics. It consists in identifying the key concepts in the text, formulas and determining their average information capacity (IDC). When assessing the semantic complexity of scientific concepts, the quantity of words in the concepts definitions, their entry into the linguistic picture of the world, and the degree of abstraction are taken into account. In addition, the structural complexity of the text, depending on the average length of words and sentences, is taken into account. As the analysis result of the textbook [9] it is found that the sections «Electrodynamics» and «Atomic and Nuclear Physics» are the most difficult to understand; their DDC, which shows the complexity of elementary statements, is high. The total amount of time and effort required to master all the educational material is characterized by IDC; the section «Electrodynamics» has significantly more IDC than other sections of physics, which is caused by the large volume and use of concepts with a high degree of abstraction. Of the all paragraphs, p. 9 «Magnetism» has the largest IDC; this is due to the objective complexity of magnetic phenomena. Using the proposed method of key concepts helps to assess the didactic complexity of various educational texts and make their comparison.

References

1. Mikk Ja.A. Optimizacija slozhnosti uchebnogo teksta: V pomoshh' avtoram i redaktoram. Moscow: Prosve-shhenie, 1981. 119 p.
2. Mayer R.V. Metod ocenki didakticheskoj slozhnosti shkol'nyh uchebnikov po razlichnym discip-linam // Standards and monitoring in education. № 5. 2020. pp. 14–19.
3. Mayer R.V. Nekotorye aspekty ocenivanija kognitivnoj slozhnosti matematicheskih ponjatij // Bulletin of the Omsk State Pedagogical University. Humanitarian Studies. 2020, № 3 (28). pp. 122-126.

ISSN 1512-1801

4. Velichkovskij B.M. Kognitivnaja nauka: Osnovy psihologii poznanija. V 2 t. T. 1. Moscow: Smysl: Publishing center «Academy», 2006. 448 p.
5. Kuznecov I.P. Mehanizmy obrabotki semanticheskoj informacii. Moscow: Nauka, 1978. 174 p.
6. Chernjahovskaja L.A. Smyslovaja struktura teksta i ee edinicy // Questions of linguistics. 1983. № 6. pp. 117–126.
7. Lukov Val.A. Lukov Vl.A. Metodologija tezaurusnogo podhoda: strategija ponimanija // Knowledge. Understanding. Skill. 2014. № 1. pp. 18–35.
8. White M.D., Marsh E.E. Content analysis: A flexible methodology // Library trends. 2006. Vol. 55. № 1. pp. 22–45.
9. Batura T.V. Matematicheskaja lingvistika i avtomaticheskaja obrabotka tekstov: ucheb. posobie. Novosib. gos. un-t. Novosibirsk: RIC NGU, 2016. 166 p.
10. Kisel'nikov A.S. K probleme harakteristik teksta: chitabel'nost', ponjatnost', slozhnost', trudnost' // Philological sciences. Questions of theory and practice. 2015. № 11 (53). Ch. 2. pp. 79–84.
11. Marchuk Ju.N. Komp'juternaja lingvistika: ucheb. posobie. Moscow: AST; Vostok–Zapad, 2007. 317 p.

Article received 2022-10-25