Analysis of High School English Curriculum Materials through Rasch Measurement Model and Maxqda

Veli Batdı[1]
Kilis 7 Aralık University

Şenel Elaldı[2]
Cumhuriyet University

Abstract

The purpose of the study is to analyze high school English curriculum materials (ECM) through FACETS analysis and MAXQDA-11 programs. The mixed methods approach, both quantitative and qualitative methods, were used in three samples including English teachers in Elazig during the 2014–2015 academic year. While the quantitative phase of the study involved the Rasch model and SPSS analysis, the qualitative phase included a case study research method. The quantitative data collection tool was the "High School English Curriculum Material Evaluation Form" developed by the researchers. Qualitative data were collected through an interview. According to the Rasch analysis results, while the material Headway coded as HEA has the highest quality, the materials coded A1.2 and A2.1 have the poorest quality. In terms of items, the most difficult item was G-M28 and for judges, J3 has the most severe while J13 has the most lenient behavior. The qualitative analysis indicated that among the themes, "deficiencies" had the most dominant dispersing of teacher comments while the theme "what to do" was not notable to every teacher. It is suggested that the balanced distribution of four language skills in ECM should be improved to achieve the objectives of the curriculum.

Keywords

High school ECM • Evaluation • Teachers' views • Rasch measurement model • MAXQDA-11 program • Mixed methods

1 Department of Educational Sciences, Kilis 7 Aralık University, Kilis Turkey. Email: veb_27@hotmail.com

2 **Correspondence to:** Şenel Elaldı (PhD), Department of Foreign Languages, Cumhuriyet University, Sivas 58140 Turkey. Email: snlelaldi@gmail.com

edam

One of the essential parts of language teaching is related to what materials are used in the teaching process, and therefore, a wide variety of teaching aids is necessary in foreign language classrooms. According to Tomlinson (2001, p. 66), the term materials encompasses "anything used to facilitate the learning of language" and may include linguistic, visual, auditory, or kinesthetic forms, which can be presented in print, through live performance or display, on a cassette, CD ROM, DVD, or the Internet. Since "materials are not simply the everyday tools of the language teachers; they are an embodiment of the aims and methods of the particular teaching/learning situation" (Zhang, 2007, p. 28), the success of the language program depends on the language teaching materials (Pakkan, 1997). "In case of inexperienced teachers, materials may also serve as a form of teacher training—they provide ideas on how to plan and teach lessons" (Richards, 2001, p. 251). Although grammar books and dictionaries were the only language teaching materials of used by teachers in the past, today, there is a great variety of language teaching materials in the market (Crystal, 1987), ranging from textbooks, supplementary texts, pictures, audio-tapes, sliders, transparencies, video tapes, smart boards to the Internet. However, despite a very wide range of commercially available materials, many teachers do not give enough importance to use them effectively and prefer to "produce their own materials for classroom use" (Howard & Major, 2005, p. 101). McGrath (2013, p. 7) differentiates foreign language materials into four categories: materials designed for language learning and teaching (e.g., textbooks, worksheets, computer software); authentic materials selected specially and exploited for teaching purposes by the classroom teacher (e.g., off-air recordings, newspaper articles); teacher-written materials; and learner-generated materials. Among all of the materials used in language classrooms, course books have been the most preferred instructional material throughout the world (Arıkan, 2008). Cunningsworth (1995) states that course books are the best resource in achieving aims and objectives regarding learner needs.

Today's language teaching throughout the world could not take place without the extensive use of commercial materials (Richards, 2005). Since teachers may lack time and expertise to design a coherent program of work, course books are designed as "the theory, practice, activities, explanations, text, visuals, content, formats, and all other elements that contribute to the finished product" (Byrd, 1995, p. 8). Furthermore, course books enable a learner to preview or review the progress about what done in class (Harmer, 2001) by reducing time needed for lesson preparation, and especially for teachers who are untrained or inexperienced, course books provide methodological support (Richards, 2001). In addition, teachers who are teaching full-time find course books invaluable because they do not have enough time to create original lessons for every class (McGrath, 2013). In spite of widespread popularity of course books, their evaluation as classroom materials remains in under-researched (Arıkan, 2009). Hutchinson and Waters (1987) suggest that materials cannot be evaluated as good

or bad: "While evaluating the materials, we try to judge the fitness of something for required purpose" (p. 96). Therefore, it is important to realize that no commercial course book is a perfect fit for a language program (Richards, 2005; Savignon, 1997). One of the prominent complaints about commercial course books is related to the transmission of cultural information of English speaking countries. Cultural information is conveyed by means, such as the use of color, photographs, cartoons, magazine-style formats as well as through the words on the page (Harmer, 2001). However, global course books produced for international market expose the debate on the role of English as a lingua franca or international language, which enables the transmission of "Western" values, a form of cultural imperialism, by the majority of learners studying English outside English-speaking countries and therefore, it is a complex issue and "how this might translate into textbooks is yet uncertain" (McGrath, 2013, p .9). Although global textbooks may have been the focus for such criticism, debates may also be explicit for 'national' textbooks which deliberately promote national values and culture ideologically for social cohesion (Lund & Zoughby, 2007). Another common criticism is that "textbooks merely grow from and imitate other textbooks and do not admit the winds of change from research, methodological, experimentation, or classroom feedback" (Sheldon, 1988, p. 239). On the other hand, "many teachers are bound by a mandated curriculum defining the content, skills, and values to be taught" (Howard & Major, 2005, p. 103). This situation may lead many teachers to follow course books slavishly to fulfill objectives outlined by a curriculum for learners. Therefore, "teachers need training and experience in adapting and modifying textbooks as well as using authentic materials and in creating their own materials" (Richards, 2001, p. 16).

"Whether the teacher uses a textbook, institutionally - prepared materials, or his/her own materials, instructional materials generally serve as the basis for much of the language input learners receive and the language practice that occurs in the classroom" (Richards, 2001, p. 251). Although modern course book packages contain "a wealth of extra material" (Harmer, 2001, p. 7), teachers are also material developers and ultimately responsible for the materials that their learners use (Tomlinson, 2001). In material selection, teachers need to have the ability to: 1) "select appropriately from what is available, 2) be creative with what is available, 3) modify activities to suit learners' needs, and 4) supplement by providing extra activities" (Dudley-Evans & St John, 1998, p. 173).

The Purpose of the Research

The aim of the present study is to analyze high school English curriculum materials (ECM) through the multi-faceted Rasch measurement model, SPSS 17 program and MAXQDA-11 program. Therefore, to fulfill the purpose of this study, the following sub-aims have been established:

Within the scope of the Rasch measurement model: (1) perform a general analysis of views on high school ECM, (2) analyze perceptions of the judges with regard to their severity or leniency, (3) analyze the items' difficulty used in the questionnaire to evaluate high school ECM, (4) analyze any bias of judges.

Within the scope of SPSS program: (5) record the participants' views regarding the efficiency level of ECM

Within the scope of Maxqada program (qualitative): (6) determine the characteristics of high school ECM, (7) obtain a sense of the views of participants regarding the relationship between high school ECM and the English Curriculum.

Method

A mixed methods approach was applied in three samples (n = 36; n = 23; and n = 10). Methodologists remarked that by combining quantitative and qualitative research, studies can maximize the strengths of each approach and develop more complete and complementary understandings (Creswell & Plano Clark, 2011). Sequential explanatory mixed methods design, which comprises two phases - first collecting and analyzing quantitative and then qualitative data- within one study (Tashakkori & Teddlie, 2003) was used to capture the views of the teachers on English course books of high schools. First, quantitative and qualitative data were collected simultaneously and then, the results were combined after the analyses for each set of data was complete (Hanson, Creswell, Clark, Petska, & Creswell, 2005).

While the quantitative phase of the study involved the Rasch measurement model and SPSS analysis, the qualitative phase included a case study research method. A comparative interpretation between the results obtained from Rasch measurement model on course materials and SPSS analysis was aimed to be in line with the level of participation. The Rasch model clarifies how a person's performance based on a specific trait predicts that person's response as right or wrong and supplies precious data for the development, modification, and monitoring of valid measurement (Boone & Scantlebury, 2006). Case studies relying on "multiple sources of evidence," inquire "a contemporary phenomenon within its real-life context when the boundaries between phenomenon and context are not clearly evident" (Yin, 1994, p. 13). In addition, in case studies describing the impact of program, it is intended to provide inference about reasons for success or failure (Davey, 1991). In this context, within the scope of this study, it was aimed to analyze high school ECM in terms of teachers' views.

Participants

The study sample was comprised of high school English teachers from a variety of high schools and private teaching institutions in Elazig city center/Turkey during the 2014–2015 academic year. Participants for the Rasch measurement model totaled 23; for SPSS analysis 36; and for MAXQDA (qualitative analysis) 10. Further, the teachers had prior experience with the materials included in this study, and therefore, were familiar with them.

The materials used in the study were:

A. Yes You Can (Prepared by the Ministry of Education) coded as;

A1.1 (Ertürk, Ateş, Alkan, & Albayrak, 2014)

A1.2 (Baydar Ertopcu, İnci, Özbıçakçı Samur, & Yalçın, 2014)

A2.1 (Ertürk, Ateş, & Alkan, 2014)

A2.2 (Baydar Ertopcu, İnci, & Özbıçakçı Samur, 2014)

B. Prepared by Turkish authors and coded as;

EB1: English Break A1 (Bektaş & Tekir, 2014)

EB2: English Break A2 (Arıkan, Saraç, Köse, Tekir, & Köse, 2011)

C. Prepared by foreign authors and coded as;

SOL: Solution (Falla & Davies, 2012)

HEA: Headway (Soars & Soars, 2011)

Data Collection

In this study, a High School English Curriculum Material Evaluation Form was prepared to collect the quantitative data (Appendix-1) in light of the review of literature and experts (1 Associate Professor of English; 1 Assistant Professor of English; and 2 English lecturers, of which 2 have a Master's degree and 1 is a PhD student; 2 Turkish teachers; and 3 English teachers). The questionnaire, which was developed with the help of Batdı's (2010) scale that he had used in his MA thesis, is a 5-point Likert-type scale with five options ranging from "Strongly Disagree," to "Strongly Agree." The 28 items' content validity ratios (CVRs) related to the evaluation of high school ECM have been presented in Appendix-1. The content validity index (CVI) value for each item was computed separately. Experts were asked to score the relevance of each item as 1=not relevant, 2=somewhat relevant, 3=quite relevant, 4=highly relevant (Davis, 1992). Next, the CVI for each item was computed as the number of

experts giving a scoring of either 3 or 4, divided by the number of experts. A CVI of .80 is considered an acceptable value for good content validity (Yurdugül, 2005). The content validity indices (CVIs) of the items were found to be 0.88, namely, greater than the 0.80 content validity criterion at a level of statistical significance of $p < .05$.

Qualitative data were collected through an interview form developed by the researchers based on the review of literature and expert opinions and consists of open-ended questions. Questions in this form are intended to explore the teachers' views on high school ECM. In analyzing the text data, Qualitative Data Analysis Software, MAXQDA, was used. MAXQDA, as a text search tool, allows "for automated searches of text for words, phrases, and co-occurring themes with more accuracy and time efficiency than hand sorting and counting" (D'Andrea, Waters, & Rudd, 2011, p. 49). This technology also enables "researchers to interrogate the data set, look for co-occurrences of codes or themes, relationships between codes, and to play with ideas in an exploratory fashion" (Lewins & Silver, 2008, p. 11).

Data Analysis

When all data were collected, a sequential analysis of the data was conducted (Tashakkori & Teddlie, 2003). First, both quantitative and qualitative data were analyzed separately and then the findings were compared in order to determine if both sets of data indicated the same overall views.

Analysis of the quantitative data. In the analysis of quantitative data, FACETS analysis program, the Rasch measurement model identified by Linacre (1993, p. 2008), was used. Three facets of the study according to the Rasch measurement model were: 1) Judges, 23 high school English teachers, 2) Items related to high school ECM (28 items), and 3) ECM (A1.1, A1.2, A2.1, A2.2, EB1, EB2, HEA, and SOL). In addition, the arithmetic means of items and factors depending on the views of the participants were calculated through SPSS analysis.

Analysis of the qualitative data. The analysis process began with the examination of the transcribed text in MAXQDA with the aim of identifying categories of statements reflecting teachers' views on English curriculum materials used in high schools. A thematic analysis, which is a synthesizing strategy used as part of the meaning-making process of case study was used (Mills, Durepos, & Wiebe, 2010).

Analysis began with reading the transcribed participant comments from beginning to end two times by the researchers in order to get an overall feel about the teachers' views on English curriculum materials (Cohen, Kahn, & Steeves, 2000) and obtain several tentative categories. Particular words and phrases were assigned into categories after reading over the comments independently. With a final reading, four models

within two main headings, *The views of the participants about the characteristics of ECM* and *The views of the participants about the relationship between ECM and English curriculum (9th, 10th, 11th, and 12th grades)* were identified as unanimously encapsulating the teachers' views. Following this, the transcribed document was loaded into MAXQDA for analysis.

Results

The results of the study obtained from qualitative and quantitative applications were presented separately.

Results regarding the Rasch Measurement Model and SPSS Data

The views of high school English teachers who participated in the study on ECM were analyzed within the framework of evaluation forms via the multi-faceted Rasch model (Lunz & Linacre, 1998). Data obtained from three facets were specified for this study: (a) A1.1, A1.2, A2.1, A2.2, EB1, EB2, HEA, and SOL curriculum materials, (b) Leniency/severity of judges, and (c) suitability of items. Data calibration map related to the facets was given in Figure 1.
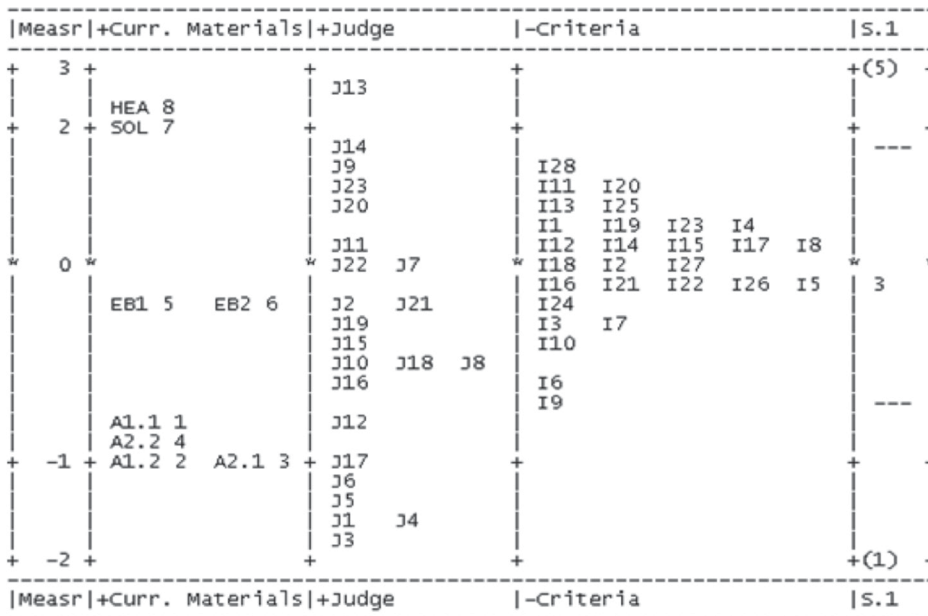
```
-------------------------------------------------------------------------------
|Measr|+Curr. Materials|+Judge        |-Criteria                    |S.1   |
-------------------------------------------------------------------------------
+   3 +                    +           +                            +(5)   +
|     |                    |  J13      |                            |      |
|     | HEA 8              |           |                            |      |
+   2 + SOL 7              +           +                            +      +
|     |                    |  J14      |                            | ---  |
|     |                    |  J9       | I28                        |      |
|     |                    |  J23      | I11  I20                   |      |
|     |                    |  J20      | I13  I25                   |      |
|     |                    |           | I1   I19  I23  I4          |      |
|     |                    |  J11      | I12  I14  I15  I17  I8      |      |
*   0 *                    *  J22  J7  * I18  I2   I27              *      *
|     |                    |           | I16  I21  I22  I26  I5     | 3    |
|     | EB1 5      EB2 6   |  J2   J21 | I24                        |      |
|     |                    |  J19      | I3   I7                    |      |
|     |                    |  J15      | I10                        |      |
|     |                    |  J10  J18 J8|                          |      |
|     |                    |  J16      | I6                         |      |
|     |                    |           | I9                         | ---  |
|     | A1.1 1             |  J12      |                            |      |
|     | A2.2 4             |           |                            |      |
+  -1 + A1.2 2    A2.1 3   +  J17      +                            +      +
|     |                    |  J6       |                            |      |
|     |                    |  J5       |                            |      |
|     |                    |  J1   J4  |                            |      |
|     |                    |  J3       |                            |      |
+  -2 +                    +           +                            +(1)   +
-------------------------------------------------------------------------------
|Measr|+Curr. Materials|+Judge        |-Criteria                    |S.1   |
-------------------------------------------------------------------------------
```

Figure 1. Data calibration map.

In Figure 1, curriculum materials, judges, and items were given separately. According to the "curriculum materials" column, while the material coded HEA has the highest quality, the materials coded A1.2 and A2.1 have the poorest quality. On the other hand,

in the "judges" column, J(udge)3 has the most severe and J13 has the most lenient behavior. When the column in which items used to evaluate curriculum materials is examined, the most difficult to answer item is G-I28 - *Questions are appropriate for taxonomy of cognitive domains (remembering, understanding, comprehension, etc.)*- while the easiest to answer item is C-I9 - *Objectives are associated with the content*.

English curriculum materials. Measurements related to high school ECM are presented in Table 1 comprehensively. According to Table 1, while the reliability co-efficient in the Rasch analysis is 1.00 which indicates a high reliability of related materials rankings, the separation index is 23.87. In line with the results, it is clear that there are statistically significant differences among English materials ($\chi2$ =4101.0, *sd* = 7, *p* = .00). The ranking of English curriculum materials from the most adequate to most inadequate is as follows: "HEA, SOL, EB2, EB1, A1.1, A2.2, A2.1, and A1.2."

Table 1

High School ECM' Measurement Report

```
| Obsvd   Obsvd Obsvd  Fair  |        Model | Infit       Outfit     |                   |
| Score   Count Average Avrage|Measure  S.E. |MnSq ZStd  MnSq ZStd | N Curr. Materials |
-------------------------------------------------------------------------------------------
|  2643    644    4.1   4.18|  2.09   .06 | 1.3   4   1.4   5 | 8 HEA 8           |
|  2617    644    4.1   4.14|  2.01   .06 | 1.5   7   1.5   7 | 7 SOL 7           |
|  1730    644    2.7   2.69| -.16    .05 | 0.8  -3   0.8  -3 | 6 EB2 6           |
|  1695    644    2.6   2.63| -.24    .05 | 0.7  -5   0.8  -5 | 5 EB1 5           |
|  1459    644    2.3   2.23| -.80    .05 | 0.9  -1   0.9  -1 | 1 A1.1 1          |
|  1418    644    2.2   2.16| -.91    .05 | 0.9  -1   0.9  -2 | 4 A2.2 4          |
|  1391    644    2.2   2.11| -.98    .05 | 0.9  -1   0.9  -1 | 3 A2.1 3          |
|  1376    644    2.1   2.09|-1.02    .05 | 1.0   0   0.9  -1 | 2 A1.2 2          |
-------------------------------------------------------------------------------------------
|  1791.1  644.0  2.8   2.78|  .00    .05 | 1.0  -0.2 1.0  -0.3| Mean (Count: 8)   |
|   500.3    0.0  0.8   0.82| 1.22    .00 | 0.2   4.0 0.3   4.1| S.D.              |
-------------------------------------------------------------------------------------------
RMSE (Model)   .05  Adj S.D.  1.22  Separation 23.87  Reliability 1.00
Fixed (all same) chi-square: 4101.0  d.f.: 7  significance: .00
Random (normal) chi-square: 7.0  d.f.: 6  significance: .32
-------------------------------------------------------------------------------------------
```

"Infit" and "outfit" statistical values related to the facets in the Rasch analysis are also given in Table 1. The quality control limit for infit and outfit values is between the range of 0.6–1.4 (Wright & Linacre, 1994). As given in Table 1, the value exceeding the determined limit for both indices (1.5) is the material coded SOL.

In the quantitative phase of the study, the arithmetic mean was computed on the basis of items and factors depending on the views of the participants through SPSS analysis (Appendix -2). When the analysis results of ECM were examined on the basis of factors, although the participants displayed the highest participation on the factor "B" belonging the material A1.2 at the level of *Mostly Disagree* ($\overline{X}$ = 2.30±.96), they displayed the lowest participation on the factor "G" at the level of *Mostly Disagree* ($\overline{X}$ = 2.01±.80). In

addition, in the evaluation of the related materials, while the participant teachers displayed the highest participation on the factors "C" and "F" belonging to the material coded HEA at the level of *Mostly Agree* ($\overline{X}_C$ = 4.19±.47; $\overline{X}_F$ = 4.19±.54), they displayed the lowest participation on the factor "G" at the level of *Mostly Agree* ($\overline{X}$ = 3.99±.58). Although Rasch analysis has a special ranking, SPSS data revealed substantial differences between these rankings. When Appendix-2 is examined in detail, such differences are likely to be seen both in items and in comparison of related factors on the basis of material. This can be interpreted as it being very important to emphasize both curriculum materials and the differences of curriculums according to the views of the participants.

Analysis of judges. Table 2 presents information regarding the leniency/severity of judges regarding ECM. When considering the judges from the severest to the most lenient, J3 is the most severe and J13 is the most lenient. The calculated value includes the standard error (RMSE) associated with the judges' severity/ leniency. When this value is at 0.09, it indicates that standard error is quite low. The reliability co-efficient associated with the judges' scoring behaviors and calculated as 0.99 highlights judges' scoring behaviors which have been performed at a high reliability.

As represented in Table 2, the Judge Separation Index is 10.22 and the reliability co-efficient is 0.99. When the hypothesis "There are statistically significant differences among judges in terms of the degrees of severity/leniency" is tested by Chi-Square ($\chi2$ = 1893, 1 sd = 22, p = .00), the null hypothesis is rejected. In other words, it is emphasized that judges have shown statistically significant differences among themselves.

Table 2

Judges' Measurement Report

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Measure	Model S.E.	Infit MnSq	ZStd	Outfit MnSq	ZStd	Nu	Judge
996	224	4.4	4.57	2.73	.11	0.9	0	1.4	2	13	J13
753	224	3.4	3.46	.59	.08	0.8	-2	1.0	0	14	J14
738	224	3.3	3.39	.49	.08	1.2	1	1.3	3	9	J9
724	224	3.2	3.32	.39	.08	0.4	-8	0.5	-7	23	J23
710	224	3.2	3.25	.29	.08	0.7	-3	0.7	-3	20	J20
677	224	3.0	3.09	.06	.08	1.1	1	1.1	1	11	J11
669	224	3.0	3.05	.01	.08	0.9	0	0.9	0	22	J22
665	224	3.0	3.03	-.02	.08	1.3	3	1.3	2	7	J7
640	224	2.9	2.90	-.19	.08	0.6	-6	0.6	-5	21	J21
639	224	2.9	2.90	-.20	.08	0.9	-1	0.9	-1	2	J2
631	224	2.8	2.85	-.25	.08	1.2	2	1.3	3	19	J19
608	224	2.7	2.73	-.41	.08	1.2	1	1.1	1	15	J15
600	224	2.7	2.69	-.47	.08	0.8	-2	0.8	-2	10	J10
592	224	2.6	2.65	-.53	.08	1.0	0	1.0	0	18	J18
590	224	2.6	2.64	-.54	.08	0.6	-5	0.6	-5	8	J8
582	224	2.6	2.60	-.60	.08	1.1	1	1.1	1	16	J16
554	224	2.5	2.45	-.80	.09	1.8	7	1.8	7	12	J12
533	224	2.4	2.34	-.96	.09	1.3	3	1.3	2	17	J17
511	224	2.3	2.22	-1.13	.09	1.0	0	0.9	0	6	J6
499	224	2.2	2.15	-1.23	.09	0.9	-1	0.8	-1	5	J5
494	224	2.2	2.13	-1.27	.09	0.8	-2	0.8	-2	1	J1
487	224	2.2	2.09	-1.33	.09	1.2	2	1.1	1	4	J4
437	224	2.0	1.82	-1.77	.10	1.2	1	1.1	0	3	J3

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Measure	Model S.E.	Infit MnSq	ZStd	Outfit MnSq	ZStd	Nu	Scorer
623.0	224.0	2.8	2.80	-.31	.09	1.0	-0.5	1.0	-0.2		Mean (Count: 23)
116.0	0.0	0.5	0.58	.89	.01	0.3	3.4	0.3	3.3		S.D.

RMSE (Model) .09 Adj S.D. .89 Separation 10.22 Reliability .99
Fixed (all same) chi-square: 1893.1 d.f.: 22 significance: .00
Random (normal) chi-square: 21.9 d.f.: 21 significance: .40

As seen in Table 2, while the infit and outfit values of J23 and J12 judges fall outside the range of 0.6 – 1.4, the accepted value proposed by Bond and Fox (2007) and Wright et al. (1994), the infit and outfit values of the other 21 judges included in the study are within the acceptable range, and are therefore deemed suitable. Since the mean square values of infit and outfit belonging to J23 and J12 are higher than the expected values, the two judges are unlikely to have consistent scoring behaviors in the evaluation of ECM.

The analysis of items used to evaluate ECM. Table 3 presents information related to whether the items to measure ECM fit the purpose or not, in other words, statistics concerning the difficulty analysis of related items.

The items related to the English materials coded *G-I28, - Questions are appropriate for taxonomy of cognitive domains (remembering, understanding, comprehension, etc.) and the course book involving speaking and writing skills, vocabulary and grammar teaching-* and coded *E-I20 - different methods (antonym, synonym, etc.) have been used in teaching vocabulary* - are evaluated as the weakest.

In addition, the easiest items found by the participants belong to the content in the course book coded *C-I9 -Objectives are associated with the content-* and in the course book coded *B-I6 - The objectives of the book are appropriate for the general principles of the National Education*.

The Standard Error (RMSE) associated with the items' analyses used to evaluate ECM is 0.10, which is the low value in determining the quality. The standard deviation value corrected for estimation error has been calculated as 0.27, which is below the critical value of 1.0. While the Separation Index is 2.83 and the reliability co-efficient is 0.89. When the hypothesis - *There are statistically significant differences in terms of item difficulties indicating the quality of ECM-* is tested by Chi-Square ($\chi2 = 251.7$, $sd = 27$, $p = .00$), the null hypothesis is rejected. In other words, it is emphasized that items evaluate different characteristics belonging to the materials and have shown statistically significant differences.

When infit and outfit values related to facets are examined, except for the material coded A-M3 in which infit or outfit data exceeded their limit values, all the materials have an acceptable level of 1.5. This result indicates that almost all items are consistent with the evaluation of the materials and their infit and outfit mean squares are within the acceptable values. The results of the analysis of ECM (A1.1, A1.2, A2.1, A2.2, EB1, EB2, HEA, and SOL) through SPSS have been presented in Appendix-1. When the data which indicates the participation levels on the basis of items and factors are examined, the participants displayed the highest participation on the factor "A" belonging the materials HEA ($\overline{X} = 4.07\pm.70$) and SOL ($\overline{X} = 4.17\pm.70$) at the level of *Mostly Agree* while they displayed the lowest participation on the same factor belonging the materials A2.1 ($\overline{X} = 2.15\pm.82$) and A1.2 ($\overline{X} = 2.16\pm.98$) at the level of *Mostly Disagree*.

Table 3

The Analysis of Items Used to Evaluate ECM

```
| obsvd   obsvd  obsvd   Fair  |        Model | Infit       outfit              |
| Score   Count Average Avrage|Measure  S.E. |MnSq ZStd  MnSq ZStd | Nu Criteria |
-----------------------------------------------------------------------------------
   458     184    2.5    2.45|   .49   .10  | 1.0   0   1.0    0   | 28 G-I28
   463     184    2.5    2.48|   .45   .10  | 0.9  -1   0.8   -1   | 20 E-I20
   470     184    2.6    2.53|   .38   .10  | 0.9  -1   0.8   -1   | 11 C-I11
   481     184    2.6    2.61|   .28   .10  | 1.2   1   1.2    2   | 25 G-I25
   483     184    2.6    2.62|   .26   .10  | 0.9  -1   1.0    0   | 13 C-I13
   485     184    2.6    2.63|   .24   .10  | 0.8  -2   0.8   -1   | 19 E-I19
   487     184    2.6    2.65|   .22   .10  | 1.3   2   1.2    2   |  1 A-I1
   490     184    2.7    2.67|   .19   .10  | 0.9   0   0.9    0   |  4 B-I4
   490     184    2.7    2.67|   .19   .10  | 0.8  -2   0.8   -2   | 23 F-I23
   497     184    2.7    2.71|   .13   .10  | 0.8  -1   0.8   -1   | 14 C-I14
   500     184    2.7    2.73|   .10   .10  | 0.8  -2   0.9   -1   | 12 C-I12
   503     184    2.7    2.76|   .08   .10  | 0.8  -2   0.9   -1   | 17 D-I17
   505     184    2.7    2.77|   .06   .10  | 1.1   0   1.1    0   |  8 B-I8
   505     184    2.7    2.77|   .06   .10  | 1.0   0   1.0    0   | 15 D-I15
   511     184    2.8    2.81|   .00   .09  | 0.7  -2   0.8   -2   | 18 E-I18
   515     184    2.8    2.84|  -.03   .09  | 1.1   0   1.1    1   | 27 G-I27
   516     184    2.8    2.84|  -.04   .09  | 1.0   0   1.0    0   |  2 A-I2
   519     184    2.8    2.86|  -.07   .09  | 0.9   0   0.9    0   | 26 G-I26
   520     184    2.8    2.87|  -.08   .09  | 0.8  -1   0.9   -1   | 21 E-I21
   521     184    2.8    2.88|  -.09   .09  | 0.9   0   1.0    0   | 22 F-I22
   527     184    2.9    2.92|  -.14   .09  | 1.1   0   1.1    0   | 16 D-I16
   528     184    2.9    2.92|  -.15   .09  | 1.2   1   1.2    1   |  5 B-I5
   539     184    2.9    2.99|  -.25   .09  | 0.8  -1   0.9    0   | 24 F-I24
   544     184    3.0    3.03|  -.29   .09  | 1.2   1   1.2    2   |  7 B-I7
   548     184    3.0    3.05|  -.33   .09  | 1.6   5   1.5    4   |  3 A-I3
   551     184    3.0    3.07|  -.35   .09  | 1.0   0   1.1    0   | 10 C-I10
   582     184    3.2    3.27|  -.63   .09  | 1.2   2   1.4    2   |  6 B-I6
   591     184    3.2    3.33|  -.71   .09  | 1.0   0   1.0    0   |  9 C-I9
-----------------------------------------------------------------------------------
| obsvd   obsvd  obsvd   Fair  |        Model | Infit       outfit              |
| Score   Count Average Avrage|Measure  S.E. |MnSq ZStd  MnSq ZStd | Nu Criteria |
-----------------------------------------------------------------------------------
  511.8   184.0   2.8    2.81|   .00   .10  | 1.0  -0.3  1.0   0.1| Mean (Count: 28)
   31.7     0.0   0.2    0.21|   .29   .00  | 0.2   1.8  0.2   1.6| S.D.
-----------------------------------------------------------------------------------
RMSE (Model)   .10  Adj S.D.   .27  Separation  2.83  Reliability  .89
Fixed (all same) chi-square: 251.7  d.f.: 27  significance: .00
Random (normal) chi-square: 27.0  d.f.: 26  significance: .41
```

As a result of the analysis, the materials (Solution and Headway) prepared by foreign authors (Falla & Davies; Soars & Soars) have the highest participation level, *Mostly Agree*. As for the materials prepared by Turkish authors, the materials in factor "E" belonging to EB1 and EB2 ($\overline{X}_{EB1}$ = 2.44±.67; $\overline{X}_{EB2}$ = 2.56±.64) and EB1 ($\overline{X}_{EB1}$ = 2.58±.77) in factor "G" have the second highest participation level. It can be said that the level of participation belonging all the other factors is at *Partially Agree*.

When the arithmetic mean of English textbooks prepared by Ministry of Education on the basis of factor is examined, all the factors range between 2.01 and 2.36. The participation level of these books is *Mostly Disagree*.

Judges' bias interaction analysis. In Table 4, interaction analysis related to the views of the judges on ECM is presented. In Table 4, Z points change between 8.90 and -5.88, indicating that judges made extremely severe or lenient evaluations on ECM. In this context, J19 gave 81 points (Z = 8.90) and exhibited severe bias to the material coded SOL, but should have given 116 points. Similarly, instead of 127 points, J14 gave 102 points (Z = 7.73) for the material coded HEA and exhibited severe bias. In addition, J4 gave 74 points (Z = 5.55) instead of 100 points for the material coded SOL. J22 gave 58 points (Z = 5.11) instead of 82 points for the material coded EB2.

Table 4*

Interaction Analysis of High School English Curriculum Evaluated with Judges

```
|Obsvd    Exp.  Obsvd  Obs-Exp| Bias+ Model            |Infit Outfit|
|Score    Score Count  Average| Measure S.E.  Z-Score| MnSq  MnSq | Sq N Curr.Mat.measr Nu Score measr|
---------------------------------------------------------------------------------------------------------
|  81     116.4   28    -1.27|  1.96   .22    8.90  |  0.7   0.7 | 151 7 SOL 7   2.01 19 J19   -.25 |
|  74     107.9   28    -1.21|  1.75   .22    7.92  |  0.7   0.7 |  95 7 SOL 7   2.01 12 J12   -.80 |
| 102     127.4   28     -.91|  1.82   .24    7.73  |  0.5   0.5 | 112 8 HEA 8   2.09 14 J14    .59 |
|  77     109.2   28    -1.15|  1.68   .22    7.65  |  0.7   0.7 |  96 8 HEA 8   2.09 12 J12   -.80 |
|  74      98.6   28     -.88|  1.23   .22    5.55  |  0.7   0.7 |  31 7 SOL 7   2.01  4 J4   -1.33 |
| 109     126.4   28     -.62|  1.31   .25    5.29  |  0.8   0.9 |  72 8 HEA 8   2.09  9 J9    .49 |
|  58      82.0   28     -.86|  1.21   .24    5.11  |  0.3   0.3 | 174 6 EB2 6   -.16 22 J22   .01 |
.........................................................................................................
|  76      54.5   28      .77| -1.13   .22   -5.11  |  0.3   0.3 |  29 5 EB1 5   -.24  4 J4   -1.33 |
| 117      93.7   28      .83| -1.37   .27   -5.12  |  0.5   0.6 | 110 6 EB2 6   -.16 14 J14    .59 |
| 138     105.2   28     1.17| -3.75   .72   -5.17  |  1.0   1.1 | 135 7 SOL 7   2.01 17 J17   -.96 |
| 126     102.2   28      .85| -1.66   .31   -5.31  |  0.7   0.8 |  47 7 SOL 7   2.01  6 J6   -1.13 |
| 132     106.6   28      .91| -2.13   .39   -5.49  |  1.7   1.6 | 136 8 HEA 8   2.09 17 J17   -.96 |
| 126     100.1   28      .92| -1.78   .31   -5.68  |  0.8   0.9 |  32 8 HEA 8   2.09  4 J4   -1.33 |
|  96      68.8   28      .97| -1.34   .23   -5.88  |  0.5   0.5 | 169 1 A1.1 1  -.80 22 J22   .01 |
---------------------------------------------------------------------------------------------------------
|Obsvd    Exp.  Obsvd  Obs-Exp| Bias+ Model            |Infit Outfit|
|Score    Score Count  Average| Measure S.E.  Z-Score| MnSq  MnSq | Sq  N  Mat.    measr Nu Score measr|
---------------------------------------------------------------------------------------------------------
|  77.9    77.9  28.0    .00|   .07   .28     .09  |  0.8   0.8 | Mean (Count: 184)                |
|  28.9    26.3   0.0    .42|   .94   .12    2.80  |  0.8   0.4 | S.D.                             |
---------------------------------------------------------------------------------------------------------
Fixed (all = 0) chi-square: 1440.7  d.f.: 184  significance: .00
---------------------------------------------------------------------------------------------------------
```

Table 4*: The uppermost and lowermost data were excluded due to extensive table data.

Besides severe biases, lenient behaviors are also exhibited by the judges included in the study. For example, J22 gave 96 points (Z = -5.88) instead of 69 for the material coded A1.1 and exhibited lenient behaviors. Similarly, J4 gave 126 points (Z = -5.68) instead of 100 points for the material coded HEA and exhibited extremely lenient behaviors.

Results Related to Qualitative Data

Results obtained from the qualitative analysis of the study have been discussed under two titles: "The views of the participants about the characteristics of ECM" and "The views of the participants about the relationship between ECM and English curriculum for 9th, 10th, 11th and 12th grades." The views of the participants regarding the related materials (A1.1, A1.2, A2.1, A2.2, EB1, EB2, HEA, and SOL) used as basic materials/textbooks as well as supplementary books / resources in high school English curriculum and coded in different ways were included in the study. Their views on the relationship between these materials and the current curriculum were also analyzed and modeled (Figure 3 and 4). Four themes related to *the views of the participants about the characteristics of ECM* emerged from the comments as "current situation," "deficiencies," "suggestions," and "what to do." They were modeled as "*Current situation of ECM and what to do; Deficiencies of ECM and suggestions*" and illustrated in Figure 2 and 3 with related codes.

The other three emerging themes related to *the views of the participants about the relationship between ECM and English curriculum (9, 10, 11 and 12 grades)* were "current situation," "deficiencies," and "suggestions" which were modeled as "*Current situation and suggestions related to the relationship between ECM and English curriculum; Deficiencies based on Language skills related to the relationship between ECM and English curriculum*" and given in Figure 4 and 5. In addition, the distribution of each theme generated by the text document of participant views through MAXQDA was illustrated in Figure 6. The document portrait in MAXQDA displays data results and enables researchers visualize which factors played an important role at what point in their interviews. Each color in the document portrait depicts a different code and each dot represents a certain fraction of the coded segments (D'Andrea et al., 2011). In this study, the data were presented with different color themes and codes on a model. While blue tones indicated "current situation," green tones indicated "suggestions," red tones "deficiencies," and purple "what to do."
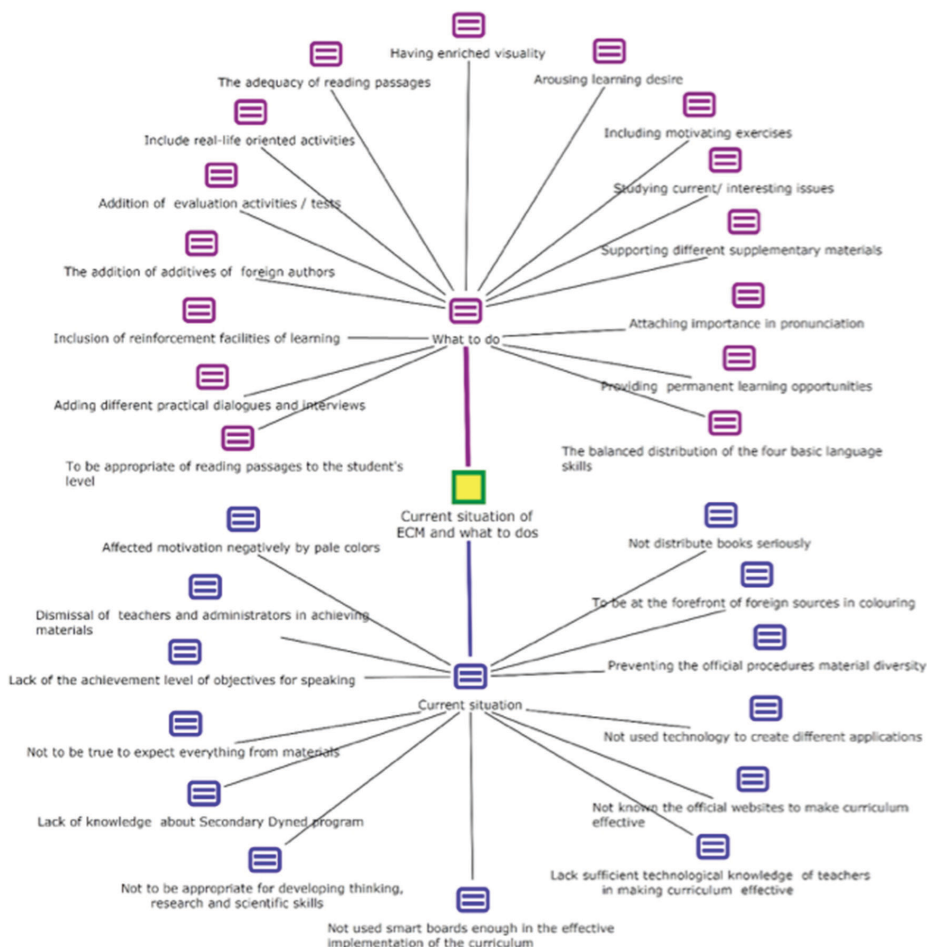
Figure 2. Current situation of ECM and what to do.

Results Related to the Characteristics of ECM

The views of the participants regarding different materials used in high school English curriculum were thematized as "current situation," "deficiencies," "suggestions," and "what to do" and modeled as A.1.1, A.1.2 under the title of "*The characteristics of ECM.*"

Current situation. The first theme emerging from the data related to the characteristics of ECM was identified by the blue dots in the document portrait. For example comments such as, *First, I want to indicate that there are great differences between the materials. Although curriculum is good at the distribution of objectives related to language skills and presenting them, I would say that the distribution of objectives in curriculum materials is not good enough and therefore, this poses a problem. In addition, diversification is not enough in visuality. When evaluated in terms of coloring according to foreign books or materials, pale colors can lead to negative effects on students' motivation* were coded in blue, and shown in the document portrait suggesting that teachers felt some materials were, at least to some extent, suitable for the current situation.

Deficiencies. The red dots represent teacher views about deficiencies related to the characteristics of ECM. Most of the codes included deficiencies of an application nature, for example, *Not include different applications in the development of social skills*, *Inability to achieve efficient learning environments, Not offer applications for the preparation for life,* and *Not offer applications for the preparation for life*. Comments, such as *I believe that diversification is needed for language skills. I think overall evaluation is necessary at times. In other words, I think there should be an overall evaluation including differences for the content of the previous units of chapters. This can lead to self-evaluation, controlling and monitoring of the objectives for both teachers and students*, were coded as a lack of overall evaluations. As can be seen in the document portrait, the dispersing of teacher comments related to deficiencies of curriculum materials was dominant. The deficiencies theme was not prominent to every teacher; however, it appeared that to those teachers who observed some necessities to diversify the use of curriculum materials in line with their beliefs about teaching, the deficiencies related to the commercial materials were in intensity ranking.

Suggestions. The green dots show comments made about improving the characteristics of ECM. Comments, such as *Textbooks are the most important factors in implementing curriculum effectively, and therefore, I would like to make some suggestions on textbooks. When evolving and changing technology has been taken into consideration, textbooks should be prepared in compliance with educational technology, particularly touch technology, curriculum, and different approaches. They should be supported by different web sites and different learning settings*, and,
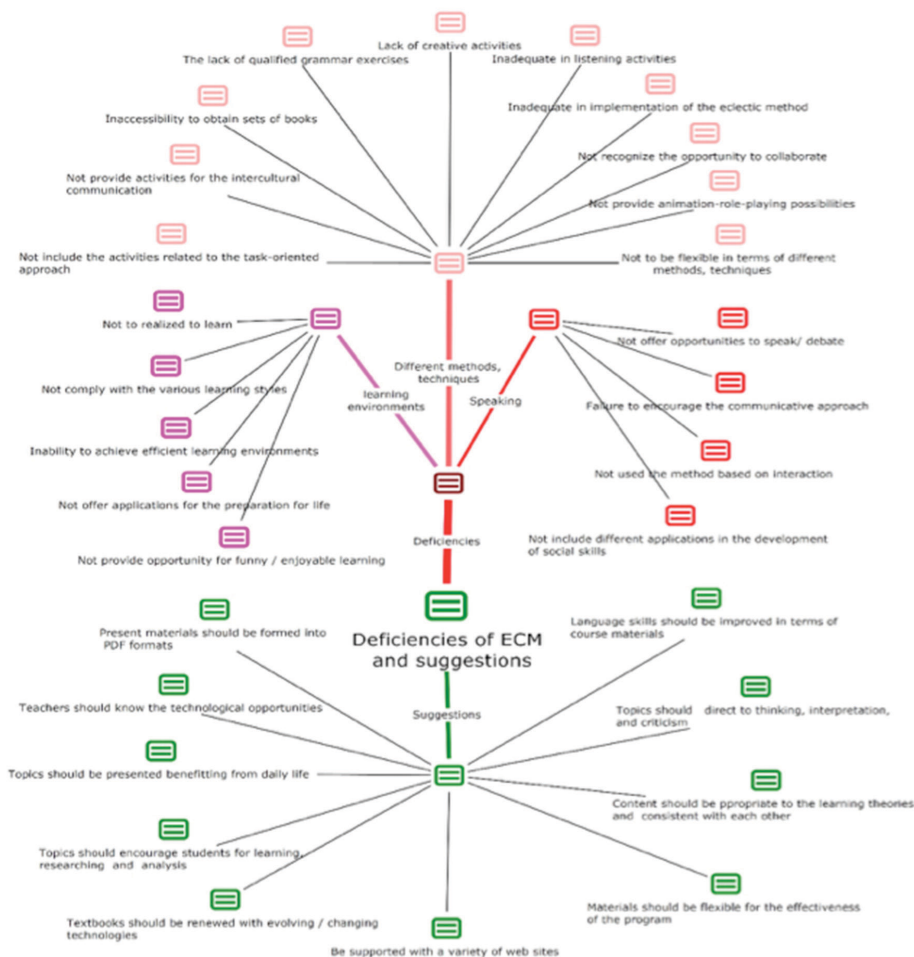
Figure 3. Deficiencies of ECM and suggestions.

individual differences stemming from different schools should be taken into account by textbooks….Moreover, textbooks should be able to meet students' needs. Social, economic and cultural integrity of subjects should be provided in textbooks and presented in a way that students would benefit in their daily lives, were coded as belonging to suggestions on textbooks. It was observed from the codes that most suggestions focused on integrating technology into teaching materials.

What to do. The purple dots indicate comments on what to do for effective English curriculum materials. Most codes were related to adding extra listening, speaking, and evaluation activities and a balanced distribution of the four basic language skills. As shown in the portrait, this theme was not notable to every teacher; however, some shared their views on what to do.
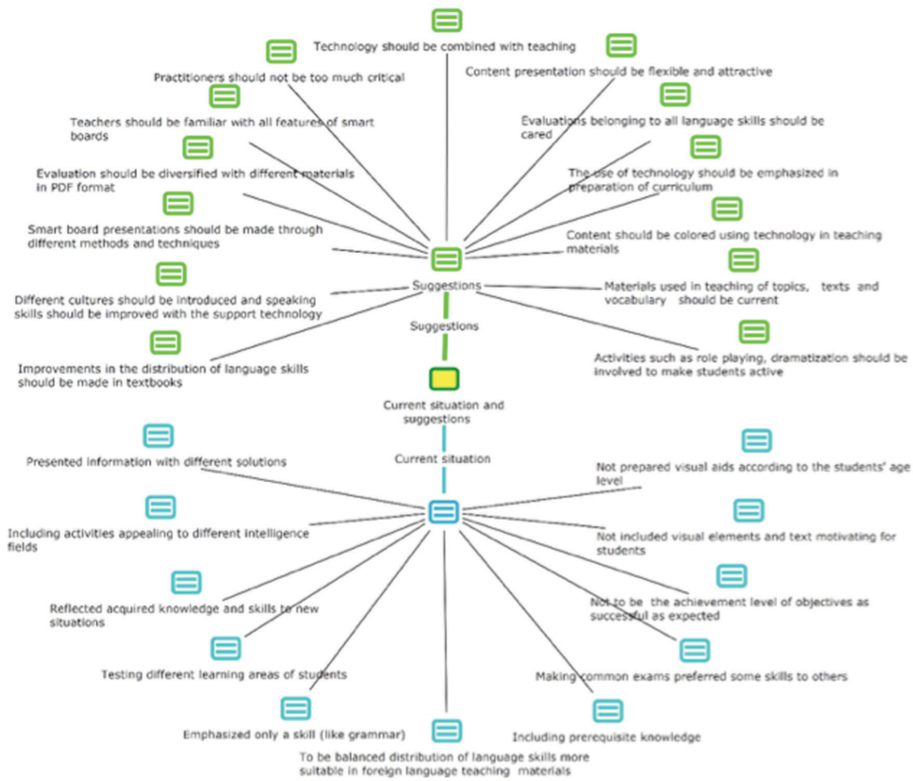
Figure 4. Current situation and suggestions related to the relationship between ECM and English curriculum.

Results Related to the Relationship between ECM and English Curriculum

The relationship between ECM and English curriculum according to the participants' views was modeled as B.1.1., B.1.2 under the relevant themes: "current situation," "deficiencies," and "suggestions."

Current situation. This theme was identified by the blue dots in the document portrait. For example, comments, such as *Frankly, I dare not to use different materials. Some explanations concerning any additional materials as "It could be possible, but you are the responsible" prevent the entry of additional materials. In other words, the presence of official procedures when needed leads to hesitation among teachers and a decrease in efficacy of practice,* and *Flexibility is very important in presenting a curriculum effectively. I think, while touch technology, such as smart boards, is applied in today's classrooms in terms of visuality, talking about a few books whether they are sufficient or not is unnecessary. A visual setting can be created easily by using this technology. It is not possible to expect everything from the materials. Moreover, many materials can be utilized in the same course through using the smart board. Textbook sets (CD, etc.) can be accessed from different sites and distributed*
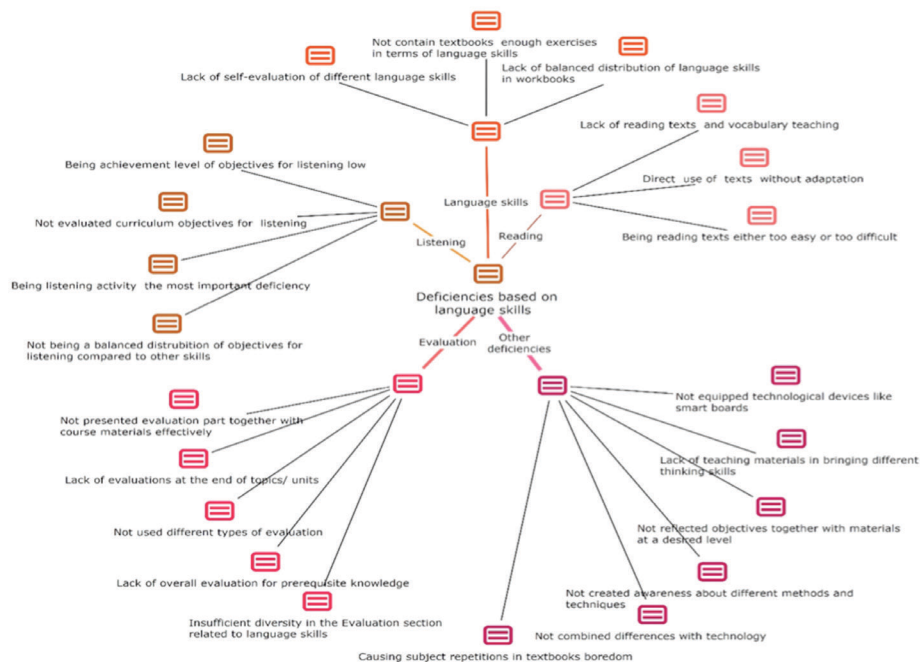
Figure 5. Deficiencies based on Language skills related to the relationship between ECM and English curriculum.

to all students and made presentations with a simple flash in diversified and colored settings. Even, there are official English sites focused on high school DynEd and online prepared by Ministry of Education and eba. We have signed an official letter to benefit them, but I wonder how many of our teacher friends have utilized them. Neither should be focused on books too much nor should be completely separate, were coded as preventing the official procedures from material diversity belonging to this theme.

Suggestions. The green dots represent comments made about the suggestions on the relationship between ECM and English curriculum. According to some comments, improvements in the distribution of language skills should be made in textbooks: *When curriculum-material relationship is evaluated in terms of objectives, the achievement of objectives seems not to be as successful as expected. Since balanced distribution of four language skills are more suitable in foreign language teaching materials, foreign books or ongoing series of books are better than National Education books in achieving the objectives of the curriculum.* Some comments were coded as current materials to be used in teaching of topics, texts and vocabulary: *Curriculum objectives are clear. Reducing them according to students' level or presenting them over students' level depends on the ability of practitioners. Objectives can be presented in accordance with student profile through technology. There are lots of classrooms in which students are better than teachers in the use of*

smart boards, but it must be opposite. Most of teachers are required to use smart boards in the best way. This situation affects curriculum materials and achievement level of all phases of curriculum indirectly. In other words, combining teaching with technology will be more effective and efficient in achieving objectives regardless of school type or different student profile.

Deficiencies. The red dots indicate teacher views regarding deficiencies stemming from the relationship between ECM and curriculum. Some teachers complained about deficiencies in evaluating curriculum objectives for listening: *Listening test is pretty compelling in common exams for both teachers and students and therefore, the curriculum's objectives for these skills lead to low level achievement of content. There is no balanced distribution between listening and other skills and this prevents ongoing teaching process. However, selected topics should be handled by the adjacency principle in every aspect and connected to each other. In short, the evaluation of curriculum objectives for listening is the most important problem.* Other teachers commented on inadequate reading texts in textbooks: *Reading is extremely important in vocabulary teaching. In order to improve students' reading, pronunciation, vocabulary knowledge, and grammar, reading texts should be emphasized to achieve the objectives of the curriculum related to these skills. Reading texts are inadequate in curriculum materials. However, resolving this issue will provide a positive contribution to gain objectives.* Some comments discussed the problem of not using different evaluation

Figure 6. The illustration of document portrait showing the themes and codes. Blue=current situation, Green=suggestions, Red=deficiencies, Purple=what to do.

types, for example, *I think the evaluation part of the curriculum cannot be presented effectively with curriculum materials. Insufficient evaluations at the end of topics or units are generally related to the measure of grammar levels, and therefore, failure to use the different types of evaluation is a serious problem.* Other comments focused on a lack of teaching materials to gain different thinking skills: *In addition, consistently engaging and motivating differences should be implemented in order to gain different thinking skills. This can be provided with versatile and possible means, such as smart boards. In language skills, curriculum materials can be said to be missing from reading, listening, and writing activities. Although the curriculum doesn't have the missing objectives in achieving these skills, it is perceived as missing due to the materials.* It can be said that all similar comments representing the negative aspects of curriculum in terms of material use for some teachers were coded as belonging to deficiencies theme and suggestions were made on improving curriculum materials.

In the software, it is likely that the structure of the study document can be seen in its integrity through the visual tools (D'Andrea et al., 2011). As shown in the document portrait of this study, the themes are of different intensities.

As can be seen in the document portrait above, most comments were associated with the "deficiencies of ECM and curriculum" which were shown in red. Comments, such as *lack of evaluation, lack of qualified grammar exercises, failure to encourage the communicative approach, causing subject repetitions in textbooks boredom,* represent the deficiencies of ECM and the English curriculum. The following theme asserted by the participants was seen to be "suggestions" that were colored in green. And the least intensity was observed in the theme of "what to do" stated by the participants with the aim to represent certain recommendations which was colored in purple. Some codes concerning the theme under consideration, such as *supporting different supplementary materials, including real-life oriented activities,* and *providing permanent learning opportunities,* can be given as examples.

Discussion

In this study, data related to high school ECM were analyzed using the multi-faceted Rasch measurement model and MAXQUDA-11. Three facets were specified as curriculum materials: leniency/severity of judges and suitability of items. According to results, while the material coded HEA has the highest quality, the materials coded A1.2 and A2.1 have the poorest quality among 8 ECM. In terms of Judges' bias on the evaluation of ECM, while J19 (for SOL), J14 (for HEA), J4 (for SOL) and J22 (for EB2) exhibited severe bias, J22 (for A1.1) and J4 (for HEA) exhibited extremely lenient behaviors. According to the items prepared to evaluate "curriculum materials," the most difficult item was G-M28 - *Questions are appropriate*

for taxonomy of cognitive domains (remembering, understanding, comprehension, etc.)- while the easiest item was C-M9 -*Objectives are associated with the content*. As for judges, J3 has the most severe and J13 has the most lenient behavior. When personal information of the judges is taken into consideration, the most lenient judge is male, graduated from Faculty of Education and has 10 years seniority as an English teacher. The most severe judge is female, graduated from Literature Faculty, has 20 years seniority as an English teacher, and has worked in different schools during her tenure. Therefore, it is likely to say that working in different types of schools and becoming experienced in the field of English teaching can lead to a tendency to evaluate curriculum meticulously and behave strictly in scoring.

According to "infit" and "outfit" values related to facets, except for the material coded A-M3 in which infit or outfit data exceeded their limit values, all the materials have an acceptable level of 1.5. This result indicates that almost all items are consistent with the evaluation of the materials and their infit and outfit mean squares are within the acceptable values (1.4 - 0.6). In other words, items evaluate different characteristics belonging to the materials and have shown statistically significant differences. As in interpreting the traditional reliability results, the closer the correlation comes to +1.00, the more reliable the test is (Bastürk, 2010, p. 57). In the current study, a reliability value of 1.00 for determination of the quality of curriculum materials, 0.99 for determination of the judges' severity/leniency levels, and 0.89 for determination of difficulty or easiness of the items has been obtained.

Qualitative findings of the study were discussed under two titles: "The views of the participants about the characteristics of ECM" and "The views of the participants about the relationship between ECM and English curriculum for 9th, 10th, 11th and 12th grades. According to the results of MAXQDA analysis, four themes related to the characteristics of ECM emerged from the comments: "current situation," "deficiencies," "suggestions," and "what to do." Three themes related to the relationship between ECM and English curriculum emerged as "current situation," "deficiencies," and "suggestions." The theme "deficiencies" had the most dominant dispersing of teacher comments, while the theme "what to do" was not notable to every teacher.

On the other hand, the crucial problems that the participant teachers complained about regarding ECM were related to insufficient listening activities in the *deficiencies* theme. Since listening is vital both in language learning and in daily communication, participants complained about insufficient curriculum objectives for the evaluation of listening comprehension. They recommended textbooks be renewed through evolving technologies in the *suggestions* theme. However, although today there is a great variety of auditory and audio-visual materials available for different classroom listening activities (e.g., DVD/VCD players, films, computers, smart boards), many

teachers still have neither the time, nor access to adequate technology to create 'authentic' audiovisual materials (McGrath, 2013). Similarly, Yel (2009) evaluated the effectiveness of English curriculum in Anatolian High Schools in Sivas. The results revealed that the curriculum contents and materials were uninteresting for the students. In addition, materials were inadequate in providing opportunities for communicative and student-centered activities. Based on the expression "a picture is worth a thousand words," it can be said that the effectiveness of visual materials in language teaching has been universally acknowledged (Pakkan, 1997). Therefore, teachers recognize the importance of using technology effectively.

The results obtained from the views of the participants about the relationship between ECM and English curriculum revealed inadequate smart board presentation techniques, insufficient reading texts and vocabulary teaching, and difficulties stemming from common exams. However, the use of smart boards in a foreign language teaching environment studied by Hall and Higgins (2005), Hur and Suh (2012), and Mathews-Aydınlı and Elaziz (2010) was found to be versatile, funny and enjoyable. According to Hur and Suh (2012), vocabulary teaching is more effective via smart boards which enhance new kinds of learning processes and support language acquisition. Therefore, it is recommended that teachers receive training in using technological materials and in creating their own materials. It is also suggested that balanced distribution of four language skills in ECM should be improved in achieving the objectives of the curriculum.

References

Arıkan, A. (2008). Topics of reading passages in ELT coursebooks: What do our students really read? *Reading Matrix*, *8*(2), 70–85.

Arıkan, A. (2009). Problems with coursebooks in EFL classrooms: Prospective teachers' opinions. *EKEV Academic Review*, *38*, 309–317.

Arıkan, A., Saraç, S., Köse, N., Tekir, S., & Köse, Ö. (2011). *English break A2* (Z. Brooks & E. Wolfson, Eds.). Ankara, Turkey: Gündüz Eğitim ve Yayıncılık.

Baştürk, R. (2010). Bilimsel araştırma ödevlerinin çok yüzeyli Rasch ölçme modeli ile değerlendirilmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, *1*(1), 51–57.

Batdı, V. (2010). *An assessment of teachers' opinions about the effectiveness of English coursebooks used in foreign language high schools (case of Elazig and Diyarbakir cities)* (Master's thesis, University of Firat, Elazığ, Turkey). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Baydar Ertopcu, F., İnci, H., & Özbıçakcı Samur, S. (2014). *Yes you can*. Ankara, Turkey: MEB Komisyon.

Baydar Ertopcu, F., İnci, H., Özbıçakcı Samur, S., & Yalçın, A. (2014). *Yes you can*. Ankara, Turkey: MEB Komisyon.

Bektaş, M., & Tekir, S. Ç. (2014). *English Break A1* (R. E. Grammer, Ed.). *Ankara, Turkey:* Gündüz Eğitim ve Yayıncılık.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model: Fundamental measurement in the human sciences*. (2nd ed.). New Jersey, NJ: Lawrence Erlbaum Inc. Publishers.

Boone, W., & Scantlebury, K. (2006). The role of Rash analysis when conducting science education research utilizing multiple-choice tests. *Science Education*, *90*, 253–269.

Byrd, P. (1995). *Materials writers guide*. Boston, MA: Heinle & Heinle.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Crystal, D. (1987). *The Cambridge encyclopedia of language*. Cambridge, UK: Cambridge University Press.

Cunningsworth, A. (1995). *Choosing your coursebook*. Oxford, UK: Heinemann.

D'Andrea, L., Waters, C., & Rudd, R. (2011). Using computer assisted qualitative software (CAQDAS) to evaluate a novel teaching method for introductory statistics. *International Journal of Technology in Teaching and Learning*, *7*(1), 48–60.

Davey, L. (1991). The application of case study evaluations. *Practical Assessment, Research & Evaluation*, *2*(9). Retrieved from http://PAREonline.net/getvn.asp?v=2&n=9

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research, 5*, 194–197.

Dudley-Evans, T., & St John, M. (1998). *Developments in ESP: A multi-disciplinary approach*. Cambridge, UK: Cambridge University Press.

Ertürk, F., Ateş, M., & Alkan, S. (2014). *Yes you can*. Ankara, Turkey: MEB Komisyon.

Ertürk, F., Ateş, M., Alkan, S., & Albayrak, Y. (2014). *Yes you can*. Ankara, Turkey: MEB Komisyon.

Falla, T., & Davies, P. A. (2012). *Solution*: *Elementary student's book.* Oxford, UK: Oxford University Press.

Hall, I., & Higgins, S. (2005). Primary school students' perceptions of interactive whiteboards. *Journal of Computer Assisted Learning*, *21*, 102–117.

Hanson, W. E., Creswell, J. W., Clark, V. L. P., Petska, K. S., & Creswell, J. D. (2005). Mixed methods research designs in counseling psychology. *Journal of Counseling Psychology*, *52*, 224–235.

Harmer, J. (2001). Coursebooks. A human, cultural and linguistic disasters? *MET*, *8*(4), 5–10.

Howard, J., & Major, J. (2005) *Guidelines for designing effective English language teaching materials*. In *Proceedings of the 9th Conference of Pan-Pacific Association of Applied Linguistics*, 101–109. Seoul, South Korea: PAAL9.

Hur, J. W., & Suh, S. (2012). Making learning active with interactive whiteboards, podcasts, and digital storytelling in ELL classrooms. *Computers in the Schools*, *29*(4), 320–338. http://dx.doi.org/10.1080/07380569.2012.734275

Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learner-centered approach.* Cambridge, UK: Cambridge University Press.

Lewins, A., & Silver, C. (2008). *Using software in qualitative research*. Los Angeles, CA: Sage.

Linacre, J. M. (1993, April). *Generalizability theory and many-facet Rasch measurement*. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA.

Lund, R., & Zoughby, K. (2007). English language textbooks in Norway and Palestine. In M. Horsley & J. McCall (Eds.), *Peace, democratization and reconciliation in textbooks and educational media* (pp. 203–211). IARTEM and Vestfold University College.

Lunz, M. E., & Linacre, J. M. (1998). Measurement designs using multifacet Rasch modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research. Methodology for business and management* (pp. 47–77). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mathews-Aydinli, J., & Elaziz, F. (2010). Turkish students' and teachers' attitudes toward the use of interactive whiteboards in EFL classrooms. *Computer Assisted Language Learning*, *23*(3), 235–252.

McGrath, I. (2013). *Teaching materials and the roles of EFL/ ESL teachers. Practice and theory.* London, UK and New York, NY: Bloomsbury Academic

Mills, A. J., Durepos, G., & Wiebe, E. (Eds.) (2010). *Encyclopedia of case study research* (Vol. I and II). Thousand Oaks, CA: Sage.

Pakkan, G. (1997). *Language teaching materials evaluation and selection, adaptation and development*. Ankara, Turkey: Sürat English Language Teaching.

Richards, J. C. (2001). *Curriculum development in language teaching*. Cambridge, UK: Cambridge University Press.

Richards, J. C. (2005, March). *Materials development and research – Making the connection*. Paper presented at a colloquium on research and materials development, at the TESOL Convention, San Antonio.

Savignon, S. (1997). *Communicative competence: Theory and classroom practice* (2nd ed.). New York, NY: The McGraw-Hill Companies.

Sheldon, L. (1988). Evaluating ELT textbooks and materials. *ELT Journal*, *37*(3), 251–255.

Soars, L., & Soars, J. (2011). *New headway: Elementary student's book*. Oxford, UK: Oxford University Press.

Tashakkori, A., & Teddlie, C. (Eds.) (2003). *Handbook of mixed methods in social & behavioral research*. California, CA: Sage.

Tomlinson, B. (2001). Materials development. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 66–71). Cambridge, UK: Cambridge University Press.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. Transactions of the Rasch measurement SIG American educational research association. *Rasch Measurement Transaction, 8*(3), 370–382.

Yel, A. (2009). *Evaluation of the effectiveness of English courses in Sivas Anatolian High School* (Master's thesis, Middle East Technical University, Ankara, Turkey). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage.

Yurdugül, H. (2005, September). *Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması*. A paper presented in XIV. National Education Sciences Congress (University of Pamukkale, Denizli, Turkey).

Zhang, Y. (2007). Literature review of material evaluation. *Sino-US English Teaching*, *4*(6), 28–31.