Bridging Theory and Practice: Exploring Bellon and Handler's Curriculum Evaluation Model and its Use in Literature*

Teoriden Pratiğe: Bellon ve Handler'ın Program Değerlendirme Modeli ve Literatürdeki Kullanımı

Seda Aydan[1*] - Hürriyet Saridemir[2]

[1] Dr., University of California San Diego, Department of Education Studies, CA, USA, https://ror.org/0168r3w48, https://orcid.org/0000-0001-5807-3956, saydan@ucsd.edu

Dr., Kaliforniya Üniversitesi San Diego, Eğitim Çalışmaları Bölümü, Kaliforniya, ABD, https://ror.org/0168r3w48, https://orcid.org/0000-0001-5807-3956, saydan@ucsd.edu

[2] Hürriyet Sarıdemir, Middle East Technical University, Faculty of Education, Ankara, Türkiye, https://ror.org/014weej12, https://orcid.org/0000-0003-2671-3263, hurriyet@metu.edu.tr

Hürriyet Sarıdermir, Orta Doğu Teknik Üniversitesi, Eğitim Fakültesi, Ankara, Türkiye, https://ror.org/014weej12, https://orcid.org/0000-0003-2671-3263, hurriyet@metu.edu.tr

* Corresponding author

Öz

Program değerlendirmesi, daha etkili programlar tasarlamak ve hâlihazırda var olanları iyileştirmek için önemli bir süreçtir. Program değerlendirme çalışmalarının bulguları sayesinde, araştırmacılar ve eğitim kurumlarının paydaşları, programın amacına hizmet edip etmediğini, öğretimin etkili olup olmadığını kontrol edebilir ve programın öğrencilere olan etkileri hakkında sonuçlar çıkarabilirler. Bu bulgular sonucunda, programın güçlü ve zayıf yönleri belirlenir. Bellon ve Handler'ın Program Değerlendirme Modeli, ele alınan eğitim programları için teorik bir çerçeve sağlayan bir modeldir. Model, değerlendirme için veri toplamak amacıyla çeşitli kaynakların kullanılmasını önerir ve öğrencileri eğitim programının merkezine yerleştirir. Ayrıca, model mevcut kaynakların ve öğrencilerin ihtiyaçlarının değerlendirilmesini tavsiye eder. Bu modeli diğer modellerden ayıran özellik, program geliştirme ve program değerlendirmeyi bir arada ele almasıdır. Bu nedenle, aslında diğer modellere kıyasla daha avantajlıdır. Ancak, avantajlarına rağmen, nadiren faydalanılan bir model olarak kalmıştır. Bu modelin sınırlı uygulanmasının olası nedenlerini anlamak amacıyla, mevcut çalışma, bu alandaki alan yazını inceleyerek, Bellon ve Handler'ın Değerlendirme Modeli'nin zaman içinde program değerlendirme çalışmalarında nasıl kullanıldığını sistematik bir şekilde raporlamaktadır. Bu nedenle, mevcut çalışma, modelin yayımlandığı yıldan itibaren Bellon ve Handler'ın Değerlendirme Modeli ile yapılan program değerlendirme çalışmalarını analiz eden bir alan yazıntaraması sunmaktadır. Bu çalışmalar, dört Yüksek Lisans tezi, dört Doktora tezi ve üç araştırma makalesinden oluşmaktadır. Alan yazın taraması sonucunda elde edilen bulgular incelendiğinde görüşmüştür ki bu modelin yardımıyla yapılan çalışmaların genellikle İngilizce öğretim programını ele almaktadır. Çalışmaların çoğunluğu Türkiye bağlamında yürütülmüştür. Bu çalışmalar hem öğrencilerden hem de öğretmenlerden alınan nicel ve nitel veriler kullanılarak yürütülmüştür. Çalışmaların çoğunluğu vaka çalışmalarıdır. Anketler, görüşmeler, sınıf gözlemleri ve sınav belgelerinin incelenmesi en yaygın kullanılan veri toplama yöntemleridir. Araştırmacılar tarafından toplanan nitel veriler, belge analizi yöntemiyle analiz edilirken, nicel veriler betimsel istatistikler elde edilerek analiz edilmiştir. Bu program değerlendirme çalışmalarının sunduğu bazı yaygın sorunlar etkileşimli öğretim materyallerinin eksikliği ve ders içeriği ile uyuşmayan verimsiz ölçme ve değerlendirme araçları olarak özetlenebilir. Model, program geliştirme ve program değerlendirmeyi bir arada ele aldığı için, bu teorik çerçeve içinde gerçekleştirilen tüm çalışmalar, programları iyileştirmek adına çözümler de sunmaktadır. Bu çözüm önerilerinden bazıları ders amaçlarının gözden geçirilmesi ve öğrencilerin ilgi alanlarına hitap eden ders materyallerinin okulun materyal birimi tarafından hazırlanmasıdır. Ayrıca, daha önce de belirtildiği gibi, bu araştırma, modelin diğer değerlendirme modellerine kıyasla nadiren kullanılmasının olası nedenlerini modelin ayırt edici özelliklerini açıklayarak incelemektedir. Model, değerlendirmeden çok geliştirmeyi önceliklendirmektedir. Bu da onu paydaşların bilgi ihtiyaçlarını karşılamayı hedefleyen, diğer program geliştirme modellerinden farklı kılmaktadır. Program değerlendirme çalışmaları çoğunlukla eğitim kurumlarını paydaşları tarafından talep edilmesi bu modelin kullanılma ihtimalini düşürmektedir. Bir diğer etken ise bu modelde yapılan çalışmaların sadece İngilizce öğretimine odaklanmış olması bu nedenle de diğer dersler için değerlendirme sürecini yürütmeye ışık tutacak çalışmalar olmamasıdır. Alan yazındaki bu eksiklik modelin yaygın kullanılmasını olumsuz etkilemektedir. Bu nedenle bu araştırmanın bulgularına dayanarak, bu modelin İngilizce dil öğretimi dışındaki çeşitli disiplinlerde daha fazla değerlendirilmesi önerilmektedir. Bu yaklaşım, modelin çeşitli çalışma alanlarındaki uygulanabilirliğini keşfetmeye olanak tanıyacak ve faydalılığını kapsamlı bir şekilde anlamayı sağlayacaktır. Ayrıca modelin başka disiplinlerde de kullanılması bu alanlardaki eğitim programlarının iyileştirilmesine de katkıda bulunacaktır. Son olarak modelin çoğunlukla Türkiye bağlamında kullanıldığı tespit edilmiştir. Bu nedenle modelin başka bağlamlardaki kullanımını ve bu kullanımını etkileri görmek amacıyla başka bağlamlarda da kullanılması ve sonuçlarının bildirilmesi bu alandaki alan yazına katkıda bulunacaktır.

Anahtar Kelimeler: Eğitim Bilimleri, Eğitimde Program Değerlendirme, Bellon ve Handler Değerlendirme Modeli, Belgelere Dayalı Araştırma

Research Article

Plagiarism Checks
This article has been reviewed by at least two referees and scanned via a plagiarism software.

Peer-Review
Single anonnymized-One internal (Editorial Board). Double anonymized-Two extarnal.

Ethical Statement
It is declared that scientific and ethical principles have been followed while carrying out and writing this study and that all the sources used have been properly cited. Lütfiye Seda Aydan-Hürriyet Saridemir

Complaints
turkisharr@gmail.com

Conflicts of Interest
The author(s) has no conflict of interest to declare.

* This article is a version of the paper titled " A Literature Review on Curriculum Evaluation Studies Using Bellon and Handler's Evaluation Model " by developing and partially changing its content, which was presented orally at the 8th Conference of Curriculum and Instruction Studies without publishing its full content.

Abstract
Curriculum evaluation is a crucial process for designing more effective programs and for bettering the already existing ones. Thanks to the findings of the program evaluation studies, researchers and stakeholders of educational institutions can assess if the program serves its aims, whether the instruction is effective, and they can have conclusions drawn on the impacts of the program on the learners. As a result of these findings, the strengths and weaknesses of the program are determined. Bellon and Handler model of curriculum evaluation is a model which provides a theoretical framework for the evaluation of educational programs. The model proposes the use of a variety of sources to collect data for the evaluation and places the learners in the center of the curriculum. Besides, the model recommends the evaluation of the existing sources as well as the needs of the students. What makes this model distinctive from the other models is its involvement of program development as well as program evaluation. Thus, in fact, it has several advantages compared to the other models. However, despite its advantages, the model has been rarely utilized. To provide an understanding of the possible reasons for the limited application of this model, the present study explores the existing body of research and systematically reports the ways in which Bellon and Handler's Evaluation Model has been employed in curriculum evaluation studies over time. Therefore, the present study presents an exploratory literature review analyzing the curriculum evaluation studies conducted with Bellon and Handler's Evaluation Model since 1982, the year the model was published. These studies consisted of four master's theses, four Ph.D. theses, and three research articles. The findings of the study indicate that studies conducted with the help of this model usually deal with English language teaching programs. They were mostly carried out in the context of Turkey. They make use of both quantitative and qualitative data from both students and teachers. The majority of the studies are case studies. Questionnaires, interviews, classroom observations, exam documents are the most employed data collection methods. The qualitative data collected by the researchers were analyzed via document analysis while the quantitative data were analyzed by obtaining descriptive statistics. Some common problems reported as a result of the evaluation were the lack of interactive teaching materials and inefficient assessment tools which do not align with the objectives of the class content. Since the model deals with program development as well as program evaluation, all the studies which were carried out within this theoretical framework offer solutions to improve the programs. Some of these solutions are revisiting the course materials and preparation of course materials appealing to the interests of the learner profiles. Additionally, as mentioned earlier, this research explores the potential factors contributing to the infrequent utilization of the model in comparison to other evaluation frameworks by elucidating its distinct characteristics. The model prioritizes enhancement over assessment, and this makes it diverge from other models aiming to fulfill stakeholders' informational requirements. Because program evaluation is frequently initiated by the stakeholders of educational institutions, the models prioritizing the informational needs of the stakeholders are preferred over this model which prioritizes improving the curricula. Another factor contributing the infrequent use of this model is the lack of guiding studies in the literature using this model to evaluate educational programs other than English language teaching. Based on the findings of this investigation, it is recommended to conduct further evaluations employing this model across diverse disciplines beyond English language teaching. This approach would facilitate an exploration of the model's applicability in various fields of study and yield a comprehensive understanding of its utility. Finally, the results of the study showed that the majority of the publications were sampled in the context of Turkey. For this reason, carrying out studies in a range of contexts can contribute to literature by shedding light into the use of the model in diverse contexts.

Keywords: Educational Sciences, Curriculum Evaluation, Bellon and Handler's Evaluation Model, Program Evaluation, Review of Literature

Introduction

Curriculum is a field which is defined as lacking clear-cut or definite boundaries (Oliva, 1997). Therefore, scholars have provided a variety of definitions of curriculum considering their philosophies and how they see or describe education. Despite the lack of agreement on the definition of the curriculum and its boundaries, studies published in the field of education so far agree on the notion that available curriculum should be reevaluated, revised and updated (Ornstein & Hunkins, 2004; Zajda, 2024). Curriculum evaluation though being a relatively newer discipline as well as being portrayed as a trans-discipline is now acknowledged as an indispensable part of program development stages (Fitzpatrick et al, 2004). Many researchers working in this area (e.g. White, 1971) validated curriculum evaluation and recognized it as an element of curriculum development process. In a widely accepted definition of curriculum evaluation, Scriven (1967) indicates that evaluation is basically appraising the worth or merit of something. Fitzpatrick et al. (2004), however, state that evaluation is the act of assessing the value of something.

Considering a range of definitions of curriculum evaluation as well as educational philosophies and expectations from the outcomes of evaluation, there are various curriculum evaluation models offered by various scholars. The curriculum evaluation model created by Bellon and Handler (1982) can be identified as one of the key curriculum evaluation models in the literature. However, the number of curriculum studies conducted with the help of this model is quite limited compared to more widely recognized models, such as Tyler's (1949) objective-based model, Stufflebeam's (1971) CIPP model, and Scriven's (1991) goal-free evaluation model. This lack of application raises important questions regarding the model's visibility, adaptability, and perceived effectiveness in educational evaluation. Given the limited number of studies employing this model, it is essential to systematically examine its use in the literature to better understand its potential contributions to curriculum evaluation. Therefore, an exploratory review of literature is necessary to reveal the studies carried out with this model in literature which provides the opportunity to see the patterns in these studies and make inferences on why the model has been underused. Thus, a deeper analysis of how this model has been applied—or why it has been overlooked—can reveal significant insights for educators and researchers. By mapping the existing research landscape, this review aims to assess the extent to which the model has been implemented, in what contexts, and with what outcomes. Furthermore, by synthesizing available studies, this review aims to establish a conceptual framework for future research on the model's use in curriculum evaluation. Without a comprehensive synthesis, researchers lack a clear foundation for understanding how the model has been operationalized in different educational contexts. For this reason, an exploratory review serves as a starting point for further empirical and theoretical investigations, providing a roadmap for future research directions (Onwuegbuzie & Frels, 2015). The present study addresses these needs in research and offers a comprehensive report of the studies conducted with this model. Additionally, this study takes a holistic approach by integrating diverse sources, including master's theses, doctoral dissertations, and journal articles. Given that exploratory reviews allow for the inclusion of varied research perspectives, this study presents a broad yet detailed picture of how the model has been employed in curriculum evaluation efforts. Such an approach ensures that both academic and applied perspectives are considered, making the findings relevant to both researchers and practitioners.

Finally, this review has direct implications for educational policy and practice. In an era where curriculum evaluation plays a critical role in shaping educational quality, policymakers and educators require diverse evaluation tools that align with different educational philosophies and institutional needs. By systematically

reviewing the limited body of research on the Bellon and Handler model (1982), this study can provide evidence-based insights that inform curriculum development strategies, teacher training programs, and institutional evaluation practices. Understanding these trends can inform future curriculum evaluation practices, particularly in contexts where adaptive, feedback-driven models are preferred over static, summative approaches.

Given these considerations, the present exploratory literature review is necessary to fill in a niche in research, as it seeks to bridge the gap between theoretical knowledge and practical application of the Bellon and Handler model in curriculum evaluation model providing a comprehensive understanding of the model's role, limitations, and potential for broader adoption in educational research and practice. With such rationale in mind, the following research questions were asked:

1.   What are the research methods (research design, data sources, and data collection tools) of evaluation studies that have been conducted using the curriculum evaluation model by Bellon and Handler (1982)?

2.   What are the purposes and results of evaluation studies that have been conducted using the curriculum evaluation model by Bellon and Handler (1982)?

1.   Theoretical Background on Curriculum Evaluation and Classification of the Models

Gredler (1996) informs that the term curriculum evaluation has originally been defined by Tyler (1949) as a fundamental stage for curriculum development for detecting the extent to which the curriculum had reached its pre-determined goals or diagnosing strengths and weaknesses in the Eight Year study in 1932.

Although the studies of curriculum evaluation are shaped around the definitions above, there are different approaches following different steps so as to conduct the evaluation. Once the literature on curriculum evaluation is taken into consideration, it is concluded that there are various models, and they have been classified according to different perspectives. Two of these classifications have been done by Gredler (1996) and Fitzpatrick et al. (2004). While Gredler (1996) created two broad categories namely the utilitarian-oriented approach and pluralist approach, Fitzpatrick et al. (2004) collected the evaluation models under 5 categories considering the aim of evaluation. These categories can be defined as objectives-oriented evaluation approach, management-oriented evaluation approach, consumer-oriented evaluation approach, and expertise-oriented evaluation approach, and participant-oriented evaluation approach. On the other hand, the model of Bellon and Handler (1982) used in this study does not belong to the categories mentioned above as it focuses on the use of evaluation for curriculum design purposes. For this reason, it stands apart from these categorizations and takes place in literature as a different model.

1.1.   Bellon and Handler's Curriculum Evaluation Model

Bellon and Handler (1982) suggest an evaluation model that integrates curriculum and development processes, and they state that curriculum development activities should be conducted by taking the evaluation information into consideration. Otherwise, the development activities would not achieve satisfactory results. Thus, in their framework, evaluation information should help the evaluator to navigate the curriculum development process.

According to Mutlu and Şimşek (2018) there are several underlying assumptions for Bellon and Handler's model that should be noted in interpreting and analyzing the design they present. First, the main goal of curriculum design and evaluation is to enrich educational programs to enable learners to have fruitful learning opportunities.

Steps to be taken to reshape the curriculum need to be assessed in terms of their likely effect on student learning. By particularly focusing on student learning opportunities, development and evaluation can be carried out. Second, program improvement activities reach the optimum level when personnel at all levels show dedication to realizing pre-set goals. Program improvement efforts are more likely to be accomplished when all appropriate personnel are actively involved in steps to achieve priority goals. Third, curriculum workers want their educational programs to be well planned and effective. Teachers, administrators and other curriculum workers should share a desire to be professionally competent. To achieve the conditions that motivate optimum performance, it is essential to enable people to work within the context of a well-organized, effectively functioning program. Therefore, school system personnel should be provided with clear information about the curriculum improvement process so that they can feel committed to curriculum development and evaluation process. The last assumption of the model is about the necessity of a well-planned, ongoing systematic process for curriculum improvement attempts. According to this assumption, a systematic approach that provides room for feedback and amendment is crucial. A systematic ongoing process ensures that the programs remain sensitive to the needs of students and communities as they keep changing.

Based on the assumptions mentioned above, development and evaluation need to proceed simultaneously. The activities should be purposeful, well organized, and adapted to local conditions and the results should be clearly communicated to all appropriate groups, along with plans for acting on major recommendations. The principal elements of this framework are the four areas of focus which include the status descriptions, the analysis activities, and the cumulative improvement components (Bellon & Handler, 1982; Mutlu & Şimşek, 2018).

1.2. The Four Focus Areas of Bellon and Handler's Model

The four focus areas, namely goals, organization, operations, and outcomes, are considered as the main dimensions of any curriculum by Bellon and Handler (1982). The first area, educational goals are general statements of purpose which give direction to the curriculum. They deal with the desired outcomes and expectations of the program. Bellon & Handler (1982) underline that program goals should be based on clearly identified learner needs.

The second area of focus stated by Bellon and Handler (1982) is organization. In the framework of this model, organization refers to a review and analysis of several factors playing a role in the way the program is operated effectively. These factors include the specific programs which are put forward, the resources to be used, the ways in which processes like interacting and planning are managed, and the structure or arrangement of the school or district.

The operations area is about day-to-day functioning of programs. The implementation process of the curriculum is studied so as to assess how well general educational goals are being translated into course or unit goals and then specific learning objectives during the instruction. At this stage, assessing the effectiveness of the instruction is an essential concern and assessment of instruction provides insights on how effective the instruction is. For this reason, a study of program operations helps teachers improve effective strategies to enhance student learning (Aktaş & Gündoğdu, 2020).

The final area of focus is the outcomes of the program in Bellon and Handler's Model. At this stage, attention should be given to the unintended as well as intended program results to see the effects of the curricula on student learning, motivation or attitudes. How the outcomes affect other programs, or influence teachers and

students should also be carefully considered. (Bellon & Handler, 1982; Mutlu & Şimşek, 2018)

All in all, these four focus areas provide a theoretical framework for the curriculum evaluators to consider while assessing the worth and merits of the curriculum providing them also a ground to merge curriculum development and assessment (Mutlu, 2020).

1.3. Stages of Bellon and Handler Model

Bellon and Handler Model provides a neat organization of the curriculum evaluation through the stages they suggested in the model. Status description is the first step of this process, and it refers to the initial assessments including gathering information about program goals, organization, operations, and outcomes. Conducting a status study can involve a combination of document analyses and interviews or surveys. Curriculum guides, accreditation reports, Board of Education minutes, faculty handbooks, and course syllabi could be utilized for document analysis. Interviews and surveys can also be conducted considering the needs of the context. Obtaining information from different viewpoints like representative teachers, administrators, other staff members, parents, and students is advisable (Bellon & Handler, 1982).

According to Bellon and Handler (1982) status description should be short and straightforward. as additional descriptive information should be generated in later phases of the improvement process. Data obtained for status study may suggest possible recommendations for improving the program and this information should not be acted on until analysis activities are completed. The areas that need to be focused on and the ones that require less attention should be identified Thus, status descriptions often lead to more efficient use of the resources available for curriculum improvement (Mutlu & Şimşek, 2018).

After the current status has been determined, program evaluation efforts should be directed to analyze the effectiveness of the program. Curriculum goals, organizations, operations, and outcomes are considered and evaluated to judge what steps can be taken to better the educational program. Status study information is utilized, and additional data are gathered if needed. Observations, interviews, questionnaires, tests, and attitude scales are among the most common data collection tools to collect needed information (Bellon & Handler, 1982).

When judgments about the effectiveness of the program are made, discrepancies or mismatches that may be available between current and expected program functions can be revealed (Bellon & Handler, 1982). These mismatches can be used to make decisions on program improvement which is one of the aspects of Bellon and Handler model making it different from the other models (Yel ,2009). Bellon and Handler (1982) also underscore that appropriate criteria or standards must be considered in decision making activities in this process such as professional association guidelines, research reports, and needs assessments. When carried out well, this phase leads to recommendations for improvement that will have a positive impact on the local situation and the students.

The curriculum evaluation process, as outlined by Bellon and Handler (1982), involves an ongoing interaction between status studies and analysis activities, facilitating a systematic approach to program improvement. The process is inherently iterative, where insights gained from analysis contribute to refining the evaluator's understanding of the current program. This dynamic interaction allows evaluators to make informed modifications in key areas without needing to wait for the completion of all analytical phases before implementing necessary changes (Bellon & Handler, 1982).

A structured approach to curriculum improvement is achieved through four key components that guide recommendations and action plans. Each phase of analysis contributes to a clearer vision of the desired curriculum.

For example, goal analysis may confirm that existing objectives remain relevant or indicate the need for revision to align with evolving educational needs. These refined goals serve as benchmarks for subsequent phases, particularly in organizational analysis, which examines the alignment between program structures and the newly established objectives. Notably, some aspects of organizational analysis can commence before the finalization of goals, but the process remains incomplete until the agreed-upon objectives are fully integrated into the program's framework (Bellon & Handler, 1982; Mutlu & Şimşek, 2018).

The model's cumulative improvement approach emphasizes a gradual, success-oriented change strategy. The operations and outcomes components build upon prior analytical findings, allowing for continuous refinement and necessary adjustments at each stage. This systematic approach ensures that individual components are examined both independently and in conjunction with one another, reinforcing the effectiveness of the evaluation process. Additionally, open communication and continuous feedback among all major components are essential for fostering coherence and responsiveness throughout the curriculum improvement process. By maintaining an ongoing exchange of information, evaluators can enhance the functionality of each program element (Bellon & Handler, 1982; Yel, 2009).

A crucial aspect of this model is the role of staff development in sustaining curriculum improvements. Once areas for enhancement are identified, targeted professional development opportunities can be provided to educators and administrators. Follow-up activities further support the effective implementation of curriculum modifications, ensuring that instructional practices evolve in alignment with program objectives. These professional development initiatives do not only address localized needs but also contribute to a unified and purpose-driven curriculum framework (Bellon & Handler, 1982).

Overall, it can be concluded that the Bellona and Handler model of curriculum evaluation provides several advantages for program evaluators and curriculum developers. However, its application remains limited in the literature, which forms the foundation for the present research.

2. Method

This study employs an exploratory literature review approach to systematically collect, analyze, and synthesize prior research on curriculum evaluation studies using the Bellon and Handler Model. According to Onwuegbuzie and Frels (2015), exploratory literature reviews are particularly useful in fields where research is sparse or fragmented, as they enable researchers to map existing studies, identify patterns, and highlight gaps for future inquiry. Exploratory reviews allow for a more flexible and iterative approach to analyzing literature from multiple contexts, perspectives, and methodological orientations, which fits into the objectives of the present study. Given the limited number of studies utilizing the Bellon and Handler Model (1982) in curriculum evaluation, an exploratory review was deemed appropriate to assess the extent of existing research, synthesize key findings, and identify areas for further study.

2.1. Data Collection and Inclusion Criteria

To ensure comprehensive coverage, data were collected from five major academic databases: ERIC, Web of Science, YÖK Tez, EBSCOhost, and Google Scholar. The search spanned studies published between 1982 and 2025, utilizing the keywords "Bellon and Handler Model" and "Curriculum Evaluation". Only full-text, open-access publications were included in the review to ensure accessibility and transparency. Given the exploratory nature of this study, peer-reviewed journal articles, doctoral dissertations, and master's theses were included, while

conference proceedings and non-peer-reviewed sources were excluded to maintain academic rigor.

2.2. Data Analysis

A total of 11 studies met the inclusion criteria, comprising four master's theses, four doctoral dissertations, and three journal articles. To analyze these sources systematically, thematic analysis method was employed (Braun & Clarke, 2006). Given that the primary objective of this review was to examine how the Bellon and Handler Model (1982) has been applied in curriculum evaluation, deductive thematic analysis method was utilized, meaning that the coding framework was developed based on predefined research questions rather than emerging inductively from the data (Braun & Clarke, 2006). This approach ensured that the identified themes were directly aligned with the study's objectives and contributed to a structured synthesis of findings across the reviewed literature. The researchers collaborated on the search process to find the relevant articles and after that a blind coding process was carried out. Once the coding process was completed, the researchers got together to discuss the codes and make final amendments. During the process, it was seen that the codes that the two researchers gave to the articles and their content mostly matched. Only a few minor changes were made to the wording of the codes. Once this process was completed, the research findings section was written by classifying and synthesizing the research via the codes. Thus, by integrating exploratory literature review methodologies with deductive thematic analysis, the present study provided a systematic approach to understanding the limited body of research on the Bellon and Handler Model (1982), while also laying the groundwork for future investigations in curriculum evaluation.

In conducting our exploratory literature review using thematic analysis, we made concerted efforts to control several potential threats to the validity of our findings. These threats were systematically addressed to ensure the credibility and rigor of our analysis. To address the selection bias, we ensured that the studies included in our review were selected through a comprehensive search strategy. A definite set of inclusion and exclusion criteria were defined, and our selection process was kept transparent. This approach minimized the risk of omitting relevant studies and ensured a broad representation of the literature. To mitigate the impact of publication bias, the researchers made efforts to include not only peer-reviewed articles but also grey literature, such as reports and dissertations. This broader selection helped reduce the over-representation of published studies with significant findings. Besides, to minimize researcher bias, certain steps were taken by using a consistent and systematic coding process throughout the thematic analysis. Multiple coders were created to enhance objectivity. Additionally, we reflected on our own preconceptions and ensured transparency in how conclusions were drawn. Also, to address confirmation bias, during the thematic analysis process, we remained open to findings that might contradict our initial expectations. We documented our decision-making process throughout the review process to prevent any undue influences from prior beliefs. Finally, to address interpretation and homogeneity bias, we followed a clear and consistent coding procedure and we actively sought to include studies with diverse methodologies, theoretical frameworks, and contexts. This effort aimed to capture a wide range of perspectives and avoid an overly narrow focus that could limit the breadth of our findings (Patton, 2002; Silverman & Marvasti, 2008).

3. Research Findings

In this section, the findings of the research are presented for each research question. Evaluation studies are chronologically presented.

3.1. Evaluation Studies that have been Conducted Using Bellon and Handler's Curriculum Evaluation Model (1982)

The studies carried out within the framework of the curriculum evaluation model by Bellon and Handler (1982) are summarized in Table 1.

Table 1

The Evaluation Studies Using the Curriculum Evaluation Model by Bellon and Handler (1982)

Publication No	Title	Authors	Year	Purpose
1	An Application and an Analysis of the Impact of the Bellon and Handler Curriculum Evaluation Framework to the English as a Foreign Language Program at the Huffco Indonesia Intensive English Language Program in Balikpapan	James Gilbert Ward	1987	PhD Dissertation
2	Systematic Language Program Development and Evaluation in Turkey.	Ayşe Yumuk	1989	Master's Thesis
3	Evaluating the English Language Curriculum at a Private School in Ankara: A Case Study	Hikmet Esin Erdem	1999	PhD Dissertation
4	Evaluating the Language Improvement Courses in the Undergraduate ELT Curriculum at Eastern Mediterranean University: A Case Study	Fatoş Erozan	2005	PhD Dissertation
5	An investigation of the effectiveness of the theme-based curriculum in the 2003-2004 academic year at the Department of Basic English at METU.	Tansu Topçu	2005	Master's Thesis
6	Evaluation of the Effectiveness of English Courses in Sivas Anatolian High Schools	Arzu Yel	2009	Master's Thesis
7	An evaluation of the Language Improvement Component in the Pre-service ELT Program at a College of Education in Kuwait: A Case Study.	Ahmad Al-Nwaiem	2012	PhD. Dissertation
8	A Program Evaluation Study of the Main Course at a Preparatory Program: A Case Study	Gülçin Mutlu	2018	Research Article
9	An Extensive Evaluation Study of the English Preparatory Curriculum of a Foreign Language School	Canay Karcı Aktaş, Kerim Gündoğdu	2020	Research Article
10	Evaluation of English as a Foreign Language Program in Turkey	Serpil Tekir	2020	Research Article
11	Evaluation Of An English Preparatory Program Using Bellon And Handler Model	Zeynep Uçkaya	2022	Master's Thesis

As shown in Table 1, 11 studies have been obtained within the time range from 1982 to 2025 in the aforementioned databases. Four of these publications are Ph.D. dissertation studies, four of these publications are masters' theses,

and three of these publications are published as research articles. All these studies have utilized Bellon and Handler Evaluation Model to evaluate English Language Teaching Programs or courses provided by English Language Teaching departments. Nearly all (n=9) of the studies have evaluated English Language Teaching Programs at a Higher Education Institution. Two studies carried out by Erdem (1999) and Yel (2009) have evaluated English Language Teaching Curriculum/courses at a K-12 school. Additionally, nearly all (nine) of the studies have been carried out in Turkey. Only one study carried out by Al-Nwaiem (2012) has been carried out in Exeter, England. The contexts of the studies are Turkey, Kuwait, Indonesia and Northern Cyprus.

There are various methods used in evaluation studies that have been conducted using the curriculum evaluation model by Bellon and Handler (1982) and these are summarized in Table 2.

Table 2

The Research Design of Evaluation Studies Utilizing the Bellon and Handler's Curriculum Evaluation Model (1982)

Publication No	Design	Data Sources	Data Collection Tools
1	Case study	Instructors, Administrative staff, Program-related documents	Interviews, Document Analysis
2	A proposal for a case study	Instructors, Administrative staff, Students, committee members and the secretary at the Preparatory School, the members of the Board of Trustees, the Ministry of Labor and National Employment Agency	Interviews, Document Analysis, Survey
3	Case study	Teachers, Students, School principals, One top management member	Questionnaires, Interviews, Classroom observations, Examination documents
4	Case study	Instructors, Students	Questionnaires, Interviews, Classroom observations, Document analysis
5	Case study	Instructors, Former students	Questionnaires, Interviews, Focus group interviews
6	Descriptive study	Teachers, Students	Questionnaires
7	Case study	Instructors, Students	Questionnaires, Interviews, Diaries, Document analysis
8	Case study	Instructors, Students	Questionnaires, Interviews, Observations
9	Case study	Instructors, Students	Questionnaire, Interviews, Class Observations, Document Analysis
10	Case study	Instructors, Students	Questionnaire, Interviews, Observations, Document analysis
11	Case study	Instructors, Students	Questionnaire, Interviews, Observations, Document Analysis

When the designs of the studies are examined, it is observed that most of the studies (n=9) have been carried out as case studies. The design of one of the studies carried out by Yumuk (1989) is a proposal for a case

study, and one of the studies is a descriptive one (Yel, 2009). In terms of data sources, teachers and instructors, being the implementers of the curriculum, constitute the main data sources of all studies. Although students are involved in most studies (n=9), two studies carried out by Ward (1987), and Topçu (2005) have not included students as data sources. Instead, Ward (1987) has included administrative staff and program related documents as data sources, and Topçu (2005) have collected data from former students. In the study carried out by Yumuk (1989), administrative staff, committee members and the secretary at the Preparatory School, the members of the Board of Trustees, the Ministry of Labor and National Employment Agency constituted other data sources. Erdem (1999) included school principals and one top management member to the data sources of the study.

Majority of the studies (n=8) have utilized questionnaires and interviews as data collection tools. Ward (1987) and Uckaya (2022) collected the data for the study by document analysis in addition to interviews. Similarly, Yumuk (1989) have utilized document analysis and survey beside interviews. Erdem (1999) utilized questionnaires, interviews, and classroom observations to collect data. Topçu (2005) included questionnaires, individual and focus group interviews to the data collection tools. He also carried out a document analysis of examination documents. Erozan (2005) conducted questionnaires, interviews, and classroom observations to collect data. She carried out a document analysis of course policy sheets, course materials, and assessment tools, too. Yel (2009) has only conducted questionnaires to collect data for the study. Different from other data collection tools, Al-Nwaiem (2012) utilized diaries to collect data along with questionnaires, interviews, and document analysis of relevant written documents. Mutlu (2018) has also conducted classroom observations in addition to questionnaires and interviews. Karcı et al. (2020) and Tekir (2020) have utilized classroom observations and document analysis to collect data along with questionnaires and interviews, too.

3.2. The Purposes and Results of Evaluation Studies that have been Conducted Using Bellon and Handler's Curriculum Evaluation Model

One of the early studies with this model was conducted by Ward (1987) to provide English as Foreign Language program administrators with an efficient, effective, systematic, and comprehensive guide to curriculum evaluation. The researcher detected some problems related to the program such as some difficulties in implementation, lack of enough sources to reinforce the class content and problems faced by instructors while covering the speaking objectives. Then the researcher asked for suggestions from the participants with the help of informal meetings and reported the suggestions that came up in these meetings in the study.

Yumuk (1989) also conducted one of the early studies with this model. However, rather than conducting an evaluation, the author created a proposal for a comprehensive evaluation with the relevant model. At the first phase of her study, she stated that some questions such as how to start planning an effective program, which steps to follow, and which factors to consider while developing and evaluating language programs are still unclear in the educational context of Turkey which has been an inspiration for the research study she proposed. With the emphasis of evaluating curriculum so as to improve the program, the staff, the administrative personnel and the program goals, she underlines the appropriacy of Bellon and Handler model into her context which is Bilkent English preparatory school. She conducts a fact-finding study in her proposal which refers to status description in the model. At this point, there is a different perspective suggested by the researcher which suggests that based on the findings of this stage, the stakeholders make a vital decision that is either to evaluate the existing program or to develop a new program. This is a distinctive suggestion as this decision is taken based on the results of the evaluation in many other curriculum evaluation models. In addition, this study proposed a wider range of people

from different groups to interview with a range from faculty secretary to freshman students. What is more, it proposes an analysis of budgetary side of the program. Thus, it provides a comprehensive guideline with Bellon and Handler framework for those who want to conduct an evaluation.

The study carried out by Erdem (1999) explores the English Language curriculum at the Middle East Technical University Development Foundation Ankara School, utilizing an adapted version of the Bellon and Handler curriculum evaluation model as well. The results of this study indicated that the current traditional teacher-centered set up of the ELT curriculum needs to be stored with the student-centered one; mechanisms and opportunities for in-service training service should be offered, and an on-going curriculum evaluation system needs to be installed.

Another investigation has scrutinized the language enhancement courses within the undergraduate curriculum of the Department of English Language Teaching (ELT) at Eastern Mediterranean University. The study primarily examined the current state of these courses across five key dimensions: objectives, content and materials, delivery, assessment, and coherence among courses. It also explored the perspectives of both instructors and students regarding these aspects. The findings generally indicated that the language enhancement courses were perceived as effective across the evaluation model used in the study. However, both students and instructors proposed adjustments to enhance effectiveness and better align with student needs and expectations (Erozan, 2005).

As another example of studies conducted within the framework of Bellon and Handler Model, Topçu (2005) explored the implementation of the theme-based curriculum used in a preparatory school at a state university in the 2003-2004 academic years. The results of the study were quite striking as they reflected a serious discrepancy between the perceptions of teachers and students. Teachers, particularly those teaching the intermediate group were found to have very negative attitudes towards the program. The researcher suggested that the implementation and quality of the materials and lack of communication between teachers and administrators are very likely to be the reasons for this discontent. In terms of objectives, reading was the skill most successfully attained and reading materials were reported as the most useful ones. However, interestingly, students found handouts much more useful than books which signaled an inappropriate choice of course book for the curriculum. In addition, writing was reported to be the most problematic area in the program and writing objectives were reported to be mostly unfulfilled. What is more, due to time constraints, teachers were observed to be more active in class and pair/group work were considered ineffective by the teachers for the attainment of objectives. Thus, with the help of these findings of the evaluation, Topçu (2005) made suggestions to improve the program under scrutiny particularly in terms of writing objectives and course content.

In the study carried out by Yel (2009), English language courses for 9th, 10th, 11th, and 12th grades in Anatolian high schools in Sivas have been evaluated with an eclectic approach of curriculum evaluation. In other words, two evaluation models suggested by Bellon and Handler (1982) and Brown (1995) have structured the evaluation model of this study. The results of this study have shown that the courses were not effective enough to reach the goals and objectives of the program. According to the students' and teachers' opinions, the course content and materials were not efficient to attract students' interest, and materials were inadequate to enhance student-centered activities and a communicative environment. Besides, the teaching and learning processes were not varied, and the assessment procedures were not parallel with the objectives of the courses.

Al-Nwaiem (2012) conducted an assessment study as part of his Ph.D. dissertation, examining the Basic

Language Skills Component (BLSC) within the ELT pre-service program at a College of Education in Kuwait. Employing an adapted version of Bellon and Handler's evaluation model, the study identified several shortcomings within the college's physical environment, such as aging buildings, a shortage of classrooms, and inadequate learning resources and facilities, including the library. Additionally, it highlighted the need for revisions to the BLSC's goals and objectives. Students expressed dissatisfaction with certain aspects of the BLSC curriculum, citing topics perceived as dull and outdated textbooks. A traditional, teacher-centered approach to teaching methods was identified as another area requiring improvement. Furthermore, students indicated a preference for alternative assessment philosophies over traditional methods. Mutlu (2018) conducted a study utilizing this model to assess the English as a Foreign Language (EFL) preparatory course, referred to as Main Course (MC), at a state university. The findings indicated that Main Course generally achieved effectiveness across four key dimensions: aims and objectives, course content and materials, course delivery, and student assessment and outcomes, as reported by both instructors and students. Both qualitative and quantitative data were employed to identify gaps between the current and desired state of the course across these aspects, largely corroborating with each other. However, inconsistencies emerged between the perceptions of teachers and students, particularly regarding course content and assessment. Karcı et al. (2020) conducted a study assessing the English preparatory curriculum at Aydın Adnan Menderes University School of Foreign Languages using the Bellon and Handler model. The research revealed the absence of r philosophy of education or objectives within the English preparatory curriculum. Despite students displaying positive motivation toward learning English, they struggled to attain the intended proficiency level. Additionally, all participants highlighted the ineffectiveness of the skills courses' instruction. The researchers also noted communication challenges between the administration and other stakeholders.

Tekir (2020) utilized the Bellon and Handler Model in a study focusing on the English as a Foreign Language (EFL) program within an English preparatory program at a state university in Turkey, established for over 15 years. The objective of the research was to assess whether the current program adequately caters to the diverse needs, language proficiency levels, and preparedness of learners, influenced in part by modifications to the high school program by the Ministry of National Education. Moreover, the study highlighted a significant number of students failing the English proficiency exam post-program completion, prompting the researcher to identify areas requiring improvement within the program. The results of the study showed that the program mostly achieved its objectives related to grammar, vocabulary, critical thinking, listening and reading, but both groups of participants were not sure about the attainment of listening and speaking objectives. Both parties expressed that there is incongruence between the first and the second semester in terms of listening skills and speaking skills related objectives were not attained. As a part of the findings of the evaluation study, the instructor participants recommended systematic teaching of speaking skills and allocation of particular class hours to improve speaking skill only. They underscored that this can be achieved by utilizing from a particular course book or from worksheets created by the program coordinators or material unit. When it comes to course conduct, both student and teacher participants provided positive feedback on the instruction. The evaluation indicated a student-centered approach. Some strengths of the instruction included use of a reasoning approach by the teachers, strategy training as well as the opportunities to speak English that the students provided with and use of various tactics such as various prompting, stimulation, questioning, illustrating, recitation, keeping all students active.

The latest curriculum evaluation study with Bellon and Handler model was carried out by Uckaya (2022). The researcher utilized the course evaluation questionnaire developed by Tekir (2020) along with the data obtained

from document analysis, course evaluation interviews and classroom observations. In this mixed method study, the curriculum was evaluated in accordance with course aims and objectives, course content and materials, instruction and formative assessment to find out what amendments are needed in the addressed curriculum. The results of the study indicated that the course objectives of listening and speaking courses did not serve the students' needs and the assessment tools were not compatible with the course objectives. The instruction is mostly teacher centered. Several suggestions were made to make the instruction more student centered, such as the integration of collaborative activities, and use of authentic materials which could serve the needs of the students better. Also, assessment components were suggested being revised to match the assessment practices with the objectives of the curriculum.

In line with the results of the study mentioned above, several suggestions have been made by the researcher to improve the curriculum. These suggestions can be summarized as a change in reading, listening and speaking objectives, more training to the teachers on the use of portfolio assessment, some changes in the objectives of the first and second semester and a change in the course book preference and the alignment of the first and second semester programs. As a reasonable suggestion for future studies, the researcher advised to include the perceptions of the administrators to the future studies.

To conclude, the research findings highlight the diverse applications of Bellon and Handler's Curriculum Evaluation Model across different educational contexts, primarily focusing on English Language Teaching (ELT) programs. The evaluation studies reveal a wide range of issues within ELT programs, such as discrepancies between curriculum objectives and student needs, the necessity for curriculum alignment, and the need for improved assessment methods. Across the studies, common suggestions for improvement include a shift towards more student-centered approaches, enhancements in the alignment of objectives and materials, and more effective use of formative assessments. Furthermore, the research indicates that while the Bellon and Handler model has led to actionable recommendations, the model has been scarcely used. The possible reasons for this scarcity are discussed in the discussion and conclusion chapter.

4. Discussion and Conclusion

As stated before, the number of curriculum evaluation studies conducted with Bellon and Handler Model is relatively low compared to the other models. Taking the elements and design of the model as well as the studies carried out with this model into consideration, a couple of probable causes for this issue have been established by the researchers.

First of all, the model suggested by Bellon and Handler (1982) is not a pure evaluation model. It merges curriculum evaluation and curriculum development. To elaborate, the four focus areas of the model namely goals, organization, operations, and outcomes are analyzed, and the data are gathered on these dimensions with only one aim in mind which is to determine the incongruence between anticipated and observed outcomes of the program so as to make necessary amendments in the curriculum for providing improved learning opportunities to the learners. Thus, this evaluation model can be labeled as an improvement focused one which was also stated by Mutlu and Şimşek (2018). On the other hand, many other curriculum evaluation models are employed as a framework to conduct curriculum evaluation so that a vital decision can be made about the program which is either continuing with the current program or terminating it. As this decision is a crucial one for various stakeholders, models with such an approach are more likely to be preferred. A similar argument can be made about Ornstein and Hunkins' (2004) classification and Bellon and Handler model. To elaborate, Ornstein and Hunkins (2004)

classify curriculum and evaluation models as Technical and Humanistic models. Technical models are more technical, and they are based on collecting data on student performance and comparing data against objectives. On the other hand, Humanistic approaches focus on more human interactions than outcomes, more quality than quantity and more "why" than "what". When we consider Bellon and Handler model we can say that the model has some components from both of these approaches. For this reason, it is extensively comprehensive. It does not just focus on one aspect or one method. Besides, the evaluation does not end when the results are obtained within the framework of the model. It reaches a conclusion only when necessary, adjustments for curriculum development have been suggested. For this reason, conducting a curriculum evaluation with this model is likely to demand more time and effort than the other models which might lead to a scarcity in its preference.

In addition, again owing to its one and only focus on curriculum development, Bellon and Handler Model does take part in any of the widely accepted classifications such as the classification of by Gredler (1996) and Fitzpatrick et al. (2004). In her classification, Gredler (1996) takes decision makers and stakeholders into account and categorizes the evaluation models in accordance with whom they serve. However, the Bellon and Handler model has not been designed to serve any stakeholders or decision makers. As curriculum evaluation is often conducted with a demand or need posed by stakeholders this model is very likely to be discarded. Similarly, the model is not a part of the classification of Fitzpatrick et al. (2004) as this classification has been made depending on the purpose of the evaluation. The model of Bellon and Handler subsumes some of the categories of Fiztpatrick et al. (2004) namely objectives-oriented evaluation approach, expertise-oriented evaluation approach and participant-oriented evaluation approach. However, the model does not purely concentrate on any of these categories. For this reason, it is relatively more difficult and time consuming to conduct a study with this model rather than any other model focusing on one aspect only such as objectives or participants. As curriculum evaluation is a process that requires staff and budget, it might seem more reasonable to curriculum evaluators and stakeholders to use a more cost-effective model rather than Bellon and Handler model.

Thanks to the review of literature conducted within this study, it has been seen that Bellon and Handler model of curriculum development and evaluation has been less utilized than the many other models of program evaluation and the majority of studies have been conducted in the context of Turkey which focuses on English language programs. On the other hand, although the number of studies utilizing this framework is limited, it is appropriate to say that the studies were quite comprehensive and the majority of them included both qualitative and quantitative data. Moreover, it has been revealed that all the studies collected data from both students and teachers. Thus, perspectives of both of these parties have been included in the studies. Also, some of the studies went a step further and included the perspectives of management and administrative personnel into the study which is likely to provide a much broader perspective to the evaluation.

The striking results and discrepancies in the results have been highlighted in all the studies. Based on the discrepancies between teachers' and students' reflections on the program, and the intended and realized outcomes of the program, various suggestions to improve the programs have been made in the studies. These suggestions did not only focus on the program and its materials, but they also focused on the effectiveness of the instruction. To illustrate, Yel (2009) underscored that the results indicate that teachers are not well-informed about the goals and objectives and the approaches of English language teaching program even though they are identified in the curriculum. For this reason, the activities and methods teachers use in the classrooms are not helpful in achieving the goals and objectives. Similarly, Al-Nwaiem (2012) pointed out that students are dissatisfied with the pre-

service ELT program particularly with monotonous, unchallenging and out-of-date textbooks and they are critical about teacher-centered, traditional instruction and examinations depending on rote-learned materials. As all the studies evaluated English language teaching curricula, it has enabled researchers to find out common problematic sides of the curricula under scrutiny. To illustrate, attainment of speaking objectives has been found problematic in many studies such as Al- Topçu (2005), Al-Nwaiem (2012), Mutlu (2018), Uckaya (2022) and Yel (2009) and suggestions to address this problem have been made. On the other hand, it is also possible to state that all four skills, namely reading, writing, speaking and listening as well as grammar and vocabulary have been addressed in the studies and numerous suggestions have been made to improve these components of the curricula. Some of these suggestions include providing a variety of tasks and activities (discussions, role-plays, debates, projects etc.) to the students which will provide opportunities for them to be more active in the classroom, incorporating reading strategies or strategic reading to the classroom content, and arranging more opportunities for students to exemplify the use of different grammar structures in various contexts.

When it comes to assessment practices, it has been interesting to see that many studies revealed inconstancies between teaching and testing and highlighted this point, which has been highlighted as one of the major aims of curriculum evaluation. Some of these studies are Mutlu (2018), Karcı et al., (2020) and Topçu (2005). In order to address this issue, a common suggestion has been made by the researchers, which is employing product-oriented evaluation, but also process-oriented evaluation should be included in the curriculum. That is, incorporating alternative assessment into curriculum has been recommended by the researchers, but Topçu (2005). put forward a warning on this issue and stated that standard procedures and practices among instructors or teachers should be ensured while employing alternative assessment methods. Otherwise, the methods do not serve their aims, and it creates frustration among students.

Another significant pattern which has been observed by the researchers within the studies is the inclusion of suggestions to benefit from technology as the date of the study becomes more recent. To elaborate, Mutlu (2018) and Karcı et al., (2020) are relatively newer studies and they both advocate the inclusion of more technology into curriculum as an answer to the problems revealed as a result of evaluation. Some of these suggestions are providing language labs, smart boards, a self-access center for students and the use of interactive software during the instruction.

On a final note, it is important to discuss the limited use of the Bellon and Handler (1982) model compared to other widely recognized models in the field, such as Tyler's (1949) objective-based model, Stufflebeam's (1971) CIPP model, and Scriven's (1991) goal-free evaluation model. Mutlu and Şimşek (2018) highlight that, despite its structured and comprehensive approach, the Bellon and Handler model has not been extensively utilized in empirical curriculum evaluation studies and the limited adoption of this model raises questions about its accessibility, applicability, and perceived advantages over other well-established frameworks.

One potential reason for the underutilization of the Bellon and Handler model could be its complex and multi-dimensional structure, which requires evaluators to engage in an ongoing interplay between status descriptions, analysis activities, and four key components of curriculum improvement. Unlike Tyler's straightforward, goal-oriented approach or Stufflebeam's decision-making model, Bellon and Handler's framework demands continuous feedback loops and adjustments throughout the evaluation process. This complexity may discourage researchers and practitioners from selecting it, particularly in large-scale evaluations where a more linear and standardized approach is preferred. Additionally, the model's emphasis on immediate and

iterative modifications rather than summative conclusions might make it less appealing for policymakers, stakeholders or institutions that require clear, outcome-based evaluation reports.

Another factor contributing to the model's lower prevalence could be its limited use in curriculum evaluation literature, empirical validation and a lack of standardized instruments for its implementation. Unlike the CIPP model, which has been widely referenced, tested, and adapted in different contexts, the Bellon and Handler model has received comparatively less theoretical and empirical attention. As a result, fewer researchers may be aware of its potential applications or may find it challenging to operationalize in diverse educational settings.

Thus, it can be concluded that the number of the studies with Bellon and Handler model is scarce, they were mostly carried out in the context of Turkey, and they are all on English language teaching programs. Besides, all the studies conducted with the model focus on the evaluation of curriculum for face-to-face instruction. However, as stated by Tekin (2022), there is a need to conduct more research on online education and evaluate curriculum for online instruction as the world has moved into a direction of online instruction after the COVID-19 pandemic. For this reason, there is a need to conduct more studies using this framework in other subjects than English and in contexts other than Turkey as also stated by Mutlu and Şimşek (2018). As the model enables the researcher to evaluate the program from various aspects, conducting evaluations with this model in different subjects can be illuminating to develop productive educational programs because despite its underrepresentation in curriculum evaluation research, the Bellon and Handler model presents unique advantages, particularly in its ability to accommodate dynamic and evolving educational contexts. Its focus on continuous improvement, iterative feedback, and flexible implementation aligns with contemporary views on adaptive and responsive curriculum evaluation. Future research could explore ways to refine and modernize the model, potentially increasing its relevance in the field.

Author Contributions

Conceptualization (CRediT 1)	Author-1 (%50) - Author-2 (%50)
Data Curation (CRediT 2)	Author-1 (%50) - Author-2 (%50)
Investigation - Analysis - Validation (CRediT 3-4-6-11)	Author-1 (%50) - Author-2 (%50)
Writing (CRediT 12-13)	Author-1 (%50) - Author-2 (%50)
Writing – Review & Editing (CRediT 14)	Author-1 (%50) - Author-2 (%50)

References

References marked with an asterisk indicate studies included in the literature review.

*Al-Nwaiem, A. (2012). An evaluation of the language improvement component in the pre-service ELT Programme at a College of Education in Kuwait: A case study. *(Unpublished doctoral dissertation), University of Exeter, Exeter, UK.*

Bellon, J. J., & Handler, J. R. (1982). *Curriculum development and evaluation: a design for improvement.* Dubuque, Iowa: Kendall/Hunt Publishing Company.

Brown, J.D. (1989). Language program evaluation: A synthesis of existing possibilities. *The second language curriculum,* 222-241.

*Erdem, H. E. (1999). Evaluating the English language curriculum at a private school in Ankara: A case study. (*Unpublished doctoral dissertation), Middle East Technical University, Ankara, Turkey.*

*Erozan, F. (2005). Evaluating the language improvement courses in the undergraduate ELT curriculum at Eastern Mediterranean University: A case study. (*Unpublished doctoral dissertation), Middle East Technical University, Ankara, Turkey.*

Fitzpatrick, J.L., Sanders, J.R., & Worthen, B.R. (2004). *Program Evaluation: Alternative Approaches and Practical Guidelines.* Boston, USA : Pearson Education Inc.

Gredler, M., E. (1996). *Curriculum Evaluation*. USA: Pearson Education Company.

*Karcı Aktaş, C., & Gündoğdu, K. (2020). An extensive evaluation study of the English preparatory curriculum of a foreign language school. *Pegem Eğitim Ve Öğretim Dergisi*, *10*(1), 169-214.

*Mutlu, G. (2018). A program evaluation study of the main course at a preparatory program: A case study. *Turkish Online Journal of Qualitative Inquiry*, *9*(3), 202-239.

Mutlu, G., & Şimşek, İ. (2018). Bellon ve Handler'in Program Geliştirme Modeli ve Modele İlişkin Özellikler. *Language Teaching and Educational Research*, *1*(2), 203-212.

Mutlu, G. (2020). A Look Into Curriculum Evaluation And Curriculum Evaluation Models. In Theory and Research in Educational Sciences II (pp.227-246).Gece Publishing.

Oliva, P. F. (1997). *Developing the curriculum*. (4th Ed.). USA: Longman.

Onwuegbuzie, A. J., & Frels, R. (2015). *Seven steps to a comprehensive literature review*. Los Angeles, CA: Sage.

Ornstein, A. C., & Hunkins, F. P. (2004). *Curriculum: Foundations, principles and issues.* Englawood Cliffs, NJ: Prentice Hall.

Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods* (3rd ed.). Sage Publications.

Scriven, M. (1967). *The methodology of evaluation. In Stake, R. E., et al. (eds.), AERA Mono graph Series on Curriculum Evaluation,* No. 1, Chicago: Rand-McNally.

Scriven, M. (1991). Evaluation thesaurus (4th ed.). Sage.

Silverman, D., & Marvasti, A. (2008). *Doing qualitative research: A comprehensive guide*. Thousand Oaks: Sage.

Stufflebeam, D. L. (1971, February 24). The relevance of the CIPP evaluation model for educational accountability. [Paper Presentation]. Annual Meeting of the American Association of School Administrators Atlantic City. New Jersey. http://eric.ed.gov/?id=ED062385

Tekin, E. (2022). Covid-19 Sonrasında Uzaktan (Çevrim içi) Yabancı Dil Öğretimi İlkeleri Üzerine Bir Araştırma. *Turkish Academic Research Review*, 7(3), 604-617. https://doi.org/10.30622/tarr.1150646

*Tekir, S. (2020). Evaluation of English as a Foreign Language Program in Turkey. *Academic Studies*, *183-209.*

*Topçu, T. (2005). An investigation of the effectiveness of the theme-based curriculum in the 2003-2004 academic year at the Department of Basic English at METU. *(Unpublished master's thesis), Middle East Technical University, Ankara, Turkey.*

Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. University of Chicago Press.

*Uçkaya, Z. (2022). Evaluation of an English preparatory program using Bellon and Handler Model. (*Unpublished master's thesis), Akdeniz University, Antalya, Turkey.*

Ward, J. G. (1987). An application and an analysis of the impact of the Bellon and Handler curriculum evaluation framework to the English as a Foreign Language program at the Huffco Indonesia intensive English Language program in Balikpapan.(Unpublished doctoral dissertation), University of Puerto Rico: Rio Piedras-Puerto Rico.*

White, J. P. (1971). The Concept of Curriculum Evaluation. *Journal of Curriculum Studies*, 3(2), 101-112.

*Yel, A. (2009). Evaluation of the effectiveness of English courses in Sivas Anatolian High Schools. (*Unpublished master's thesis), Middle East Technical University, Ankara, Turkey.*

*Yumuk, A. (1989). Systematic language program development and evaluation in Turkey. (*Unpublished master's thesis), Bilkent University, Ankara, Turkey.*

Zajda, J. (2024). Curriculum design and evaluation in the global culture. In *Handbook of Curriculum Theory, Research, and Practice* (pp. 783-797). Cham: Springer International Publishing.