Journal of the College of Education for Women
مجلة كلية التربية للبنات
June 30, 2025 [Vol. 36(2)]
P-ISSN: 1680-8738; E-ISSN: 2663-547X

An Evaluation of English-Medium Instruction Program in Secondary Schools for Distinguished Students

Israa Jaafar Naji [1] , **Shaima Mahdi Saalh** [2]

Department of English language, College of Education for Women, University of Baghdad, Iraq[1]
Department of English language, College of Education for Women, University of Baghdad, Iraq. (Board of Editors)[2]

israa.naji2203m@coeduw.uobaghdad.edu.iq[1]
shaimaa.mehdi@coeduw.uobaghdad.edu.iq[2]

https://doi.org/10.36231/coedw.v36i2.1816

Received: 20 March 2024; **Accepted**: 17 May 2024; **Published**: June 30, 2025

Abstract

The study aims to identify the level of agreement in using English as a medium of instruction according to study sample responses to the Context, Input, Process, and Product model questionnaire and evaluate the English as a medium of instruction program in secondary schools for distinguished students by the study sample according to the CIPP questionnaire items. The descriptive quantitative study included 109 teachers who actively attended the program in Iraq during the 2023–2024 academic year. To achieve the aims of this study, a five-point Likert scale questionnaire with 47 items based on Stufflebeam's (2015) CIPP model was prepared, validated, and administered to the participants in person and online. The data has been analyzed using SPSS. The findings have revealed that the teachers generally evaluated many aspects of the program positively. For example, the context of the course book prepares students scientifically and linguistically for their future. However, the study also revealed that (EMI) implementation poses challenges for teachers and students, including the absence of specific guidelines for EMI implementation and teachers' in-service training.

Keywords: CIPP model, English-Medium Instruction EMI, Evaluation, Secondary Schools for Distinguished Students, Stufflebeam

P-ISSN: 1680-8738; E-ISSN: 2663-547X مجلة كلية التربية للبنات

تقويم برنامج اللغة الإنكليزية واسطة للتعليم في المدارس الثانوية للمتميزين

اسراء جعفر ناجي[١] ، شيماء مهدي صالح[٢]

قسم اللغة الإنكليزية، كلية التربية للبنات، جامعة بغداد، العراق[١]
قسم اللغة الإنكليزية، كلية التربية للبنات، جامعة بغداد، العراق. (عضو هيأة التحرير)[٢]

israa.naji2203m@coeduw.uobaghdad.edu.iq[1]
shaimaa.mehdi@coeduw.uobaghdad.edu.iq [2]

https://doi.org/10.36231/coedw.v36i2.1816

تاريخ الإستلام: 20\3\٢٠٢٤، تاريخ القبول: 17\5\٢٠٢٤، تاريخ النشر الإلكتروني: ٣٠\6\٢٠٢٥

المستخلص :

تهدف الدراسة إلى التعرف على مستوى الاتفاق في استخدام اللغة الإنجليزية كوسيلة للتعليم وفقًا لاستجابات عينة الدراسة لاستبيان نموذج السياق والمدخلات والعمليات والمنتج وتقييم برنامج اللغة الإنجليزية كوسيلة للتعليم في المدارس الثانوية للطلاب المتميزين من قبل عينة الدراسة وفقًا لبنود استبيان CIPP. شملت الدراسة الكمية الوصفية ١٠٩ معلمًا حضروا البرنامج بنشاط في العراق خلال العام الدراسي ٢٠٢٣-٢٠٢٤. لتحقيق أهداف هذه الدراسة، تم إعداد استبيان مقياس ليكرت من خمس نقاط مكون من ٤٧ عنصرًا بناءً على نموذج Stufflebeam's (2015) CIPP والتحقق من صحته وإدارته للمشاركين شخصيًا وعبر الإنترنت. تم تحليل البيانات باستخدام برنامج SPSS. كشفت النتائج أن المعلمين قاموا عمومًا بتقييم العديد من جوانب البرنامج بشكل إيجابي. على سبيل المثال، يعمل سياق كتاب الدورة على إعداد الطلاب علميًا ولغويًا لمستقبلهم. ومع ذلك، كشفت الدراسة أيضًا أن تنفيذ (EMI) يفرض تحديات على المعلمين والطلاب، بما في ذلك عدم وجود إرشادات محددة لتنفيذ (EMI) وتدريب المعلمين أثناء الخدمة.

الكلمات المفتاحية: نموذج CIPP، اللغة الانكليزية واسطة للتعليم، تقويم، المدارس الثانوية للمتميزين، Stufflebeam

1. Introduction

English is the global language for research and publication and the principal medium for teaching in universities and educational institutions (Flowerdew & peacock, 2001). Many educational institutions have switched to English as their major language of instruction in recent years. According to research conducted by Crystal (2012) and Nunan (2003), this shift is caused by the current status of the English language globally. In recent years, many educational institutions have transitioned to using English as their primary language of instruction. This shift is due to the current global state of the English language, as highlighted by research conducted by Crystal (2012) and Nunan (2003).

In Asian countries, the rapid growth of social media has had an impact on how people perceive English. Many believe that English serves three primary purposes: promoting national development, gaining knowledge about other cultures worldwide, and engaging in global communication (Chang, 2011). In various countries, including Iraq, there is a growing trend to teach subjects using the English language in universities and secondary schools, as it is a universal language for effective global communication. This approach is called English as a Medium of Instruction (Altae, 2020). The EMI educational program teaches knowledge in English to improve both students' English and their knowledge-specific skills. Thus, the main objective of an EMI program is for students to acquire both language and content without having to choose between the two (Coleman, 2006). The mastery of English in the classroom is considered "ongoing challenge that is typical of the ordinary course of academic life" (Saalh and Kadhim, 2020: 229). Decision-makers in the Iraqi Ministry of Education have recognized the importance of English language education for distinguished students, "Language is the communication tool the society uses to negotiate common meaning and build social cohesiveness and communication surrounding ideas, behaviors, and activities" (Saalh and Esmaeel, 2022: 5). As a result, the Iraqi Ministry of Education has introduced English as a medium of instruction (EMI) to accommodate the students' future needs. Secondary schools for distinguished students have been teaching scientific subjects like physics, mathematics, chemistry, and biology in the English language since 2016 (Attia et al., 2018). This decision was taken in alignment with the English Medium Instruction (EMI) program for these subjects. The implementation of EMI policies in foreign language (FL) school environments is expected to improve academic and linguistic competition among students, thus enhancing their overall performance. Teacher proficiency, linguistically and professionally, has a core importance in any educational program (Saalh, 2014). Moreover, the teacher has a direct

impact on students' language (Mahdi, 2009). Therefore, the implementation of the EMI program faced several obstacles, one of which was the varying degrees of English proficiency among students and teachers (Anggraini, 2018). This problem is particularly recognized when students graduate to the secondary level and the medium of instruction changes from Arabic to English, even though English was taught as a subject in earlier grades. Answering questions in English can be a burden for some students, and teachers also recognize that some pupils may struggle to understand the material due to a lack of English competence (Floris, 2014). The researcher, as a teacher at the Secondary School for Distinguished Students, noticed the challenges that both students and teachers face during the implementation of the EMI program. To better understand this phenomenon, conducting scientific research to evaluate its characteristics is essential. Therefore, it is crucial to explore the obstacles that teachers encounter while incorporating the program from their perspectives. If we could gain a better understanding of the situations in these schools and language learning, it would eliminate confusion and significantly improve education. Academic researchers and evaluators have not conducted any research-based evaluations of the EMI program since its implementation in 2016–2017, except for an Attia et al. (2018) study by a group of experts from the Ministry of Education. The study assessed the use of EMI to teach scientific subjects in first-grade intermediate schools and sought to highlight the program's positive and negative aspects. The students showed a positive attitude during the experiment, but the study also revealed some challenges and obstacles that must be addressed to ensure the success of the program. However, the previous study limited its scope to the first year of the program's implementation, lacked specific data collection models, and was conducted in Arabic. Therefore, we conducted the current study to address the gaps and offer a more comprehensive evaluation of the program. The current study will evaluate the program after eight years of application in secondary schools for distinguished students in Iraq, relying on the teachers' perspective. The research employs the CIPP evaluation model developed by Stufflebeam in 197. This model is preferred over others because it encompasses a variety of evaluations, including context, input, process, and product evaluations, and is suitable for foreign language curricula. As a result, this study aims to provide answers to questions within this framework. Furthermore, to improve the English language program at Iraqi institutions, there is a need to conduct such an evaluation. Brown (1989) emphasized the importance of evaluation, as it ensures that other elements remain relevant. Program evaluation is crucial in both formal and online education settings. It involves collecting, analyzing, and interpreting data to determine the effectiveness and usefulness of a program (Lync, 1990).

Program evaluation research plays a crucial role in the field of education, as it offers several benefits. Firstly, it provides accurate data and detailed information about how students experience learning English in school. Secondly, it improves the program's overall effectiveness (Tayibnapis, 2008). As a result, this study aims to investigate the implementation and effectiveness of the English-medium instruction program in secondary schools for Distinguished Students, acknowledging the importance of program evaluation research.

The study aims to identify the teachers' perspectives on using English as a medium of instruction in secondary schools for distinguished students.

The study's significance lies in providing the program's stakeholders with a comprehensive understanding and deep insights into the program. The information gathered during the study could be used for various purposes, such as revising the program, ascertaining whether it is functioning as expected, and informing interested parties about the program. Furthermore, the researcher hopes that this type of investigation will have scientific value for future studies as well as for those interested in evaluating educational programs. The Ministry of Education could also use it as a framework for its program development studies.

2. Theoretical Framework

2.1 Key Words

2.1.1Evaluation

The evaluation was defined by Stufflebeam (2002, p.280) as "the process of delineating, obtaining, providing, and applying descriptive and judgmental information about the merit and worth of some object's goals, design, implementation, and outcomes to guide improvement decisions, provide accountability reports, inform institutionalization/ dissemination decisions, and improve understanding of the involved phenomena".

Moreover, Rallis and Bolland (2004) described program evaluation as a methodical investigation that offers data for decision-making and facilitates assessments of merit, worth, value, or significance. Sanders (1994) defined Program evaluation as a comprehensive and systematic assessment aimed at determining the worth or suitability of an educational activity, either on an ongoing basis or during a particular period. In accordance.

Worthen (1990) defined Program evaluation as "the determination of the worth of a thing consisting of those activities undertaken to judge the worth or utility of a program (or alternative programs) in improving some specified aspect of an educational system" (p.42). While researchers emphasize that program evaluation serves a variety of functions, one of its most important goals is to give information about the program during the evaluation process.

2.1.1.1 Educational Evaluation and its Purposes

Generally speaking, evaluating educational programs is an essential responsibility for anyone overseeing an education program. To Mertens (2019), program evaluation is a process of gathering and integrating evidence that results in a conclusion statement about the value, appropriateness, significance, or quality of a program, product, person, policy, proposal, or plan in the form of an empirical aspect (problem) and a normative aspect (statement of values).

Evaluation is a crucial part of the education process. Consistent with Rea-Dickins and Germaine (1992), evaluation plays a significant role in teaching and learning. They identified three key reasons for evaluation: accountability, development, and teacher development .

The first objective of evaluation is to determine whether a program is effective and efficient. This type of evaluation usually takes place after a program has been completed.

The second objective, which is development, aims to improve the quality of a program during its execution. It focuses on enhancing the skills of the program staff and teachers and involves collecting various kinds of data such as surveys, records, and interviews.

The third objective of evaluation is teacher development, which aims to formalize and broaden teachers' understanding of the teaching and learning processes in the classroom. It also raises awareness among teachers and practitioners about what happens in the classroom, rather than what is expected .

Further, Henry and Junles (2000) stated an acritical point of view that evaluation serves different social purposes, and it is a social action accompanied by a set of activities intended to facilitate people's understanding of policies and programs that aim to meet their needs.

Others like Cronbach (1963) highlighted the significance of evaluation as a means of improvement. He suggested that evaluation should be used to enhance educational programs while they are still ongoing, rather than just focusing on assessing the final product for market value.

2.1.1.2 Evaluation Models

Several evaluation models were developed and implemented in the field of educational program evaluation during the 1940s, 1950s, and 1960s. Among these, four models gained popularity and recognition, namely Tyler's objective model, Stake's responsive model, Scriven's Goal-Free model, and Stufflebeam's CIPP model.

- Tyler's Objective Model: Tyler (1949) suggested that evaluation serves the purpose of matching pre-behavioral objectives with actual outcomes. This indicates that evaluation involves assessing whether the intended outcomes have been achieved in practice. The process of evaluation helps to identify areas of success and areas that need enhancement, enabling individuals and administrations to modify and improve their strategies for better results.

- Stake's responsive model: Stake (2011) referred to his responsive model as sacrificing measuring precision to enhance the practical value of the findings for individuals involved with the program. A program is considered responsive if it prioritizes its actions over its objectives, meets the audience's requirements, and considers the different value perspectives while evaluating its success or failure (Stake, 1975).

- Scriven's goal-free model: This model is "a model in which official or stated program goals and objectives are withheld or screened from the evaluator" (Youker & Ingraham, 2013, p.51). Instead of concentrating on goals, goal-free evaluation looks at the results of the educational program, both those that were planned and those that weren't. The goal-free evaluator is unresponsive to what the teaching designers say about what they want to achieve. Instead, the goal-free evaluator looks at the results that the designers' educational program achieved (Popham, 1995).

- The CIPP model: The acronym CIPP, which stands for Context, Input, Process, and Product, was created by Daniel Stufflebeam in the late 1960s. This model has been widely tested and refined over time, with the initial aim of promoting accountability and establishing standards for educational programs in the United States (Stufflebeam, 2000, 2003). The CIPP model offers a comprehensive framework for evaluating programs and projects, taking into account their context and inputs, the processes used to implement them, and the resulting products or outcomes. Its flexibility and adaptability make it a valuable tool for assessing and improving a wide range of initiatives across different fields and sectors.

2.1.2 CIPP Stages

The model consists of four interconnected stages: context, input, process, and product. The four components have vital roles in the planning, implementation, and assessment of a project, as stated by Stufflebeam (2000, 2003).

- Context Evaluation

Stufflebeam (2000, p. 280) stated that "Context evaluations assess needs, problems, and opportunities as bases for defining goals and priorities, and

judging the significance of outcomes".

Context evaluation involves assessing requirements and needs in a specific environment to make informed decisions (Stufflebeam & Shinkfield, 1985). Clients utilize feedback from this evaluation to establish or clarify goals based on the intended beneficiaries' identified needs. They then analyze the degree to which the program's outcomes effectively meet the identified needs of the beneficiaries.

School districts can use context evaluation to interact effectively with the public and create a shared understanding of the district's benefits, drawbacks, requirements, potential clients, and critical problems. This approach can also help convince a funding agency of a project's value, set staff development goals, prioritise schools for support, and guide parents or advisors in identifying areas that need attention. (Gredler, 1996).

- Input Evaluation

The second phase of the model is input evaluation, which involves gathering data and determining the best way to allocate resources to achieve program objectives. As stated by Stufflebeam (2000), input evaluations are used to assess various approaches to meet needs and plan programs while allocating resources. To achieve its goals, an institution must conduct an input evaluation that includes a program description, as well as planning and allocation of resources required to achieve those goals.

Stufflebeam and Shinkfeld (1985) explained that input evaluation helps in decision-making by outlining and analyzing different procedural designs. In simpler terms, Stufflebeam (1971) argued that input evaluation aims to assist clients in exploring various options based on their specific needs and circumstances and supports them in developing an effective plan.

- Process Evaluation

Process evaluation provides feedback on the implementation of a program. The primary goal of this type of evaluation is to help staff identify problems with operations and fix the plan or activities as needed (Stuffelbeam et al., 2000).

It serves two main purposes: first, to provide information to external stakeholders who want to learn more about the program, and second, to help program evaluators interpret the program outcomes (Gredler, 1996).

Formative evaluation and process evaluation are closely related. Formative evaluation relies on data generated by activities in process evaluation. Process evaluation is an ongoing check on the plan, documenting the process and looking out for any changes, important steps that may be missing, or

mistakes made in the implementation of specific procedures (Stuffelbeam et al., 2000).

It should, however, also include instructions on how to change the program to meet participant needs better and reduce costs. At last, an "accountability report" requires product evaluation (Stufflebeam and Shinkfeld, 1985).

- Product evaluation

Product evaluation is a process that involves identifying and assessing the short-term, long-term, intended, and unintended results of a program or project. The primary aim of this kind of evaluation is to connect findings to objectives and evaluations of context, input, and process. Therefore, the evaluation of a product should assess how well the objectives were addressed and the overall impact of the programs (Stufflebeam, 2000, 2003, 2014).

The aims of the product evaluation just like summative evaluation are to compare the results achieved with the targeted needs. If possible, it also compares the results with those of other competitive programs. Moreover, it evaluates the context, inputs, and processes of the effort and guides whether to continue, modify, adapt, or terminate the effort based on these assessments (Stuffelbeam,2003).

2.1.2.١ Strengths and Drawbacks of the CIPP Model

The CIPP model is particularly beneficial for evaluation due to many factors:

Firstly, its broad history and continuing modifications. Given its non-specific nature, the model can be applied to a wide range of evaluation circumstances, regardless of the program or solution. According to Stufflebeam (2014), this tool is utilized by a wide range of professionals including evaluators, program specialists, researchers, developers, administrators, committees, task groups, and individuals at various levels of authority.

Secondly, The CIPP model is an effective evaluation tool due to its clear guidelines established by Daniel Stufflebeam and his colleagues. Stufflebeam (2000, 2003, 2014) provided detailed and comprehensive instructions for evaluation staff, as well as guidance on the appropriate timing, rationale, and methodology for its use.

Thirdly, the model includes four distinct dimensions (Context, Input, Process, and Product) that may be employed as a comprehensive framework for evaluating programs or projects. Alternatively, these dimensions can be employed individually to accommodate specific evaluation needs. Context

evaluation is utilized in the process of making planning decisions to establish objectives. Input evaluation assists in organizing decisions for designing instructional procedures. Process evaluation is employed to implement decisions, and monitor procedures. Product evaluation is applied in recycling decisions to evaluate and react to the outcomes resulting from actions (Popham, 1995).

The last point of the CIPP model's strengths is that it is considered a helpful tool for evaluators to formulate important questions during the evaluation process. They can create multiple inquiries for each component of the model. As per Harrison (1993), the model allows evaluators to intervene in the process before and throughout the program and also has the option to evaluate individual components separately.

Although CIPP is commonly applied, it is important to realize its limitations before implementing it for evaluation.

Firstly, it appears to have similarities to the need assessment. Context evaluation shares similarities with needs assessment since it also addresses needs.

Secondly, the implementation of the model is time-consuming when the entire model is utilized (Vo,2018).

Thirdly, the CIPP evaluation model has a potential weakness in the evaluator's occasional inability to respond to some significant questions or issues. Worthen, et al. (1997) stated that in planning evaluation procedures, evaluators need to reflect on the resources and period available. another model may have to be considered if this model requires more periods or resources than are available.

2.2 English-medium Instruction (EMI)

Dearden (2014, p.4) defined EMI as "The use of the English language to teach academic subjects in countries or jurisdictions where the first language (L1) of the majority of the population is not English". This definition refers to the practice of using the English language to teach subjects in school for instance mathematics, Physics, Chemistry and Biology in nations or regions where the majority of the population's first language is not English.

Similarly, Madhavan and McDonald (2014) considered that EMI is the teaching of subjects using the English language where English is not the formal language. There are many educational settings in which 'content' is taught in a language other than the home language of the students.

2.2.1 Reasons for Employing EMI in an Education Setting

In light of its position as a global lingua franca and its widespread use in academia, English has become an essential medium for instruction in several regions of the world. Teaching materials, such as scientific books and journals that are written only in English, are an important factor that strengthens English's place in the academy (Jenkins, 2007). The need for an EMI policy in higher educational and secondary school settings is apparent for many reasons.

Firstly, EMI is highly desirable due to the social value of English, particularly for students from mainstream, mother-tongue-based primary schools who possess a good command of the language. This has been observed by Kulung (2021) in his study.

Secondly, Kulung (2021) assumed that teaching students in their target language, at a young age, leads to better learning outcomes and effectiveness. During early life, students can easily absorb multiple languages without any confusion or mental competition, making it an ideal time to teach them.

Thirdly, using EMI in the secondary school curriculum equips students with the necessary skills to excel in university courses, particularly in the domains of medicine and engineering. This has been asserted by Altae (2020) in his study.

Fourthly, employing EMI is believed to enhance students' proficiency in the language, which is closely linked to a nation's economic progress. This makes it a desirable quality for governments to encourage, as highlighted by Ali (2013).

Fifthly, more EMI courses are needed to help incoming foreign students and local students who are preparing for future academic exchanges abroad. This is because academic exchange programs are becoming more common, as noted by Brumfit (2004).

Finally, students need to be prepared for professional settings where English is used as a language of interaction. English is used to communicate in trade, travel, banking, government agencies, and the sharing of culture, technology, and information. The purpose of EMI is to enhance students' proficiency in the skills required by the work sector, as observed by Coleman (2006) and Crystal (2004).

2.2.2 Benefits of Adopting the EMI Program

Numerous studies have explored the benefits of implementing English as a Medium of Instruction in certain educational settings, such as universities

and international schools:

- *Linguistic improvement and communication skills.* Previous studies have found that EMI is beneficial for improving English language skills, vocabulary, and proficiency; it also opens up more opportunities for international communication; and it has a positive effect on the larger community (Seitzhanova et al., 2017).

- *Access to academic resources.* Implementing EMI in education facilitates the enrollment of foreign students, enhances the university's rating, provides professional training for non-English speakers, fosters reflective practice, and improves learning and pedagogical aspects. Additionally, EMI programs ensure the availability of English materials and the existence of superior curricula (Karvonen, 2017).

- *Career advancement.* Implementing the EMI policy effectively helps people prepare for the competitive job market, reach their professional goals, pass internationally recognized English tests (like the TOEFL and IELTS), and advance their careers by opening up better job opportunities (Seitzhanova et al., 2017).

2.2.3 Challenges of Adopting the EMI Program

Numerous academic research studies have provided evidence that adopting English as the primary language of instruction presents several challenges :

- *lack of proficiency among Teachers.* A major challenge of adopting English as a language of instruction is that teachers, despite their academic qualifications, often lack the English language skills required to teach the language effectively (Dearden, 2014).

- *Lack of understanding among both teachers and students.* When the medium of communication in a classroom is a foreign language, there can be a lack of understanding or awareness of important issues among both teachers and students. Teachers may mistakenly attribute a student's difficulty in comprehending the language to a lack of intelligence or unwillingness to cooperate. (Alidou et al., 2006).

- *lack of support from the government.* This lack of support from the authorities for EMI manifests itself in various forms, including inadequate training for teachers, insufficient learning resources, and limited funding for schools. Consequently, this leaves schools and teachers with the daunting task of providing quality education to students who are expected to learn in a language that is not their first language (Karvonen, 2017).

2.3 Secondary Schools for Distinguished Students

Secondary Schools for Distinguished Students are educational institutions that accept students who have attained a primary school average of 95% or higher based on their knowledge and intelligence levels (Al-Azza, 2002). These schools are specialized in providing education to such students. Almost all teachers in these schools hold a Bachelor's degree, Master's degree, and some even have a Doctorate. In 2016/2017, the Iraqi Ministry of Education decided to use the English language to teach scientific subjects like Physics, Chemistry, Biology, and Mathematics in these schools, starting from the first intermediate grade until the sixth secondary grade graduation. This decision aimed to enhance academic standards for Distinguished students (Attia et al., 2018).

2.4 Related Works

Qomariah (2022) conducted a study entitled "Evaluation of the Implementation of English Medium Instruction" aimed to evaluate the implementation of the English medium instruction (EMI) educational program at Harapan Ibu Islamic Elementary School. It also aimed to provide recommendations for improvement of the program. The instruments used by the researchers to collect data are interviews, observation, and documentation. The evaluator used a structured method and interview guidelines to interview ten informants using the CIPP model. The study sample is chosen using the purposive sampling method. The overall participants in the study are ten informants: one principal, one vice principal, two teachers, two parents, two students and two alumni. The findings of this study indicate that: the suitability of the program context to the expectations of stakeholders is considered very good. The program's process is regarded as good, as indicated by the interaction between teachers and students using English. The products of the program are considered very good as shown by students' academic achievements. The program is regarded as good because it has adequate and good facilities, infrastructure, and standardized quality. The unavailability of guidelines and the absence of training or development for teachers are shortcomings that need to be improved.

Erdogan & Mede (2021) conducted a study entitled "Evaluating an English Preparatory Program Using the CIPP Model and Exploring the Motivational Beliefs for Learning" aimed at evaluating the English Preparatory School Language Program at a Turkish state university from the perceptions of students and instructors. The motivational beliefs of students for learning were also examined. The quantitative data were obtained through a Likert-scale questionnaire adapted from Stufflebeam's (1971) CIPP model was used to gather data from the students and instructors. Also, semi-structured

interviews were carried out in this study. The researchers used purposeful sampling to select 54 elementary-level students who speak Turkish as their mother tongue. Of these participants, 64.8% were male and 35.2% were female. In addition, 33 instructors with more than 2 years of teaching experience were included in the study. The study's findings indicated that speaking and listening skills needed to be developed, despite the participants' satisfaction with the materials, assessment, and teaching strategies. Furthermore, the students regarded control beliefs, intrinsic goal orientation, and task value as vital motivational beliefs. In conclusion, the program's strengths and weaknesses were identified through the perspectives of both students and academics.

The present study and the two previous studies mentioned are descriptive ones. The aims of Erdogan & Mede, 2021 and Qomariah, 2022 studies are similar to the aim of this study since they attempt to find out the teachers' perspectives on the EMI program. Regarding the study sample, Erdogan & Mede, 2021 was conducted in a university, while the current was conducted in secondary schools for distinguished students.

3. The Analytical Part

3.1 Methodology of the Study

It is worth mentioning that this research was a descriptive quantitative nature which employed a survey research design. A survey is conducted by asking participants a series of questions using a questionnaire to collect data.

3.2 Population and Sample

The population of this research was all the teachers who teach scientific subjects in 13 Iraqi secondary schools for distinguished students, except for the Kurdistan region, for the academic year 2023-2024. The sample of this research includes (109 teachers) from 13 Iraqi secondary schools for distinguished students.

3.3 Instrument

In the present study, a questionnaire based on Stufflebeam's (1971) CIPP model was prepared using a Likert scale. 48-item questionnaires which were divided into four dimensions of the CIPP model. The questionnaire included 11 items related to context, 10 items on input, 12 items on the process, and 15 other items evaluating the product.

3.3.1 Validity

The current study focuses on two types of validity: face and construct validity:

- Face validity: the initial version of the questionnaires of the current study is exposed to a jury of ten experts in the field of ELT to ensure

their face validity. The experts show their approval of the appropriateness of the items for the topic and sample concerned after some recommendations, some items are modified or removed in the formulation of almost every statement.

- Construct Validity: It refers to the extent to which a test measures the specific concept it was intended to evaluate. It is a crucial factor in determining the overall validity of a research method. The assessment of construct validity becomes even more important when researching something that cannot be directly measured or observed (Trafimow, 2022).

3.3.2 Pilot Administration

A sample of (20) participants was chosen randomly from the Al-Diwaniyah Secondary School for Distinguished Girls, City Centre, Diwaniyah Education Directorate to represent the pilot study. After conducting the pilot study, it became apparent that certain items were not easy to understand. As a result, some modifications were made to these items to reveal the ambiguity. Also, the pilot study revealed that the time needed for completing and submitting the responses ranges between (35-45) minutes to finish and send the responses.

3.3.3 Reliability

As far as the current study is concerned, two kinds of reliability are adopted:

- Test-Retest: Concerning the current study, The CIPP scales (teachers' questionnaire), are administered to a sample of (20) participants (the pilot sample) chosen randomly from the Al-Diwaniyah Secondary School for Distinguished Girls, City Centre, Diwaniyah Education Directorate. After two weeks, the same measures are applied to the same sample. Using Pearson correlation coefficient to estimate reliability or r- value, between the two sets of responses, results show that r- values for the CIPP scales are (0.89) which are indicators of good reliability since the values are higher than (0,70) (Litwin, 1995).
- Internal Consistency: The four scales of the study (context scale, input scale, process scale, and product scale) are found to be of high internal consistency with R-values of (0.76, 0.79, 0.77 and 0.79), respectively., all the values were observed to be above .60, which is generally considered acceptable for reliability.

3.4 Final Application

The questionnaire was administered to a total of 109 teachers who teach scientific subjects using the EMI program in Iraqi secondary schools for distinguished students. The survey was conducted offline and online using questionnaires. Thus, the data was obtained in two ways: Firstly, the

questionnaire was printed and administered to 41 participants personally at four separate schools in the Diwaniyah Education Directorate during the last week of December 2023. The entire process took approximately five successive weeks to complete. In the second week of February, the researcher administered questionnaires to participants in the study. The questionnaires were accessed via a link created using Google Forms, which was then shared with teachers from sixteen secondary schools across different governorates in Iraq through electronic media. The Google Form was open for four days, and during this time, 68 respondents participated. The participants were given ample time to read and consider the questions before submitting their responses.

3.5 Statistical Methods

To achieve the aims, the statistical methods used in the current study are calculated by the SPSS (the statistical package of the social sciences) version 24 package and these methods are:

1. **Means:** to identify the responses on each item of the questionnaire.
2. **Pearson correlation coefficient** is used to find out the item-total correlations for the CIPP scale .
3. **Alpha Cronbach:** equation to assess the internal consistency for the study instruments.
4. **Discrimination Power:** It measures the discrimination power of schema activation and literary analysis.

4. Presentation and Discussion of Results

The questionnaire findings were presented separately for each aspect, categorized under four stages of the CIPP evaluation model - context, input, process, and product. As shown in the following tables:

4.1 Context Evaluation

In the first section of the questionnaire, 11 items concern the program's context. The survey results indicate that, in general, teachers have a positive perception of most items. Considering the context, Teachers acknowledge the conciseness of the objectives and students' preliminary knowledge of the English language with the content of the program. In Table (1), the teacher's perspectives on the context aspect were presented. The statistics for this part revealed that the teachers, in general, had positive perceptions about the context part of the program.

To be more specific, the highest mean score was 3.95, indicating that most teachers agreed with item 1:

"The context of the course book is consistent with EMI program objectives in enhancing students' English language proficiency, preparing them for

higher education and the job market."

The lowest mean score was 2.37 which indicated that teachers disagreed with item 9:

"The context of the course book aligns the topics' difficulty with their duration in the annual calendar."

On the other hand, the average mean was 3.44, which indicated that teachers partially agreed with item 3:

"The context of the course book is based on accurate scientific translation from Arabic to English."

However, a number of the teachers believed that the context of the course book was not based on an accurate scientific translation from Arabic to English.

Table (1)

Teachers' Perspectives on Context Evaluation

Item Rank	Items	Strongly Agree	Agree	Undecided	Disagree	Strongly disagree	per cent.	Mean	Std. Deviation	T-value	Item direction
1	Context1	24	69	2	14	0	78	3.95	0.87	17.35	Agree
2	Context2	24	62	5	17	0	77.2	3.86	0.94	15.01	Agree
3	Context8	41	33	14	11	10	75.4	3.77	1.3	10.184	Agree
4	Context4	23	53	11	19	3	73.6	3.68	1.08	11.4	Agree
5	Context6	12	57	13	25	2	69.6	3.48	1.02	9.96	Agree
6	Context3	24	41	8	31	5	68.8	3.44	1.24	7.89	Agree
7	Context11	8	62	9	29	1	68.6	3.43	0.99	9.78	Agree
8	Context5	14	50	11	29	5	67.2	3.36	1.143	7.84	Undecided
9	Context10	10	43	13	39	4	63	3.15	1.121	6.03	Undecided
10	Context7	3	47	8	48	3	59.8	2.99	1.05	4.88	Undecided
11	Context9	1	25	10	50	23	47.4	2.37	1.09	-1.279	Disagree

4.2 Input Evaluation

In the second section of the questionnaire, 10 items relate to the input of the EMI program. Upon reviewing the results, it was found that teachers were generally confident about the input aspect of the program, including the tools and materials used during the lesson. However, they were undecided about some other items, such as the adequacy of educational conditions.

To be more specific, the highest mean score for the part was (3.71) which showed that teachers agreed with item 5:

"*Utilizing a technological platform in conjunction with traditional in-school lessons has the potential to improve student communication and facilitate idea sharing."*

The lowest mean score was 2.37 which indicated that teachers were undecided about item 6:

"The availability of adequate qualified teachers to fulfil the educational demands."

On the other hand, the mean average was 3.48, indicating agreement among teachers with item 1:

"The program allows the use of visual and auditory materials to understand the scientific subject."

Table (2)

Teachers' Perspectives on Input Evaluation

Item Rank	Items	Strongly Agree	Agree	Undecided	Disagree	Strongly disagree	Per cent.	Mean	Std. Deviation	T-value	Item direction
1	Input5	24	54	8	21	2	74.2	3.71	1.074	11.73	Agree
2	Input9	22	52	14	20	1	73.6	3.68	1.03	11.99	Agree
3	Input4	26	44	10	26	3	71.8	3.59	1.17	9.68	Agree
4	Input3	12	64	10	21	2	71.6	3.58	0.98	11.44	Agree
5	Input2	10	65	10	22	2	70.8	3.54	0.98	11.12	Agree
6	Input1	14	51	18	25	1	69.6	3.48	1.01	10.05	Agree
7	Input10	18	48	8	31	4	68.4	3.42	1.17	8.13	Agree
8	Input8	24	36	9	35	5	67.2	3.36	1.27	7.08	Undecided
9	Input7	9	31	13	41	15	64	3.2	1.23	5.95	Undecided
10	Input6	11	36	19	32	11	60.8	3.04	1.2	4.66	Undecided

4.3 Process Evaluation

In the third section of the questionnaire, 11 items concern the program's Process. Analyzing the findings of the process, it can be seen that, the teachers had positive attitudes towards the way delivering materials, creating a learning atmosphere and performance to provide more active participation of the students. This indicated that the teaching methodology of the program was well-accepted and considered a positive point. The

highest mean score was (4) which showed that teachers agreed with item 11:

"The formative tests administered in English are clear to the students."

In contrast, item 7 received the lowest mean score (2.1), indicating disagreement :

"Mixing English and Arabic distracts students understanding."

The mean average was (3.72), which showed that teachers agreed with item 8:

"Teachers can effectively communicate their ideas and promote students' understanding by using the English language."

Table (3)

Teachers' Perspectives on Process Evaluation

Item Rank	Items	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree	Per cent.	Mean	Std. Deviation	T-value	Item direction
1	Process11	31	61	5	10	2	80	4	0.93	16.79	Agree
2	Process4	24	67	5	10	3	78.2	3.91	0.94	15.67	Agree
3	Process1	22	65	3	17	2	76.2	3.81	0.99	13.72	Agree
4	Process10	25	58	6	17	3	75.6	3.78	1.06	12.64	Agree
5	Process2	16	67	10	14	2	74.8	3.74	0.927	14	Agree
6	Process8	19	55	21	13	1	74.4	3.72	0.92	13.74	Agree
7	Process3	15	64	10	19	1	73.4	3.67	0.95	12.81	Agree
8	Process6	11	58	14	20	6	68.8	3.44	1.08	9.13	Agree
9	Process5	9	32	15	49	4	58.8	2.94	1.108	4.11	Undecided
10	Process9	18	39	11	39	18	54.2	2.71	1.17	1.84	Undecided
11	Process7	0	14	12	54	29	42	2.1	0.94	-4.422	Disagree

4.4 Product Evaluation

The fourth section of the questionnaire includes 15 items that pertain to the program's product evaluation. This evaluation is the final stage of the model and involves gathering data about the outcome of the EMI program. The results of this section revealed that teachers expressed positive and negative perspectives on the outcome of using English as a medium of instruction in secondary school for distinguished students, with challenges facing them. The highest mean score was (4.32) which showed that teachers strongly agreed with item 1:

"*The EMI program improved students' vocabulary in math, chemistry, biology, and physics."*

On the contrary, the lowest mean score (1.95), indicated disagreement with item 13:

"*The EMI program does not place an additional burden on teachers, or increase their workload."*

The mean average was (3.67), which showed that teachers agreed with the item7:

"The EMI program facilitates lesson progression and fosters a positive classroom environment."

Table (4)

Teachers' Perspectives on Product Evaluation

Item Rank	Items	Strongly Agree	Agree	Undecided	Disagree	Strongly disagree	per cent.	Mean	Std. Deviation	T-value	Item direction
1	Product1	48	53	5	1	2	86.4	4.32	0.77	24.75	Strongly agree
2	Product12	32	66	5	4	2	82.4	4.12	0.8	21.08	Agree
3	Product2	24	70	9	3	3	80	4	0.817	19.18	Agree
4	Product4	28	63	8	7	3	79.4	3.97	0.92	16.76	Agree
5	Product5	27	65	5	6	6	78.6	3.93	1	14.8	Agree
6	Product6	22	64	11	8	4	76.8	3.84	0.95	14.71	Agree
7	Product8	22	56	21	8	2	76.2	3.81	0.91	15.04	Agree
8	Product7	13	64	17	13	2	73.4	3.67	0.9	13.52	Agree
9	Product10	14	62	12	18	3	72.2	3.61	1	11.54	Agree
10	Product9	4	46	7	37	15	57.6	2.88	1.21	3.29	Undecided
11	Product11	4	43	9	30	23	55.4	2.77	1.27	2.22	Undecided
12	Product3	1	31	14	48	15	51.2	2.56	1.07	0.85	Disagree
13	Product14	3	30	11	41	24	50.2	2.51	1.191	0.121	Disagree
14	Product15	5	22	8	48	26	47.6	2.38	1.185	-1.09	Disagree
15	Product13	1	12	6	52	38	39	1.95	0.97	-5.9	Disagree

5. Conclusion

The results of the study remarked some important information concerning the program. To begin with, the teachers perceived that many aspects of the program functioned well and there are positive and negative teachers' perspectives regarding using English as the medium of instruction at Iraqi secondary schools distinguished students.

The dimension on which teachers had the most positive opinions was

determined as context. The study's quantitative findings show that teachers acknowledge the conciseness of the objectives and students' preliminary knowledge of the English language with the content of the program, preparing them for higher education and the job market. Teachers are satisfied that the context of the course book prepares students scientifically and linguistically for their future lives besides enhancing students' analytical and critical thinking abilities. However, the most negative opinions were related to the duration of the academic course, which teachers believe is inadequate. Students need more time is needed to overcome the topics' difficulty and complete the content.

As quantitative data suggests, teachers had generally positive remarks on the input dimension, teachers were confident in the input aspect of the EMI program such as using visual and auditory materials, presentations, and extra materials, along with a technological platform used during the lesson. However, they are undecided about some other items, such as the availability of adequate support, including the teacher's guidebook, scientific textbooks, continuous in-service training by the Ministry of Education, and qualified staff to teach scientific subjects using the English language.

Teachers generally expressed positive opinions about the process dimension in the quantitative findings, it can be seen that they were satisfied with the way the materials were delivered, the learning atmosphere created and the performance that provided more active participation of the students. This indicates that the teaching methodology of the program was well-accepted. However, teachers perceived negative consequences when it came to effectively communicating their ideas and promoting students' understanding by using only the English language in the class. This means that teachers often allow the use of Arabic to ensure student comprehension.

When it comes to product dimension, data has been approved that the EMI curriculum is generally effective However, it is inefficient in solving certain problems. This suggests that the language skills and scientific vocabulary that students have acquired throughout the program are sufficient to prepare them for academic life at university. Additionally, most students can understand exam instructions and questions. It is worth noting that there are several problems with the product dimension. Firstly, there is a lack of proficiency in English among both students and teachers. Secondly, the EMI program has placed an additional burden on teachers causing an increase in their workload. Lastly, the literary translation of scientific subject curricula from Arabic to English has resulted in incoherent scientific content. These problems need to be addressed before we can determine whether the EMI program is successful.

Here, the answers to the second research question represents the conclusions arrived at regarding this research question.

6. Recommendations of the Study

In light of the study results, the following recommendations are suggested:

1. The Ministry of Education should pay attention to employing teachers with comprehensive training programs in the same governorate where they reside before the start of the course.
2. The contextualization of language for English-medium instruction (EMI) courses must be considered to avoid memorization-based learning, especially for terminology courses. A revision of foreign language policy and planning may be necessary and officially announced.
3. The Ministry of Education needs to address the issue of inadequate supplementary materials, such as the teacher's guidebook, scientific textbooks, and laboratory equipment. and provide adequate support to students.

7. Suggestions of the Study

In the light of the findings of the study, further studies need to be undertaken as follows:

1. Investigating the perspectives of students and parents.
2. Investigating the instructional problems with experimental, correlational and curriculum evaluation research is important for deeper understanding.
3. Conduct further studies with a bigger sample size for all the secondary schools for distinguished students in Iraq, the study would be more comprehensive.
4. Conduct further studies with a larger scale and a variety of research instruments being combined to gain deeper insights into the nature of EMI.
5. Investigating the impact of the EMI phenomenon on society, that encompasses both sociological and pedagogical aspects.

References

Al-Azza, S. H. (2002). *Raising the gifted and talented*. Amman International House, and Dar Al-Thaqafa for Publishing and Distribution.

Ali, N. L. (2013). A changing paradigm in language planning: English-medium instruction policy at the tertiary level in Malaysia. *Current Issues in Language Planning, 14* (1), 73-92.

Alidou, H., Boly, A., Brock-Utne, B., Diallo, Y. S., Heugh, K., & Wolff, H. E. (2006). *Optimizing learning and education in Africa: The language factor*. Paris: ADEA.

Altae, M. (2020). An overview of the stages of development of the Iraqi

English language curriculum. *Social Sciences & Humanities Open, 2*(1), 100047.

Anggraini, D. R. (2018). Challenges in teaching English to young learners. *LINGUA: Journal Bahasa dan Sastra, 18*(2), 141–146. https://ejournal.unsri.ac.id/index.php/lingua/article/view/11017.

Attia, R. K., Abd al-Hussein, A., Janani, M., H., & Hamoudi, S., T. (2018). An evaluation study of the experience of teaching scientific subjects in the first grade of intermediate schools for distinguished students in the English language. *Dirasat Tarbawiya, 11* (41).

Brumfit, C. J. (2004). Language and higher education: Two current challenges. *Arts and Humanities in Higher Education, 3*(2), 163-173.

Chang, Y. Y. (2011). English-medium instruction for subject courses in tertiary education: Reactions from Taiwanese undergraduate students. *Taiwan International ESP Journal*, *2*(1), 55-84.

Coleman, J. A. (2006). English medium instruction in European higher education. *Language Teaching*, *39*(1), 1-14.

Cronbach, L. J. (1963). Course improvement through evaluation. *Teachers college record*, *64*(8), 1-13.

Crystal, D. (2012). *English as a global language*. Cambridge University Press.

Crystal. D. (2004). The past, present and future of World English. In A Gardt & B. R. Huppauf (Eds.), *Globalization and the future of German* (pp. 27-45). Berlin: Mouton de Gruyter.

Dearden, J. (2014). *English as a medium of instruction-a growing global phenomenon*. British Council.

Erdogan, G., & Mede, E. (2021). Evaluating an English Preparatory Program Using CIPP Model and Exploring the Motivational Beliefs for Learning. *Journal of Education and Educational Development, 8*(1), 53-76.

Floris, F. D. (2014). Learning subject matter through English as the medium of instruction: Students' and teachers' perspectives. *Asian English, 16*(1), 47-59.

Flowerdew, J., and Peacock, M. (2001). (Eds.). *Research perspectives on English for academic purposes*. Cambridge: CUP.

Gredler, M. E. (1996). *Program evaluation*. NJ: Prentice Hall.

Harrison, A. S. (1993). *An evaluation model for middle school counselling and guidance*. Old Dominion University.

Henry, G., & Junles, N. (2000). *Making sense of evaluation: A reader in evaluation theory and practice*. Thousand Oaks, CA: Sage Publications. Models of Evaluation: An Overview.

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford University Press.

Karvonen, H. (2017). *English as a medium of instruction benefits and challenges as viewed by founders of international schools in Ethiopia*. (Unpublished MA Thesis), University of Turku, Turku, Finland.

Kulung, B. R. (2021). *English as a medium of instruction at secondary level education: Challenges and pitfalls* (Doctoral dissertation, Department of English).

Madhavan, D., & McDonald, J. (2014). *Webinar: English as a medium of instruction (EMI): Philosophies and policies*. Retrieved from http://www.oecd.org/edu/imhe/foodforthoughtenglishasamediumofinstructionwebinar.

Mahdi, S. (2009). Self-efficacy of English language teachers and its effect on their pupils' achievement in Baghdad secondary schools/Karkh2. *Journal of Educational and Psychological Researches, 6*(21), 374-390.

Mertens, D. M. (2019). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (5th ed.). SAGE Publication.

Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia- Pacific Region. *TESOL Quarterly, 37*(4), 589-613.

Popham, W. J. (1995). *Educational evaluation* (3rd ed.). Boston: Allyn and Bacon.

Qomariah, A. R., Driana, E., & Setiadi, H. (2022). Evaluation of the implementation of English medium instruction. *Journal Penelitian dan Evaluasi Pendidikan, 26*(1).

Rea-Dickins, P. & Germaine, K. (1992). *Evaluation*. Oxford: Oxford University Press.

Saalh, S. M., & Esmaeel, D. H. (2022). EFL Student- Teachers' Perception in the Culture of Thinking. *Journal of the College of Education for Women, 33*(2), 1–12. https://doi.org/10.36231/coedw.v33i2.1584

Saalh, S.M. & Kadhim, S.H. (2020) The EFL Students' Academic Buoyancy in Reading and Listening Skills. *Asian EFL Journal. 27*(4.4), 226-253. http://www.scopus.com/inward/record.url?eid=2-s2.0-85096076667&partnerID=MN8TOARS_

Saalh, S.M. (2014). The effect of combining video lectures and Kolb experiential learning on EFL student-teachers' ability to teach communicative course book and their teaching competencies. *Journal of College of Education for Women, 25*(4), 1139- 1156.

Sanders, J. R. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). SAGE Publications.

Seitzhanova, A., Plokhikh, R., Baiburiev, R., &Tsaregorodtseva, A. (2017). English as the medium of instruction: Modern tendency of education in Kazakhstan. *PRADEC Conference Proceedings, 3* (1), 74-77.

Stake, R. E. (1975). *Evaluating the arts in education: A responsive approach.* Columbus: Charles E. Merill Publishing Company.

Stake, R. E. (2011). Program evaluation particularly responsive evaluation. *Journal of Multidisciplinary Evaluation, 7*(15).

Stufflebeam, D. L. (1971). The use of experimental design in educational evaluation. *Journal of Educational Measurement, 8*(4), 267-274.

Stufflebeam, D. L. (2000). The CIPP model for evaluation. *In Evaluation models: Viewpoints on educational and human services evaluation* (pp. 279-317). Dordrecht: Springer Netherlands.

Stufflebeam, D. L. (2002). The CIPP model for evaluation. In: D. L. Stufflebeam, G. F. Madaus, T. Kellaghan (Eds.). *Evaluation models: Viewpoints on educational and human services evaluation.* (2nd ed). (pp. 279-317). New York, Kluwer Academic Publishers.

Stufflebeam, D. L. (2003). The CIPP model for evaluation. In: T. Kellaghan & D. L. Stufflebeam (Eds.), *The international handbook of educational evaluation*. (pp.31-62). Boston, Kluwer Academic Publishers.

Stufflebeam, D. L. (2014). *CIPP model for evaluation: An improvement and accountability oriented approach*. DL Stufflebeam & CLS Coryn, EVALUATION THEORY, MODELS, AND APPLICATIONS, 309.

Stufflebeam, D. L., & Shinkfield, A., J. (1985). *Systematic evaluation: A self-instructional guide to theory and practice.* Boston, Kluwer Academic Publisher.

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

Worthen, B. (1990). Program evaluation. H. Walberg & G. Haertel (Eds.), The International encyclopedia of educational evaluation. (pp. 42-47). Toronto, ON: Pergamon Press.

Youker, B. W., & Ingraham, A. (2013). Goal-Free Evaluation: An Orientation for Foundations' Evaluations. *The Foundation Review, 5*(4). https://doi.org/10.9707/1944-5660.1182.