



مجلة الجمعية السعودية العلمية للمعلم
Journal of the Saudi Scientific Association for the teacher

دورية علمية نصف سنوية - محكمة

المجلد الثالث- العدد الأول

رجب ١٤٤٧هـ - يناير 2026م

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

المشرف العام

رئيس جامعة الملك خالد

أ. د. فالح بن رجاء الله السلمي

نائب المشرف العام

وكيل الجامعة للدراسات العليا والبحث العلمي

أ. د. حامد مجدوع القرني

المشرف على وحدة المجلات والجمعيات العلمية

أ. د. محمد سحيم أبو حسان

رئيس هيئة التحرير

أ. د. عبدالله بن علي آل كاسي

رئيس هيئة التحرير

أ. د. عبدالله بن علي آل كاسي

أستاذ المناهج وطرق تدريس العلوم بجامعة الملك خالد

هيئة التحرير

أ. د. أحمد بن محمد سعد الحسين

أستاذ المناهج وطرق تدريس الدراسات الاجتماعية والوطنية بجامعة الإمام محمد بن سعود الإسلامية

أ. د. محمد بن عبدالله محمد عسيري

أستاذ علم النفس التربوي بجامعة تبوك

أ. د. مرضي بن غرم الله الزهراني

أستاذ المناهج وطرق تدريس اللغة العربية بجامعة أم القرى

أ. د. محمد بن زيدان عبدالله آل محفوظ

أستاذ المناهج وطرق تدريس العلوم الشرعية بجامعة الملك خالد

أ. د. محمد بن يحيى صفحي

أستاذ التربية الخاصة بجامعة جازان

أ. د. عائشة بنت بليهش العمري

أستاذ تقنيات التعليم جامعة طيبة

مدير التحرير

أ. د. عاصم محمد إبراهيم عمر

أستاذ المناهج وطرق تدريس العلوم بجامعة سوهاج، مصر

الهيئة الاستشارية

أ. د. إبراهيم بن عبدالله بن إبراهيم العبيد

جامعة الإمام محمد بن سعود الإسلامية

أ.د. صالح بن يحيى بن مفرح الزهراني

جامعة جدة

أ. د. خالد عبد اللطيف محمد عمران

جامعة سوهاج

أ. د. مفرح بن سعيد صالح آل كردم

جامعة الملك خالد

Dr. Michael Brody

Montana State University

أ. د. سامي بن فهد بن راشد السندي

جامعة القصيم

أ.د. حمد بن عبدالله بن مطلق القميري

جامعة الأمير سطام بن عبدالعزيز

د. محمد محمود محمد القسيم

الجامعة الهاشمية

أ. د. راشد حسين محمد العبدالكريم

جامعة الملك سعود

أ.د. ناصر عبدالله ناصر الشهراني

جامعة الملك خالد

معلومات عامة عن المجلة وتاريخ التأسيس:

دورية علمية محكمة نصف سنوية تصدر عن الجمعية السعودية العلمية للمعلم (جسم) بجامعة الملك خالد. تنشر إسهامات الباحثين في مجال التعليم والتعلم، وبصورة خاصة كل ما يتعلق بإعداد المعلم وتطويره المهني. وتهدف المجلة إلى تحقيق العديد من الأهداف أهمها: الإسهام في نشر المعرفة من خلال طرح ودراسة القضايا المتعلقة بالتعليم والتعلم، ونشر البحوث والدراسات العلمية المحكمة في مجال التعليم والتعلم، وإيجاد قناة نشر علمية تخدم الباحثين في شتى المجالات المتعلقة بالمعلم وبرامج إعداده وتأهيله وتطويره، والإسهام في عرض وتحليل وقراءة الكتب في مجال التعليم والتعلم والمتعلقة برسالة المجلة وأهدافها. وقد تأسست المجلة في عام ١٤٤٤هـ بموافقة مجلس جامعة الملك خالد في اجتماعه الثالث بتاريخ ٧ / ٤ / ١٤٤٤هـ بالقرار رقم (٤٤/٣/١١) المتضمن الموافقة على إنشاء المجلة، وتشكيل هيئة تحريرها اعتباراً من ١ / ١ / ٢٠٢٣م.

رؤية المجلة:

التميز والريادة في نشر الأبحاث والدراسات في مجال التعليم والتعلم.

الرسالة:

نشر الأبحاث والدراسات العلمية المحكمة في مجال التعليم والتعلم وفق المعايير العلمية للنشر.

الأهداف:

١. الإسهام في نشر المعرفة من خلال طرح ودراسة القضايا المتعلقة بالتعليم والتعلم.
٢. نشر الأبحاث والدراسات العلمية المحكمة في مجال التعليم والتعلم.
٣. إيجاد وعاء نشر علمي يخدم الباحثين في شتى المجالات المتعلقة بالمعلم وبرامج إعداده وتأهيله وتطويره.
٤. الإسهام في عرض وتحليل الكتب وملخصات رسائل الماجستير والدكتوراه في مجال التعليم والتعلم.

الشروط، والقواعد، والتعليمات، والحقوق، والإجراءات الخاصة بالنشر في المجلة:

أولاً: الشروط والقواعد الخاصة بالنشر في المجلة:

١. أن يقع البحث ضمن أحد مجالات النشر بالمجلة.
٢. خلو البحث من الأخطاء اللغوية والنحوية.
٣. أن يسهم البحث في تنمية الفكر التربوي وتطوير تطبيقاته محلياً أو عربياً أو عالمياً.
٤. أن يلتزم الباحث في بحثه بأخلاق البحث العلمي، وحقوق الملكية الفكرية.
٥. ألا تزيد نسبة الاستدلال العلمي باستخدام برنامج iThenticate عن (٢٠٪).
٦. لا تتم كتابة اسم الباحث أو الباحثين في متن البحث صراحةً، أو بأي إشارة تكشف عن هويته أو هويتهم، ويمكن استخدام كلمة الباحث أو الباحثين بدلاً من ذلك.
٧. الآراء الواردة في البحوث المنشورة تعبر عن وجهة نظر الباحثين فقط، ولا تعبر بالضرورة عن رأي المجلة.

ثانياً: تنظيم البحث

أ. البحوث التطبيقية:

يورد الباحث أو الباحثون مقدمة تبدأ بعرض طبيعة البحث، ومدى الحاجة إليه، ومسوغاته، ومتغيراته، متضمنة الدراسات السابقة بشكلٍ مدمج دون تخصيص عنوان فرعي لها. يلي ذلك استعراض مشكلة البحث، ثم تحديد أهدافه، وبعد الأهداف تورد أسئلة البحث أو فروضه. ثم تعرض منهجية البحث؛ مشتملةً على: مجتمع البحث، وعينته، وأدواته، وإجراءاته، متضمنةً كيفية تحليل بياناته. ثم تعرض نتائج البحث ومناقشتها وتفسيرها، والتوصيات المنبثقة عنها.

ب. البحوث النظرية:

يورد الباحث أو الباحثون مقدمةً يمهّد فيها للفكرة المركزية التي يناقشها البحث، مبيّنًا فيها: أديبات البحث، وأهميته، وإضافته العلمية إلى مجاله. ثم يعرض منهجية بحثه، ومن ثم يُقسّم البحث إلى أقسام على درجة من الترابط فيما بينها، بحيث يعرض في كل منها فكرة محددة تكوّن جزءًا من الفكرة المركزية للبحث. ثم يختتم البحث بم خلاصة شاملة متضمنةً أهم النتائج التي خلص إليها البحث.

ثالثاً: التوثيق

توضع قائمة المراجع في نهاية البحث باتباع أسلوب التوثيق المعتمد في المجلة.

١. أن يكون التوثيق في متن البحث وقائمة المراجع وفق نظام جمعية علم النفس الأمريكية (APA) الإصدار السابع.
٢. يلتزم الباحث بترجمة أو رومنة^١ (Romanization /Transliteration) توثيق المقالات المنشورة في الدوريات العربية الواردة في قائمة المراجع العربية (مع الإبقاء عليها في قائمة المراجع العربية)، وفقاً للنظام التالي:
 - أ) إذا كانت بيانات المقالة المنشورة باللغة العربية الواردة في قائمة المراجع (التي تشمل اسم، أو أسماء المؤلفين، وعنوان المقالة، وبيانات الدورية) موجودة باللغة الإنجليزية في أصل الدورية المنشورة بها، فتكتب كما هي في قائمة المراجع، مع إضافة كلمة (In Arabic) بين قوسين بعد عنوان الدورية.
 - ب) إذا لم تكن بيانات المقالة المنشورة باللغة العربية موجودة باللغة الإنجليزية في أصل الدورية المنشورة بها، فيتم رومنة اسم، أو أسماء المؤلفين، متبوعة بسنة النشر بين قوسين، ثم يتبع بعنوان المقالة إذا كان متوافقاً باللغة الإنجليزية في أصل المقالة، وإذا لم يكن متوافقاً فتتم ترجمته إلى اللغة الإنجليزية، ثم يتبع باسم الدورية التي نشرت بها المقالة باللغة الإنجليزية إذا كان مكتوباً بها، وإذا لم يكن مكتوباً بها فيتم ترجمته إلى اللغة الإنجليزية. ثم تضاف كلمة (In Arabic) بين قوسين بعد عنوان الدورية.
 - ج) توضع قائمة بالمراجع العربية بعد المتن مباشرةً، مرتبة هجائياً حسب الاسم الأخير للمؤلف الأول، وفقاً لأسلوب التوثيق المعتمد في المجلة.
 - د) يلي قائمة المراجع العربية، قائمة المراجع الإنجليزية، متضمنةً المراجع العربية التي تم ترجمتها، أو رومنتها، وفق ترتيبها الهجائي (باللغة الإنجليزية) حسب الاسم الأخير للمؤلف الأول، وفقاً لأسلوب التوثيق المعتمد في المجلة.

الجبر، سليمان. (١٩٩١م). تقويم طرق تدريس الجغرافيا ومدى اختلافها باختلاف خبرات المدرسين وجنسياتهم وتخصصاتهم في المرحلة المتوسطة بالملكة العربية السعودية. مجلة جامعة الملك سعود - العلوم التربوية، ٣(١)، ١٤٣ - ١٧٠.

Al-Jabr, S. (1991). The evaluation of geography instruction and the variety of its teaching concerning the experience, nationality, and the field of study at intermediate schools in the Kingdom of Saudi Arabia (in Arabic). *Journal of King Saud University-Education sciences*, 3(1), 143-170.

رابعاً: تعليمات النشر في المجلة

● يلزم تنسيق البحث تبعاً لما يلي:

١. لا يتجاوز البحث المقدم للنشر (٣٠) ثلاثين صفحة، وبما لا يزيد عن (٨٠٠٠) ثمانية آلاف كلمة.
٢. أن يتضمن البحث ملخصين: أحدهما باللغة العربية، والآخر باللغة الإنجليزية، بشرط ألا يزيد أي منهما عن (٢٥٠) كلمة، وأن يكتب كل منهما في صفحة مستقلة، متبوعاً بكلمات مفتاحية لا تزيد عن خمس كلمات تعبر عن محاور البحث.
٣. تكون أبعاد جميع هوامش الصفحة (٢,٥) سم، ما عدا الهامش الأيمن (٣,٥) سم، والمسافة بين الأسطر والفقرات "مفرد"

^١ (يقصد بالرومنة: النقل الصوتي للحروف غير اللاتينية إلى حروف لاتينية، تمكن قراء اللغة الإنجليزية من قراءتها، أي: تحويل منطوق الحروف العربية إلى حروف تنطق بالإنجليزية).

٤. الخط المستخدم في المتن للكتابة باللغة العربية (Traditional Arabic) بحجم (١٦)، وللكتابة باللغة الإنجليزية (Times New Roman) بحجم (١٢)، وتكون العناوين الرئيسية في اللغتين بولد (Bold).
٥. يكون نوع الخط المستخدم في الجداول والأشكال باللغة العربية (Traditional Arabic) بحجم (١٢)، وباللغة الإنجليزية (Times New Roman) بحجم (١٠)، وتكون العناوين الرئيسية في اللغتين بولد (Bold).
٦. يلتزم الباحث/ الباحثون في البحوث المكتوبة باللغة العربية باستخدام الأرقام العربية (١، ٢، ٣...) في جميع ثنايا البحث.
٧. يكون ترقيم صفحات البحث في منتصف أسفل الصفحة، ابتداءً من صفحة الملخص العربي ثم الملخص الإنجليزي وحتى آخر صفحة من صفحات البحث ومراجعته.
٨. توضع قائمة بالمراجع العربية بعد المتن مباشرة، مرتبة هجائياً حسب الاسم الأخير للمؤلف الأول، يليها مباشرة قائمة المراجع الأجنبية، وذلك وفقاً لأسلوب التوثيق المتبع في المجلة.

خامساً: حقوق المجلة وحقوق الباحث أو الباحثين

١. تقوم هيئة تحرير المجلة بالفحص الأولي للبحث، وتقدير أهليته للتحكيم، أو الاعتذار عن قبوله حتى تنطبق عليه شروط النشر، أو رفضه دون إبداء الأسباب.
٢. تنتقل حقوق طبع البحث ونشره إلى المجلة عند إشعار الباحث بقبول بحثه للنشر، ولا يجوز نشره في أي منفذٍ آخر ورقياً أم إلكترونياً، دون الحصول على إذن كتابي من رئيس هيئة التحرير.
٣. لا يحق للباحث/ الباحثين التقدم بطلب لسحب البحث بعد إبلاغه/ إبلاغهم بوصول البحث إلى المجلة.
٤. هيئة التحرير الحق في ترتيب البحوث المقدمة عند النشر لاعتبارات فنية.
٥. هيئة التحرير الحق في اختصار أو إعادة صياغة بعض الجمل والعبارات لأغراض الضبط اللغوي ومنهج التحرير.
٦. يبلغ الباحث بعدم قبول بحثه بناءً على تقارير المحكمين دون إبداء أسباب.
٧. ترسل نسخة إلكترونية للباحث/ الباحثين من العدد المنشور فيه بحثه/بحثهم، ونسخة إلكترونية أيضاً لمستلة البحث.

سادساً: إجراءات النشر في المجلة

١. إرسال البحث إلكترونياً بصيغة (word) وبصيغة (PDF) طبقاً للشروط والقواعد والتعليمات الخاصة بالمجلة والمذكورة أعلاه، ويرفق مع البحث سيرة ذاتية للباحث/ الباحثين؛ إن كانت مراسلته/ مراسلتهم المجلة هي الأولى لهم.
٢. إرسال البحث إلكترونياً من خلال موقع المجلة الإلكتروني [/https://journals.kku.edu.sa/ssjt/ar](https://journals.kku.edu.sa/ssjt/ar)
٣. أن يوقع الباحث/ الباحثون إقراراً يفيد أن البحث لم يسبق نشره، وأنه غير مقدم ولن يقدم للنشر في جهة أخرى حتى تنتهي إجراءات تحكيمه ونشره في المجلة، أو رفضه، وأنه غير مستل من أية دراسة أياً كان نوعها.
٤. إشعار الباحث عبر البريد الإلكتروني باستلام بحثه خلال خمسة أيام من تاريخ إرساله للمجلة.
٥. إشعار الباحث بإرسال البحث للتحكيم في حال اجتياز بحثه للفحص الأولي أو إعادته للباحث في حال رفضه.
٦. إرسال البحث المقدم للنشر - في حال اجتيازه للفحص الأولي - إلى محكمين من ذوي الاختصاص يتم اختيارها بسرية تامة، وذلك لبيان مدى أصالته وجدته وقيمة نتائجه وسلامة طريقة عرضه ومن ثم مدى صلاحيته للنشر.
٧. بعد التحكيم، ترسل تقارير المحكمين للباحث/ الباحثين لإجراء التعديلات التي أوصى بها المحكمون.
٨. بعد عمل التعديلات، يعاد إرسال النسخ الأصلية للبحث والنسخة المعدلة على البريد الإلكتروني للمجلة لمراجعة البحث في صورته النهائية من هيئة التحرير.
٩. إشعار الباحث بقبول بحثه للنشر إلكترونياً على موقع المجلة.

افتتاحية العدد

بسم الله الرحمن الرحيم

بحمد الله وتوفيقه، تطلّ مجلة الجمعية السعودية العلمية للمعلم على قرائها مع بداية المجلد الثالث، العدد الأول لعام ٢٠٢٦م، مواصلةً رسالتها العلمية في دعم البحث التربوي والتعليمي، وتعزيز دور المعلم والباحث في تطوير التعليم في المملكة العربية السعودية.

تجدد المجلة التزامها بتقديم أبحاث رصينة تسهم في تحقيق مستهدفات رؤية المملكة ٢٠٣٠، ولا سيما ما يتعلق بتطوير التعليم ورفع جودة مخرجاته، ودعم الابتكار، وتعزيز مكانة البحث العلمي في الحقل التربوي. وتواصل المجلة مسيرتها الرائدة في خدمة المجتمع العلمي من خلال نشر دراسات نوعية تعالج قضايا التعليم والتعلم، وتقديم حلولاً وممارسات مبنية على الأدلة.

يضم هذا العدد مجموعة من البحوث المتنوعة التي شارك فيها أعضاء هيئة التدريس في الجامعات، والمتخصصون ذوو الخبرة في التعليم العام، إضافة إلى إسهامات واعدة من طلاب وطالبات الدراسات العليا. وتعكس هذه المشاركات ثراء الحقل التربوي وتنوع اتجاهاته البحثية، كما تجسد الجهود المستمرة لتعزيز جودة التعليم وتطوير ممارساته.

وتدعو المجلة الباحثين والمتخصصين للانضمام إلى مجتمعها العلمي، سواء بالمشاركة في التحكيم أو بنشر أبحاثهم، بما يسهم في ترسيخ مكانة المجلة ورفع تصنيفها محلياً وعالمياً. كما تتطلع إلى استمرار التعاون البناء مع الباحثين والمؤسسات التعليمية لتحقيق التميز العلمي المستدام.

وفي الختام، تتقدم هيئة التحرير بخالص الشكر والتقدير لكل من أسهم في إعداد هذا العدد، سائلين الله التوفيق لمواصلة مسيرة العطاء العلمي، وأن نلتقي بكم في أعداد قادمة تحمل مزيداً من الإثراء والتميز. والله ولي التوفيق.

رئيس هيئة التحرير

أ. د. عبدالله بن علي معيض آل كاسي

جدول المحتويات

الصفحة	عنوان البحث
27-1	مبادئ التربية الاقتصادية المستنبطة من القرآن الكريم لمواجهة ارتفاع مستوى الإنفاق في الأسرة السعودية د. عليا بنت علي بن محمد العقيلي..... فاعلية برنامج تدريبي في تخفيف الضغوط النفسية لدى العاملين في مجال رعاية الأشخاص ذوي اضطراب طيف التوحد في مراكز الرعاية النهارية في منطقة عسير
53-28	أ. آسية علي خضران العُمري، د. أحمد حسن خضري..... أ نموذج مقترح لتدريس العلوم قائم على التكامل بين أنموذجي دورة التقصي الشائبة (CICIM) والتعلم ثنائي الموقف (DSLIM) وأثره على تنمية مهارات التفكير الجانبي وخفض التحيز المعرفي لدى طلاب الصف الثاني المتوسط
86-54	د. محمد علي طاهر صلوي، أ.د. ناصر بن عبدالله الشهراني..... أثر استخدام تقنية الواقع المعزز في تدريس العلوم على تنمية الاستيعاب المفاهيمي لدى طالبات المرحلة الابتدائية
115-87	د. هديل سعيد عبدالرحمن آلسرور..... تصور مقترح لتدريس العلوم باستخدام تقنية الواقع المعزز بالمرحلة الابتدائية.
141-116	أ. أميرة خالد عبدالله عسيري، أ.د. عاصم محمد إبراهيم، د. علي بن جبران محمد الحراملة..... فاعلية وحدة دراسية مقترحة في مادة العلوم قائمة على التعليم الأخضر في تنمية استيعاب المفاهيم البيئية ومهارات حل المشكلات وتعزيز الشغف البيئي لدى طلاب المرحلة المتوسطة
170-142	د. يحيى إدريس عبده صميلي.....

الصفحة	عنوان البحث
194-171	دور المناهج التعليمية في تعزيز قيم النزاهة ومكافحة الفساد لدى الطلاب أ. سعد فرحان محمد العتيبي، د. إبراهيم عبدالله محمد يحيى.....
216-195	دور المناخ التنظيمي في تعزيز جودة أداء المعلمين في مدارس الهيئة الملكية بمدينة ينبع الصناعية أ. علي محمد القرني.....
237-217	تصورات أعضاء هيئة التدريس لدور الكفايات الرقمية في تطوير برامج الإعداد التربوي بكلية التربية أ. طلال غازي معاضد المطيري، أ. سعد فرحان محمد العتيبي، أ. ناصر فهيد ناصر الغري.....
262 -238	The Application of International Speaking Assessment Criteria among Saudi Female English Language Teachers. د. دلال عبدالله فهد القيعاوي.....

The Application of International Speaking Assessment Criteria among Saudi Female English Language Teachers.

د. دلال عبدالله فهد القيعاوي

أستاذ مناهج وطرق تدريس اللغة الإنجليزية المشارك، جامعة جدة

dalqaiawi@uj.edu.sa

Abstract

This study aims to examine the degree to which Saudi English Language teachers utilize the International Speaking Assessment Criteria in their instructional practices. The research sample comprised Saudi female English language teachers in the public education system of Saudi Arabia. The research employed a mixed methods approach with an embedded design, integrating qualitative and quantitative methodologies. The methods employed included a questionnaire and a semi-structured interview. The research findings indicate that Saudi teachers of English as a foreign language are not implementing international standards for assessing students' speaking skills. They depend on the assessment guidelines established by the Ministry of Education, which are included at the beginning of the textbooks for each educational level (elementary, intermediate, and secondary). The assessment methods differ among teachers and teachers reported that the scores assigned for the speaking examination are insufficient. The research concluded that these criteria represent distinct, segmented skills that do not constitute a reliable framework and that it may result in students' low speaking proficiency because it lacks real-life applications and authentic evaluation methods. The results also show no significant differences between the means of teachers concerning education level.

Key Words: Speaking Skill – International Assessment Criteria – Speaking Skills Assessment

درجة تطبيق المعايير الدولية في تقييم مهارة التحدث لدى معلمات اللغة الإنجليزية في المملكة العربية السعودية

الملخص:

يهدف البحث الحالي إلى تحديد درجة تطبيق معلمات اللغة الإنجليزية بالمملكة العربية السعودية للمعايير الدولية في تقييم مهارة التحدث باللغة الإنجليزية لدى الطالبات، وقد تكونت عينة البحث من جميع معلمات اللغة الإنجليزية بمدارس التعليم العام في المملكة العربية السعودية. واتبع البحث المنهج المختلط ذو التصميم المضمن لتصميم رئيسي للأساليب المختلطة بالبحث، وتكونت أدوات البحث من استبيان تم توزيعه على معلمات اللغة الإنجليزية بجميع مراحل التعليم العام (الابتدائي والمتوسط والثانوي) وذلك لمعرفة درجة تطبيق المعلمات للمعايير الدولية في التقييم، بينما تم تطبيق المقابلة المنظمة بهدف الوصول إلى فهم كامل ومعرفة شاملة لأساليب التقييم المستخدمة من قبل المعلمات في تقويم مهارة التحدث. وأظهرت النتائج أن معلمات اللغة الإنجليزية بالمملكة لا يتبعن معايير دولية أو معيارية محددة وإنما يتبعن التعاميم التي تصلهن من الوزارة والمهارات المطلوبة والمحددة لهن في بداية الكتب الدراسية لكل مرحلة، والتي هي عبارة عن مهارات محددة ومنفصلة عن بعضها لا تشكل هيكل معتمد يمكن الاعتماد عليه في تقييم مهارة التحدث باللغة الإنجليزية. واستنتجت الباحثة أن هذا ربما كان أحد أبرز الأسباب في ضعف الطالبات في مهارة التحدث باللغة الإنجليزية حيث لا تتم مراعاة ترابط المهارات ولا التطبيقات الواقعية أو التقويم الحقيقي في هذا الأسلوب من التقويم. كما أظهرت النتائج عدم وجود فروق ذات دلالة إحصائية بين متوسطات استجابات معلمات اللغة الإنجليزية بالنسبة للمستوى التعليمي.

الكلمات المفتاحية: مهارة التحدث، معايير التقييم الدولية، تقييم مهارة التحدث

Introduction

Speaking is a crucial skill among the various competencies of any language. This skill involves converting cognitive messages and emotions into verbal expressions. It is the primary medium of communication that facilitates interaction among individuals. Speaking is a dynamic activity that combines language skills with social ability. Yang (2010) asserts that the main objective of oral English training is the enhancement of the ability to engage in natural conversation in professional or everyday contexts. Speaking encompasses more than mere word choice; it involves tone, pitch, and volume, all of which can profoundly influence the conveyed message. It entails many aspects which are: Pronunciation, Vocabulary, Grammar, and Fluency. Pronunciation has to do with the way words are pronounced which may affect the understanding of what was said. Vocabulary is the list of words used by a speaker and is the means of putting an idea into proper words. Grammar offers a framework of sorts for constructing sentences so that the message is coherent. Fluency, means the smoothness and the facility of communication which can improve the interaction between the listener and the speaker. When these components are combined they form the basis of good verbal communication or speaking.

According to Brown & Yule (1983) “One of the most difficult aspects of language learning, according to traditional views, is learning to speak — that is, to produce language orally in the target language.” To become a professional speaker and to learn each component of the speaking skills of a professional speaker, students must know the rules of the assessment of speaking skills. Holroyd (2000) suggest that assessment should be viewed as part of the learning process and not as an event that occurs at the end of the learning process (cited in Ribaya Emperador-Garnace et al., 2021, p. 42). Systematic assessment enhances students’ rehearsal of their speaking skills which in turn results in enhanced fluency and confidence. The assessment of speaking skills enables students to engage in authentic communication scenarios while developing their ability to convey ideas effectively. The assessment process develops students' ability to assess their learning, their identity, and their ability to learn languages. The assessment also enables teachers to identify weaknesses of the teaching strategy and to provide additional support to students for improvement.

The assessment of speaking skills prepares students for actual communication events and proper expression of their ideas and leads to the improvement of their skills as well

as their identity and learning power. Molla et al. (2017) report that assessment makes instructors contemplate challenging questions that challenge the effectiveness of their teaching methods and the possibility of applying new strategies to enhance teaching and improve learners' achievement. However, proceeding with speaking assessments can be time-consuming, particularly in large classrooms. It is a time-consuming process because it involves listening to every student, providing him/her with feedback, and documenting the outcomes. Teachers who are non-native speakers of the target language may face significant challenges in assessing speaking skills because they may not have the necessary knowledge about pronunciation, intonation, and cultural aspects that are required to give proper and fair marks to their students. Teachers must ensure that the criteria they set for assessment are valid and reliable. The criteria should be well-defined and systematically developed to achieve precise and rigorous results.

However, the development of speaking assessment criteria is not an easy process. Ulker (2017) declared that it is not an easy task to develop speaking assessment criteria because it is not only the identification of the possible criteria that can be used to measure students' speaking proficiency but also the limitation of the criteria to the most critical and relevant ones to give the required outcomes. Among the four language skills (speaking, listening, reading, and writing) speaking is reported to be the most difficult skill to assess (Richards, 2008; Navidinia et al., 2019). Fulcher (2018) pointed out that speaking assessment is a complex process because oral production is a creative product that needs a strong ground to align with continuous changes. Therefore, speaking assessment criteria should focus on particular aspects and should adopt well-structured criteria (Gulchiroy et al., 2021). These criteria should be highly valid and reliable and well developed to encompass all aspects of good speaking skills. Examples of well-defined criteria include the Test of English as a Foreign Language (TOEFL), the International English Language Testing System (IELTS), and the Common European Framework of Reference for Languages (CEFR). The researcher found out that they are found on a predetermined set of factors: accuracy, vocabulary range, delivery, pronunciation, responsiveness, and the use of several linguistic instruments that cover nearly every aspect of speaking skills assessment. "These criteria are used to delineate the knowledge and skills shown by students during the oral examination" (Ulker, 2017). Other studies suggest that standardized tests serve as excellent predictors of students'

retention, provide consistent and comparable data about students learning, and help monitor students progress throughout their educational journey (Benjamine and Pashler, 2015 and Jakirlic and Hoppe, 2021). Moreover, the results of these tests are used as a criterion to qualify Saudi students for scholarship.

The following are brief descriptions of these criteria

1. IELTS Speaking Band Descriptors which is an international speaking assessment criterion that encompass Fluency and coherence, Lexical resource, Grammatical range and accuracy, and Pronunciation (British Council, 2014).
2. The TOEFL Essentials speaking test is made up of 3 task types with a total of 19 recordings. For the first 6 recordings, the teste will read one side of a dialog, hear a recording, and then see the expected response on the screen. S/he must read that response aloud. The second task type is listen and repeat. S/he will hear 8 sentences aloud and must repeat them. The final speaking task is a simulated interview, where s/he will watch a video with 5 questions about oneself and be given time to respond to each one (Ulker, 2017).
3. In CEFR Criteria students' speaking skills are assessed using scales that are linked to the CEFR scale. The assessor gives 0–5 marks for each of the following criteria: Grammatical Resource, Lexical Resource, Discourse Management, Pronunciation, and Interactive Communication. Marks for each of these criteria are doubled (Roca-Varela, 2013).

The Problem

The quality of assessment methodologies, instruments, or rubrics greatly affect the assessment process and the evaluation of speaking skills, thus affecting the quality of the overall speaking skill outcomes. Molla et al., (2017) emphasize that assessment determines students' performance and teachers' actual classroom practices. Teachers must utilize an authorized assessment tool to effectively guide students in developing specific skills and achieving desired outcomes. According to Duque-Aguilar (2021), students' assessments should be equated with the teaching of the sub-skills of speaking. Appropriate assessment tools and their application frequency determine the success of language teaching and learning. The assessment of speaking should be closely connected with teachers' instructions to enable them to make decisions based on student's linguistic skills and course goals and to select the appropriate speaking tasks (Fulcher, 2018; Ginther, 2012; Shaaban, 2005). These methods will illustrate to students that assessment

is fundamental to the learning process and represents a crucial element of their education (Jones, 2011).

Research shows that despite the increasing record supporting English instruction in Saudi Arabia, second language acquisition remains considerably below expectations (Akhter, 2020; Alqahtani, 2016; Alrasheedi, 2020; Younes, 2016). Saudi students' skills in speaking and other related skills fall short of the anticipated standards (Abker, 2020; Alshammari, 2020; Alshammari, 2021). In Saudi Arabia, there is a widely held perception that EFL students can attain high levels of proficiency in writing reading, and listening, but they cannot speak English as compared to their peers from around the globe (Kashmiri, 2020). Extensive study has been undertaken on the pedagogical method of developing speaking skills among Saudi students (Alsiyat, 2021; Faiyaz and Haq, n.d.; Mohammed Ali et al., 2019) while less research has been undertaken on how teachers should assess their students during spoken English instruction “Instead of assessing four basic skills, only writing skills are assessed in the public sector of educational institutions” (Faiyaz and Haq, n.d., p. 200). Furthermore, in an exploratory study adopting a focused interview, the researcher found that the criteria adopted in assessing students' performance are not clearly defined and that there are no specific criteria used for the assessment of speaking skills among Saudi EFL teachers. Therefore, the researcher, supported by the results of other research, like (Alharbi and Surur, 2019; Rashid, Khan, and Sheikh, 2022, Alshaki, 2024; Alghanmi, 2023) believes that Saudi EFL teacher's perception of the appropriateness and effectiveness of the assessment methods plays a vital role in determining degree and frequency of assessment practices in English language speaking skills classes. The present research intends to fill this research gap and to achieve the following objectives.

Research Objectives

1. Identify the method authorized by international criteria for assessing the speaking skills of Saudi students in EFL classrooms.
2. Determine the degree of Saudi English Language teachers' application of international speaking assessment criteria among students.

Research Questions

The present research intends to answer the following questions:

1. What are the methods authorized by international criteria for assessing the speaking skills of students in EFL classrooms?

2. What is the degree of Saudi English Language teachers' application of international speaking assessment criteria among students in EFL classrooms?

Research Methodology

The present study adopted a mixed methods approach, involving the use of both qualitative and quantitative methodologies to get a comprehensive knowledge of the assessment methods employed by Saudi EFL teachers.

Table (1): Qualitative and Quantitative Methods Adopted in Research

No	Method	Type	Objective
1	The Descriptive Method	Quantitative	- Analysis of International Criterion - Description of Saudi EFL teacher's responses to the Questionnaire
2	The phenomenological approach	Qualitative	- Achieving a better understanding of the assessment methods adopted by Saudi EFL teachers

The embedded design constitutes a primary approach within mixed methods of research. The process involves the concurrent gathering of both quantitative and qualitative data, with the qualitative data integrated within the quantitative framework. This design is optimal for research emphasizing quantitative data while also seeking to comprehend how qualitative data provides additional insights. This design enables the researcher to investigate research questions from various perspectives.

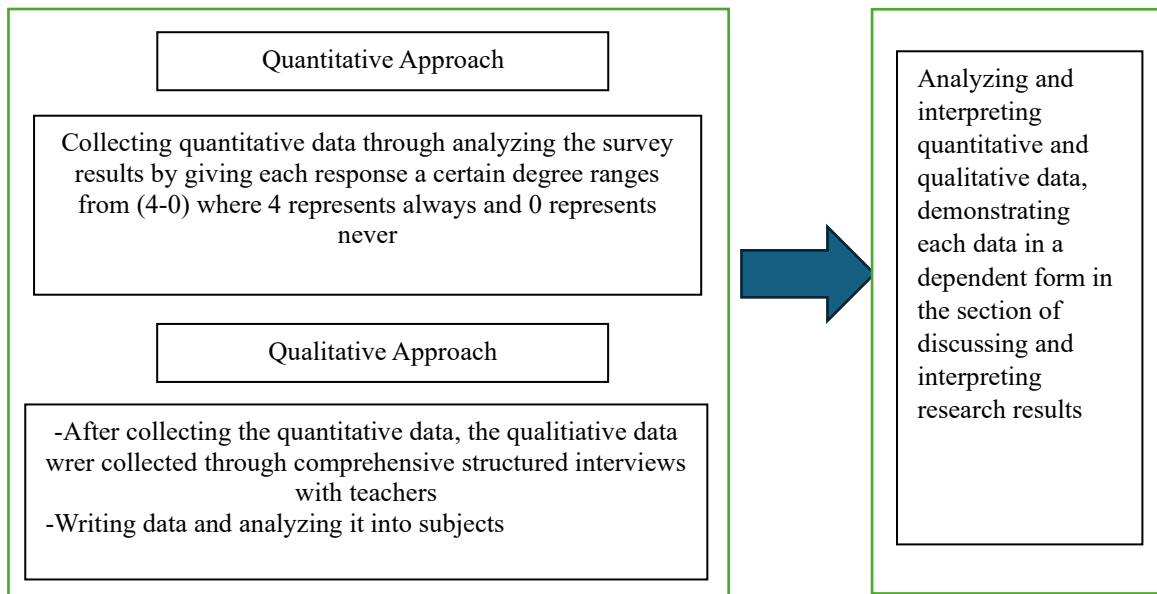


Figure (1): Design of the Mixed Method Approach

Community and sample

The research community encompasses all Saudi EFL teachers in all three levels of Saudi Public Education (i.e., Elementary, Intermediate, and Secondary) covering all five regions of Saudi Arabia. The sample was the convenience sample including the teachers who responded to the questionnaire and they reached to 112 respondents. While the focused group of the interview reached to 20 teachers.

Research Tools

1. The Quantitative tool (Teachers' Questionnaire).
2. The Qualitative tool (Teachers' Interview).

1. The Quantitative tool (Teachers' Questionnaire)

- a- **Determining objective of the questionnaire:** The objective of the questionnaire is determined by assuring that it will help to give answers to the research questions and achieve its objective that is to determine the degree of Saudi English Language teachers' application of international speaking assessment criteria among students.
- b- **Defining dimensions of the questionnaire:** The dimensions of the questionnaire are derived from the international criterion for assessing English language skills, as outlined below: IELTS, TOEFL, and CEFR encompass the primary dimensions of speaking skills assessment, specifically Fluency and Coherence, Lexical Resource, Grammatical Range and Accuracy, Pronunciation, Sociolinguistic skills, and Non-verbal techniques. The primary dimensions of the questionnaire are established through an examination of the selected international criteria, classified into nine key dimensions: Grammar range, Accuracy, Vocabulary range, Fluency, Coherence, Pronunciation, Conversational Skills, Sociolinguistic Skills, and Non-Verbal Expressions.
- c- **Constructing the Questionnaire:** After establishing the nine key dimensions, the researcher determined the indicators of the nine dimensions which encompass 27 indicators distributed as follows:

Table (2): Number of the Questionnaire Dimensions and Indicators

Dimension	Indicators
Grammar range	4
Accuracy	3
Vocabulary Range	4
Fluency	2
Coherence	4
Pronunciation	4
Conversational Skills	4

Dimension	Indicators
Sociolinguistic Skills	2
Non-Verbal Expressions	2

d- **Estimating grades for teachers' Responses:** The questionnaire grades are distributed on a five Likert points scale (always, very frequently, occasionally, rarely, and never) ranging from 4 to 0 for positive indicators and from 0 to 4 for negative indicators.

e- **Validity of the Questionnaire:**

The validity of the study tool was verified by performing two procedures:

1. Content validity

After completion of the initial questionnaire, it was sent to a group of 11 referees from the faculty of Saudi universities specializing in English Language Curriculum and Instruction. This was done to ensure the linguistic reliability, clarity of meanings, and relevance of each indicator to its assigned dimension, along with adding appropriate amendments and suggestions for improvement of the questionnaire. The judgment's findings suggested revisions to several questionnaire questions for clarity for Saudi EFL teachers in public education. Nonetheless, the majority of the questionnaire questions surpassed the estimated percentage of the referee's consensus, which is 90%.

2. Internal validity

Following the validation of the questionnaire's content, it was administered to additional Saudi EFL teachers from outside the chosen study sample. Upon data collection, the Pearson Correlation coefficient was computed to assess the internal validity of the questionnaire. The correlation coefficient was computed between the score of every indicator on the questionnaire and the total score of the corresponding dimension. Table 3 presents the validity findings of the questionnaire for each dimension.

Table (3): Pearson Correlation Coefficient between each Indicator and the related Dimensions

Dimension	Indicator	Pearson Correlation	Indicator	Pearson Correlation
Grammar rang	1	664.0**	2	711.0**
	3	632.0**	4	642.0**
Accuracy	1	512.0*	2	630.0**
	3	674.0**	2	
Vocabulary range	1	621.0**	4	509.0*
	3	604.0**	2	508.0*
Fluency	1	527.0*	2	512.0*
Coherence	1	608.0**	4	614.0**
	3	520.0*	2	503.0*
Pronunciation	1	501.0*	4	616.0**
	3	631.0**	2	522.0*
Conversational Skills	1	674.0**	4	703.0**
	3	681.0**	2	693.0**
Sociolinguistic Skills	1	504.0*	2	512.0*

Dimension	Indicator	Pearson Correlation	Indicator	Pearson Correlation
Non-Verbal Expressions	1	515.0*	2	508.0*
		*Significant at .05	**Significant at .01	

Table 3 indicates that the correlation coefficients between the indicators and the total score for each dimension of the questionnaire are both good and acceptable for scientific research, as all are significant at a level below 0.05. Table 4 presents the results of construct validity, illustrating the relationship between the scores of each dimension and the total score of the questionnaire.

Table (4): Values of construct validity between the score of each dimension and the total score of the questionnaire.

Dimension	Pearson Correlation
Grammar rang	*0.528
Accuracy	**0.655
Vocabulary range	*0.531
Fluency	*0.549
Coherence	*0.530
Pronunciation	**0.04
Conversational Skills	**0.598
Sociolinguistic Skills	*0.538
Non-Verbal Expressions	*0.496

Table 4 indicates that the correlation coefficients between the dimension scores and the total score of the questionnaire are robust for scientific research, as all coefficients achieved statistical significance at a level below 0.05, thereby demonstrating the validity of the questionnaire's internal structure.

3. Reliability of the Questionnaire:

The questionnaire's reliability was verified using Cronbach's Alpha formula for each dimension individually and for the total indicators, after its application to the exploration sample of the research. Table number 5 presents the reliability coefficients of the questionnaire across all dimensions.

Table (5): Values of reliability coefficients according to Cronbach's Alpha equation for the various axes and dimensions of the questionnaire.

Questionnaire dimensions	Alpha Cronbach
Grammar range	85.0
Accuracy	82.0
Vocabulary range	84.0
Fluency	78.0
Coherence	79.0
Pronunciation	85.0
Conversational Skills	85.0
Sociolinguistic Skills	80.0
Non-Verbal Expressions	81.0
Total score of the questionnaire	0.86

The results of Table 5 indicate that Cronbach's Alpha coefficients are appropriate for scientific research purposes, as the reliability coefficients on the questionnaire ranged between (0.78) and (0.85), while the reliability coefficients on the total score of the questionnaire reached (0.86).

4. **Time for answering the Questionnaire:** The Mean of time required for answering the questionnaire was calculated and reached 15 minutes.

5. **The final form of the Questionnaire:** The final form of the questionnaire comprised 29 indicators divided into 9 dimensions. The maximum grade of the questionnaire reached 116 and the minimum grade reached to (0).

6. **Questionnaire Scale:** Each indicator has five levels, with a score of (0) for “Never”, a score of (1) for “Rarely”, a score of (2) for “Occasionally”, a score of (3) for “Very frequently”, and a score of (4) for “Always”. The responses were monitored according to a five-point Likert scale to determine the degree of agreement with the indicators of each dimension of the questionnaire, as shown in Table (6).

Table (6): Degree of agreement according to the five-point Likert scale

No	Degree of approval	Means
1	Never	1: <1.8
2	Rarely	1.8: < 2.60
3	Occasionally	2.6: < 3.4
4	Very frequently	3.4: <4.2
5	Always	4.21: <5

2. The Qualitative tool (Teachers' Interview)

The qualitative part of the research included the researcher conducting in-depth interviews with Saudi EFL teachers practicing the assessment of speaking skills among students. The embedded design employs a qualitative method to support the findings of the quantitative instrument. In-depth interviews are an interactive qualitative inquiry approach that relies on verbal participation, so they're appropriate for exploring participants' perspectives and experiences. Some open-ended questions were prepared related to teachers' perspectives and practices during the speaking skills assessment. The questions are as follows:

1. How often do you assess your students' speaking skills?
2. How do you assess your student's speaking skills?
3. Do you consider certain criteria when assessing your students?

4. What type of assessment do you use to evaluate a student's speaking skills (in-class exercises or a specific test)?

Results

The answer to the first question of the research focuses on the main dimensions and the specific indicators of assessments that are deduced from international criteria.

Q.1. What are the international assessment criteria that should be adopted by Saudi EFL teachers to assess the speaking skills of their students in EFL classes?

The analysis focused on the predominant international standards recognized in Saudi Arabia. The three assessed criteria included IELTS, TOEFL, and CEFR. The dimensions of the questionnaire are based on these international criteria within the section assessing English Speaking Skills. The investigation uncovered several dimensions: Fluency and Coherence, Lexical Resource, Grammatical Range and Accuracy, Pronunciation, Sociolinguistic Skills, and Non-verbal Techniques. The questionnaire dimensions are subsequently developed by presenting them to a panel of referees specialized in English language Curriculum and Instruction for assessment of content validity. The review produced the following dimensions: Grammar Range, Accuracy, Vocabulary Range, Fluency, Coherence, Pronunciation, Conversational Skills, Sociolinguistic Skills, and Non-Verbal Expressions, which are further detailed by 32 indicators as presented in Table2.

Q.2. What is the degree of Saudi EFL teachers' inclusion of international Speaking skills assessment criteria into their teaching practices?

To answer the second question, two steps are followed:

1. A questionnaire was distributed to 112 Saudi EFL teachers in public education schools across all five regions of Saudi Arabia. The frequencies, Means, Standard Deviations, and Rankings of the indicators for each dimension were computed independently. Table 7 presents a comprehensive examination of the grades across each dimension.

Table (7): Results of the Grammar Range Dimension

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
		Always	Very frequently	Occasionally	Rarely	Never				
I evaluate the appropriate use of subordinate clauses within clauses and indicator	F	0	25	20	10	0	27.3	75.0	1	Occasionally
	%	0	5.45	4.36	2.18	0				
I determine the complexity of the verb phrase used	F	0	24	22	9	0	26.3	73.0	2	Occasionally
	%	0	6.43	0.40	4.61	0				
I assess the complexity of other indicators	F	0	21	22	12	0	16.3	76.0	3	Occasionally
	%	0	2.38	0.40	8.21	0				
I estimate the length of the spoken sentences	F	0	21	22	11	1	14.3	80.0	4	Occasionally
	%	0	2.38	0.40	0.20	8.1				
Mean	Degree of Agreement						St. Deviation			
3.20	Occasionally						0.77			

The results of Table 7 indicate that the general Mean of the indicators of the Grammar Range dimension, which includes 4 indicators, amounted to (3.20), with a standard deviation of (0.77), which means that the participants agreed with an (Occasionally) degree to all the indicators of this dimension.

Table (8): Results of the Accuracy Dimension

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
		Always	Very frequently	Occasionally	Rarely	Never				
I define the range of sentence structures, especially how to move elements around for information focus	F	0	27	18	9	1	20.3	80.0	1	Occasionally
	%	0	1.49	7.32	4.16	8.1				
I evaluate the communicative effect of errors like its effect on intelligibility and precision or expression.	F	1	23	20	10	1	14.3	91.0	2	Occasionally
	%	8.1	8.41	4.36	2.18	8.1				
I identify the error density of sentences like the number of grammatical errors in a given speech	F	1	23	19	7	5	05.3	04.1	3	Occasionally
	%	8.1	8.41	5.34	7.12	1.9				
Mean	Degree of Agreement						St. Deviation			
3.13	Occasionally						0.96			

The results of Table 8 indicate that the general Mean of the indicators of the Accuracy dimension, which includes 3 indicators, amounted to (3.13), with a standard deviation of

(0.96), which means that the participants agreed with an (Occasionally) degree to all the indicators of this dimension.

Table (9): Results of the vocabulary range Dimension

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
		Always	Very frequently	Occasionally	Rarely	Never				
I value the variety of words used	F	0	40	11	4	0	3.65	0.61	1	Very frequently
	%	0	7.72	0.20	3.7	0				
I assess the adequacy and appropriateness of vocabulary concerning the requirements of the speaker's attitude	F	0	31	18	6	0	3.45	0.68	2	Very frequently
	%	0	4.56	7.32	9.10	0				
I assess the adequacy and appropriateness of vocabulary concerning the requirements of style	F	0	29	18	8	0	3.38	0.78	3	Occasionally
	%	0	7.52	7.32	5.14	0				
I assess the adequacy and the appropriateness of vocabulary concerning the requirements of referential meaning	F	0	26	20	9	0	3.30	0.74	4	Occasionally
	%	0	3.47	4.36	4.16	0				
Mean	Degree of Agreement						St. Deviation			
3.44	Very frequently						0.73			

The results of Table 9 indicate that the general Mean of the Vocabulary Range dimension indicators, which includes 4 indicators, reached (3.44), with a standard deviation of (0.73), which means that the participants agreed with (Very frequently) with the Vocabulary Range dimension. The previous table shows that two indicators obtained an agreement degree (Occasionally), while two indicators obtained an agreement degree (Very frequently).

Table (10): Results of the Fluency Dimension

Indicator	Degree					Mean	St. Deviation	M	Degree of Approval	
		Always	Very frequently	Occasionally	Rarely					Never
I assess speech rate so as not too slow to keep	F	0	29	21	5	0	3.44	0.66	1	Very frequently
	%	0	7.52	2.38	1.9	0				

Indicator	Degree						Mean	St. Deviation	M	Degree of Approval
	Always	Very frequently	Occasionally	Rarely	Never					
links between words or propositions in mind.										
I evaluate speech continuity so the flow of speech will not be interrupted by pausing during the search for appropriate words	F	0	28	21	6	0	3.40	0.68	2	Occasionally
	%	0	9.50	2.38	9.10	0				
Mean	Degree of Agreement						St. Deviation			
3.42	Very frequently						0.68			

The results of Table 10 indicate that the general Mean of the two indicators of the Fluency dimension was (3.42), with a standard deviation of (0.68), which means that the participants agreed with (Very frequently) to the Fluency dimension. The previous table shows that one indicator obtained an agreement (Occasionally) degree and the other indicator obtained an agreement degree (Very frequently).

Table (11): Results of the Coherence Dimension

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
	Always	Very frequently	Occasionally	Rarely	Never					
I consider the logical sequencing of spoken sentences	F	0	37	13	5	0	3.58	0.66	1	Very frequently
	%	0	3.67	6.23	1.9	0				
I determine the use of cohesive devices within and between spoken sentences.	F	0	32	17	5	1	3.45	0.61	2	Very frequently
	%	0	2.58	9.30	1.9	8.1				
I assess the relevance of spoken sentences to the general purpose of a turn	F	0	30	15	8	2	3.32	0.56	3	Occasionally
	%	0	5.54	3.27	5.14	6.3				
I identify the ability to divide speech into meaningful utterances or chunks within spoken sentences	F	0	26	18	10	1	3.25	0.62	4	Occasionally
	%	0	3.47	7.32	2.18	8.1				
Mean	Degree of Agreement						St. Deviation			
3.40	Occasionally						0.62			

The results of Table 11 indicate that the general Mean of the Coherence dimension indicators, which number 4 indicators, reached (3.40), with a standard deviation of (0.62), which means that the participants agreed with (Occasionally) with the Coherence dimension. The previous table shows that two indicators obtained an agreement degree (Occasionally), and two indicators obtained an agreement degree (Very frequently).

Table (12): Results of the Pronunciation Dimension

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
		always	Very frequently	Occasionally	Rarely	Never				
I assess the overall effect of accent on intelligibility	F	0	39	8	8	0	3.56	0.73	1	Very frequently
	%	0	9.70	5.14	5.14	0				
I evaluate the appropriate use of rhythm and stress timing, and the linking of sounds to produce connected speech	F	0	37	11	6	1	3.52	0.76	2	Very frequently
	%	0	67.3	20.0	10.9	8.1				
I value the production of sounds at the word and phoneme levels	F	0	35	12	7	1	3.47	0.97	3	Very frequently
	%	0	6.63	8.21	7.12	8.1				
I track the development of topics during speech	F	0	29	18	8	0	3.38	0.93	4	Occasionally
	%	0	7.52	7.23	5.14	0				
Mean	Degree of Agreement						St. Deviation			
3.48	Very frequently						0.91			

The results of Table 12 indicate that the general Mean of the Pronunciation dimension indicators, which numbered 4 indicators, reached (3.48), with a standard deviation of (0.91), which means that the participants agreed (Very frequently) with the Pronunciation dimension. The previous table shows that 3 indicators obtained an agreement degree (Very frequently), while one phrase obtained an agreement degree (Occasionally).

Table (13): Results of the Conversational Skills Dimension

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
		Always	Very frequently	Occasionally	Rarely	Never				
I value student's initiative in turn-taking and topic control	F						3.61	0.62	1	Very frequently
	%	0	38	13	4	0				
	F	0	1.69	6.23	3.7	0	3.58	0.71	2	

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
		Always	Very frequently	Occasionally	Rarely	Never				
I track cohesion with	%	0	38	12	4	1	3.50	0.63	3	Very frequently
I assess students to distinguish between register and style	F	0	1.69	8.21	3.7	1.8				
I evaluate conversation maintenance including clarification, repair, checking, and pause fillers	F	0	2.58	5.34	3.7	0	3.41	0.80	4	Very frequently
	%	0	33	13	8	1				
Mean	Degree of Agreement						St. Deviation			
3.52	Very frequently						0.70			

The results of Table 13 indicate that the general Mean of the indicators of the Conversational Skills dimension, which number 4 indicators, reached (3.52), with a standard deviation of (0.70), which means that the participants agreed with (Very frequently) for the Conversational Skills dimension, and it is clear from the previous table that all the indicators of this dimension obtained an agreement degree (Very frequently).

Table (14): Results of the Sociolinguistic Skills Dimension

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
		always	Very frequently	Occasionally	Rarely	Never				
I assess eye contact and facial expressions	F	0	36	16	4	0	3.56	0.63	1	Very frequently
	%	0	6.36	1.29	3.7	0				
I value the use of cultural references	F	1	30	16	8	1	3.36	0.80	2	Occasionally
	%	1.8	5.45	1.29	5.14	1.8				
Mean	Degree of Agreement						St. Deviation			
3.46	Very frequently						0.74			

The results of Table 14 indicate that the general Mean of the indicators of the Sociolinguistic Skills dimension, which number 3 indicators, amounted to (3.46), with a standard deviation of (0.74), which means that the participants agreed with (Very frequently) in the Sociolinguistic Skills dimension. The previous table shows that one indicator obtained an agreement degree (Occasionally), while another indicator obtained an agreement degree (Very frequently).

Table (15): Results of the Non-Verbal Expressions Dimension

Indicator	Application Degree						Mean	St. Deviation	M	Degree of Approval
	Always	Very frequently	Occasionally	Rarely	Never					
I consider the effect of body posture and relevant gestures	F	0	40	11	4	0	3.64	0.61	1	Very frequently
	%	0	7.72	0.20	3.7	0				
I assess eye contact and facial expressions.	F	1	36	16	3	0	3.60	0.54	2	Occasionally
	%	1.8	5.65	1.29	5.5	0				
Mean	Degree of Agreement						St. Deviation			
3.54	Very Frequently						0.80			

The results of Table 15 indicate that the general Mean of the indicators of the Non-Verbal Expressions dimension, which numbered 2 indicators, reached (3.54), with a standard deviation of (0.80), which means that the participants agreed with (Very frequently) for one of the Non-Verbal Expressions dimension, and occasionally for the other. However, it is clear from the previous table that the total result of this dimension obtained an agreement degree (Very frequently).

The following table presents a comprehensive overview of the aggregate findings across all dimensions of the questionnaire.

Table (16): Order of All the questionnaire Dimensions

Dimension	Means	Std. Deviations	M	Degree of Approval
Grammar range	3.20	0.77	8	Occasionally
Accuracy	3.13	0.96	9	Occasionally
Vocabulary range	3.44	0.73	5	Very frequently
Fluency	3.42	0.68	6	Very frequently
Coherence	3.40	0.62	7	Occasionally
Pronunciation	3.48	0.91	3	Very frequently
Conversational Skills	3.52	0.70	2	Very frequently
Sociolinguistic Skills	3.46	0.74	4	Very frequently
Non-Verbal Expressions	3.54	0.58	1	Very frequently
Total score for all dimensions	3.39	0,76	...	Occasionally

The results related to Table 16 indicate that the Mean for all dimensions of the questionnaire was (3.39), with a standard deviation of (0.76). This Mean falls at the third level of approval degrees which is (Occasionally). Non-verbal expressions came in first place with a Mean of (3.54), and an approval degree (Very frequently). Conversational Skills came in second place with a Mean of (3.52), and an approval degree (Very frequently). Pronunciation came in third place with a Mean of (3.48), and an approval degree (Very frequently). Sociolinguistic Skills came in fourth place with a Mean of (3.46), and an approval degree (Very frequently). Vocabulary came in fifth place with a

Mean of (3.44), and an approval degree (Very frequently). Fluency came in sixth place with a Mean of (3.42), and an approval degree (Very frequently). Coherence came in seventh place with a Mean of (3.40), and an approval degree (Occasionally), and Grammar Range came in eighth place with a Mean score of (3.20), and an approval degree (Occasionally), and Accuracy came in ninth place with a Mean of (3.13) and an approval degree (Occasionally).

2. To confirm the findings obtained from the questionnaire, the researcher conducted a focused group interview. A group of 21 EFL teachers were asked to respond to four open-ended questions. The data obtained from the interview was classified and encoded into three primary categories as follows:

Category one: Approach of assessment

This category includes answers provided by Saudi EFL teachers in public education about the type and frequency of assessment methodologies used to assess their students' speaking skills. This category includes the first and second questions, which are as follows:

The first question is: "How often do you assess your students' speaking skills?" According to the answers given by the teachers, the researcher perceived that teachers are using both formative and summative assessments in assessing the speaking skills of students in EFL classrooms and they are using it twice during the term. Using both types of assessments in a foreign language classroom is a comprehensive method to help students learn and test their knowledge. These two kinds of tests have different but complimentary objectives, and using them together improves both teaching and student performance.

The second question is: "What type of assessment do you use to evaluate student's speaking skills?" The answers of the teachers suggest that they use both methods of assessment. They use the formative evaluation through in-class exercises and the summative evaluation. However, teachers declared that this depends on the available time, the marks specified for speaking skills assessment, and the number of students in class.

Category two: The adopted Criteria

This category contains information deduced from the answers to **the third question which is:** "Do you consider certain criteria when assessing your students' speaking skills?"

Most of the participants reported that they are applying the instruction given by the Ministry of Education. These cover certain elements and are offered at the beginning of each textbook in every level of public education. The researcher returned to the textbook of each level and analyzed the dimensions and related indicators to be assessed. Furthermore, the researcher examined their relation to international criteria that are IELTS, TOEFL, and CEFR. The results of the analysis demonstrated that the instruction covers only pronunciation, vocabulary, and communication but there is no reference to any international criteria. There are certain suggestions and examples in the textbook that teachers should follow and adopt but they do not cover most of the international criteria indicators. According to teachers' responses and answers this can be a reason for students' low level of performance regarding EL speaking skills.

Category three: The targeted language aspects

This category covers the answers to the fourth question and focuses on the language aspects that are targeted by EFL teachers during the assessment practices of speaking skills. **The fourth question is** "What language aspects do you consider when assessing your students?". Teachers enumerated only some of the indicators covered in the questionnaire. They mentioned that the language aspects they usually target are vocabulary, coherence, communication, fluency, and grammar. However, not all the teachers mentioned the same aspects, and the percentage differed from one teacher to the other. Table 17 illustrates the frequency and percentage of language aspects targeted by Saudi EFL teachers according to the participants' responses in the interview.=

Table (17): Frequency and Percentage of Language Aspects covered by Teachers

Language Aspect	Frequency	Percentage
Vocabulary	15	71.4
Fluency	13	61.9
Communication	11	52.4
Grammar	9	42.9
Pronunciation	7	33.3
Coherence	5	23.8

The table above elucidates the percentage of language aspects addressed by Saudi EFL teachers and delineates the specific language features involved. The results indicate that the absence of a standardized criterion for evaluating speaking skills leads to different responses among teachers in the aspects they emphasize, as they rely on personal experience and tailor their assessment methods based on introductory instructions and elements in the textbooks rather than a specific criterion. Teachers also include the marks

established for the speaking skills, and the allotted time and number of students in class. The findings indicate that the biggest proportion pertains to the assessment of Vocabulary due to its prominence at the beginning of the textbook and because it is recommended by the Ministry of Education.

Discussion

Analyzing the results of the questionnaire and the answers to the questions of the interview, the researcher concluded that EFL teachers in Saudi Public Schools are not adopting an international certified criterion in the assessment of the speaking skills of their students. However, they have certain instructions and guidelines from the Ministry of Education, and they are also provided with specific elements and recommendations at the beginning of each textbook which they follow in the assessment practices of speaking skills. Some of these elements are identical to the main dimensions in the international criteria but still, they are not enough for establishing consensus among EFL teachers and achieving the goal of proficient speaking. These results align with the studies of Shaaban (2005), Tracing (2011), and Duque-Aguilar (2021) who reported that teachers still have problems assessing communicative functions, forms of assessment, and language elements to assess. The challenges encountered by EFL teachers when assessing communicative functions mostly stem from their reliance on textbook exercises and the absence of realistic assessment methods. This conclusion also aligns with that of (Zaim et al., 2020) who pointed out that teachers often lack the knowledge to conduct accurate assessments and that this is due to the absence of a simple yet efficient tool for assessment.

Moreover, results indicated that the assessment methods differ from one teacher to the other according to students' proficiency level, time and marks specified for speaking skills, and number of students in the classroom. Pronunciation, vocabulary, and short conversations are the elements required in the textbooks and recommended by the Ministry of Education through administrative circulars. These differences among teachers create a weakness in speaking skills assessment and consequently, students' speaking proficiency as students will not be able to figure out what is expected from them or what learning outcomes are required. Teachers do not have enough knowledge to assess language skills authentically, so they need to follow reliable and trusted tools, and this is verified in international criteria like IELTS, TOEFL, and CEFR as described by Ülker (2017). This also aligns with Qizi and Gayratovna (2021) who recommended that teachers should emphasize specific criteria in assessing their students.

Adhering to a defined criterion is anticipated to ensure fairness, objectivity, transparency, accountability, consistency in grading, professionalism, and credibility, as well as improved identification of learning gaps. However, Faiyaz and Haq (no date) noted that speaking skills assessment receive insufficient attention in the Saudi schools. Furthermore, teachers highlighted in the interview that speaking skills receive insufficient emphasis, accounting for only 5 out of 40 marks in the final exam. They noted that they typically adhere to book criteria that emphasize specific elements rather than certified or international criteria.

In conclusion, using international criteria for evaluating EFL speaking skills is an urgent need as it standardizes assessments, mitigates discrepancies between student and teachers' assessments, and enhances learning outcomes which improves the overall effectiveness of language acquisition.

Recommendations

According to the results of the research and the analysis of teachers' responses, the researcher recommends that Saudi EFL teachers must apply certain assessment strategies to be able to achieve the targeted outcomes of speaking proficiency in EFL classes and the following are some of these recommendations:

1. Employing explicit, transparent and distinct criteria that include all components mandated by international assessment criteria. These elements are Pronunciation, Grammar, Vocabulary, Fluency, Coherence, and communication skills.
2. Applying both formative and summative assessment techniques while observing students across various contexts (e.g., collaborative tasks, presentations, informal discussions).
3. Employing diverse assessment techniques, including structured assessments (e.g., role-plays, presentations), informal assessments (e.g., class discussions, pair work), as well as self-assessment and peer evaluations to promote reflection and involvement.
4. Emphasizing both fluency and accuracy by allowing students to communicate freely with minimal correction to enhance fluency while offering targeted feedback on grammar, pronunciation, and vocabulary.
5. Providing constructive comments through integrating balanced positive reinforcement with suggestions for improvement.

6. Promoting self-monitoring and reflection via the use of recordings of students' speaking to facilitate their assessment and assist in establishing personal objectives for speaking improvement.
7. Tailoring tasks to students' competency levels and using scaffolding strategies to assist struggling students and novices.
8. Conducting research that addresses other skills of English as a foreign language like listening, reading and writing.
9. Research can be constructed to describe the degree of speaking skills assessment across different education levels.

References

- Abker, I. (2020). Difficulties in Pronouncing English Morphemes among Saudi EFL Students at Albaha University. A Case Study in Almandag. *Arab World English Journal*, 11(2), 395–410.
- Kashmiri, H. (2020). Communication challenges: Saudi EFL Speaking Skills and strategies to overcome speaking difficulties. *Arab World English Journal*, 267, 1–61. <https://doi.org/10.24093/awej/th.267>
- Ahmad, M. (2019). Specifying Criteria for the Assessment of Speaking Skill: A Library Based Review. *INTERNATIONAL JOURNAL OF EDUCATIONAL SCIENCES*, 24(1–3). <https://doi.org/10.31901/24566322.2019/24.1-3.1075>
- Akhter, T. (2020). Problems and Challenges Faced by EFL Students of Saudi Arabia during COVID-19 Pandemic. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 12(5).
- Alghanmi, B. (2023). *Self-assessment in EFL speaking classroom and its effect on achievement, self-regulated learning, and critical thinking: students' voices from Saudi Arabia*. Unpublished PhD thesis, University of Glasgow.
- Alharbi, A. and Surur, R. (2019). The Effectiveness of Oral Assessment Techniques Used in EFL Classrooms in Saudi Arabia From Students and Teachers Point of View. *English Language Teaching*; Vol. 12, No. 5. Published by Canadian Center of Science and Education
- Alqahtani, S. S. (2016). Enhancing the Saudi EFL Students' Pronunciation of the English Phoneme /v/ via Immersion in Virtual Platforms. *Arab World English Journal*, 7(3), 463–478.
- Alshakhi, A. (2024). Speaking Skill Assessment Instrument Validity: An Investigation into Instructors' Perceptions. *Journal of Echohumanism*. Volume 3 (7).
- Alrasheedi, S. (2020). Investigation of Factors Influencing Speaking Performance of Saudi EFL Learners. *Arab World English Journal*, 11(4), 66–77.

- Alshammari, H. (2020). Chinese Language in Saudi Arabia: Challenges and Recommendations. *English Language Teaching*, 13(2), 75.
- Alshammari, H. (2021). Assessing the Reading Skills of the Saudi Elementary Stage EFL Learners. *Advances in Language and Literary Studies*, 12(1), 55.
- Alsiyat, F. (2021). Speaking Challenges of Saudi EFL Learners. *International Journal of Social Science and Human Research*, 04(12). <https://doi.org/10.47191/ijsshr/v4-i12-30>
- Benjamin, A. and Pashler, H. (2015). The Value of Standardized Testing: A Perspective From Cognitive Psychology. <https://journals.sagepub.com/home/BBS>.
- Brown, G. & Yule, G. (1983). *Teaching the spoken language* (Cambridge University Press., Ed.).
- British Council. (2014). *IELTS Scoring Guide*.
www.Ieltsessentials.Com/Pdf/Bandcoredescriptorsspeaking.
- Burns, A. (2012). A holistic approach to teaching speaking in the language classroom. *Symposium 2012*, 165–178.
- Kaid, A. and Haq, A. (n.d.). *Spoken English Skills and Saudi EFL Undergraduates: A Case Study*.
- Fulcher, G. (2018). Assessing spoken production. *Tesol International Association*, 1–6.
- Ginther, A. (2012). Assessment of speaking. *The Encyclopedia of Applied Linguistics*.
- Holroyd, C. (2000). Are assessors professional? Student assessment and the professionalism of academics. *Active Learning in Higher Education*.
- Jakirlic, N. and Hoppe, E. (2021). Standardized Tests as Predictors of Success in Health Professions Education: a Scoping Review. *Optomtric Education*: volume (47)1.
- Jones, D. (2011). *Speaking and Listening: Planning and Assessment* (Jones D. & Hodson P., Eds.). Routledge.
- Mohammed Ali, J., Shamsan, M., Guduru, R. and Yemmela, N. (2019). Attitudes of Saudi EFL Learners towards Speaking Skills. *Arab World English Journal*, 10(2), 253–364. <https://doi.org/10.24093/awej/vol10no2.27>
- Navidinia, H., Mobaraki, M., and Malekzadeh, F. (2019). Investigating the effect of noticing on EFL students' speaking accuracy. *International Journal of Instruction*, 12(1), 83–98.
- Rashid, A., Khan, S., Sheikh, F. (2022). Investigating EFL Teachers' Perceptions of the Use of YouTube in the Development of Saudi EFL University Students' Speaking Skills. *Journal of English Language Literature and Education*, volume 3(4).
- Richards, J. (2008). *Teaching listening and speaking from theory to practice*. Cambridge University Press.

- Roca-Varela, M. (2013). How are spoken skills assessed in proficiency tests of general English as a Foreign Language? A preliminary survey. *Palacios, I. M.* , 53–68.
- Shaaban, K. (2005). Assessment of young learners. *English Teaching Forum*, 43(1), 34–40.
- Tracing, A. (2011). *Ways to Proficiency in Spoken English*' (Pawlak M., Waniek-Klimczak E., & Majer J, Eds.).
- Ulker, V. (2017). *Journal of Education and Practice* www.iiste.org ISSN (Volume 8, Issue 32). Online. www.iiste.org
- Yang, X. (2010). How to achieve authentic context in classroom oral English teaching. *Journal of Language Teaching and Research*, 1(3), 339–342.
- Younes, Z. and Albalawi, F. (2016). Investigating the Factors Leading to Speaking Difficulties: Both Perspectives of EFL Saudi Learners and Their Teachers. *Arab World English Journal*, 7(2), 268–287.
- Zaim, M., Refnaldi, R. & Arsyad, S. (2020). Authentic assessment for speaking skills: Problem and solution for english secondary school teachers in Indonesia. *International Journal of Instruction*, 13(3), 587–604.
<https://doi.org/10.29333/iji.2020.13340a>



مجلة الجمعية السعودية العلمية للمعلم
Journal of the Saudi Scientific Association for the teacher

دورية علمية نصف سنوية - محكمة

المجلد الثالث- العدد الأول

رجب ١٤٤٧هـ - يناير 2026 م