مجلة النخبة للدراسات والأبحاث

الإصدار الثالث – العدد الأول – البحث الثاني

10 / يناير / 2024م

Jordan - Amman
ejsr@alno5ba.com
www.ejsr.alno5ba.com

" تقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية في ضوء معايير جودة اختبارات اللغة الإنجليزية العالمية"

إعداد الباحث:

محمد علي عبد الله العباسي

إشراف:

أ.د هاشم بن أحمد الصمداني

أستاذ المناهج وطرق تدريس اللغة الانجليزية بجامعة أم القرى

متطلب تكميلي لنيل درجة الماجستير في تقويم المناهج والبرامج التعليمية

الفصل الدراسي الأول

2023م/ 1445هـ

مقدمة :

تحظى اللغة الإنجليزية باهتمام كبير في عصـرنا الحالي نظرًا لكونها اللغة المسـتخدمة في أغلب المجالات على مسـتوى العالم، فاللغة الإنجليزية هي الأكثر انتشارًا في مجالات التجارة والأعمال والطب والعلوم على اختلاف أنواعها، كما أصبحت اللغة الإنجليزية من أهم وسائل الاتصال بين الشعوب على المستوى العالمي (الفلاحات، 2020).

وتعد اللغة الإنجليزية من اللغات الأسـاسـية للتواصـل على المسـتوى العالمي، حيث يتم الاعتماد عليها في التعبير عن الأفكار والثقافات، كما أنها اللغة الأهم التي يتم من خلالها اكتسـاب المعارف والمهارات المختلفة وفي شـتى المجالات، ويرجع ذلك إلى غزارة المؤلفات التي يتم نشـرها باللغة الإنجليزية، ولا شـك في أن الصعوبات التي يواجهها الأفراد حول العالم في استخدام اللغة الإنجليزية يؤثر وبصورة واضحة على قدرة هؤلاء الأفراد على التواصل والتطور في مجالات الحياة المختلفة، وبذلك فقد أصـبح للغة الإنجليزية قيمة جوهرية في حياة شـعوب العالم بأسـره، لما تحمله هذه اللغة من علوم ومعارف حديثة، ولما تلعبه من دور في تسـهيل التواصـل العالمي (طاشكندى، 2018).

يعد التقويم من العناصـر الرئيسـة في المناهج الدراسـية، والتقويم الناجح هو الذي يؤدي إلى تغيير في بعض الأهداف وتعديل البعض الآخر، وإلى تحسـين وتطوير طرائق التدريس والوسـائل المتبعة، ويلقي الضـوء على المشـكلات التي تواجه العملية التربوية وعلى جوانب القوة والضـعف في عناصـر المنهج (جبارة، 2017، 60). التقويم مكون ضروري للاستقصاء عن كيفية تعلم الطلبة، وكيفية تحقيقهم لمعايير قياس الإنجاز المحددة، مع تحديد مستوياتهم بدقة، كما أن التقويم يقدم التعزيزي اللازم للطلبة، من خلال تفسير تقدمهم وإنجازهم، وبالإضافة إلى منحهم الوثائق التي تؤكد تقدمهم في المستويات التعليمية (Maxwell, 2012,688).

وتعتبر الاختبارات من أكثر وسائل التقويم شيوعًا في التعرف على ما يمتلكه الأفراد من قدرات ومعارف ومهارات، والاعتماد عليها في اتخاذ قرارات مهمة (البوعبيد، 2016)، كما تعد الاختبارات اللغوية إحدى أهم أدوات القياس والتقويم كونها تهدف إلى قياس سـلوك الطالب وكفايته اللغوية، حيث أشـار زكي (2016) أنها مدخلات لغوية يقوم المعلم بتصـميمها في صـيغة أسـئلة موجهة للمتعلم، بهدف قياس أدائه من كفايات تواصـلية ونحوية ومهارات لغوية، ويقاس هذا الأداء بدرجات كمية قياسـية وهي مجموع ما حصـل عليه الطالب من درجات موزعة على ورقة الأسـئلة، أو درجة كيفية تقييمية تسـتخدم في قياس الأداء الشـفهي تحدثًا، وعلى ضوء هذه الدرجات يتم اتخاذ قرار ترفيع المتعلم أو إبقائه في المستوى اللغوي نفسه.

وقد أكدت هيئة تقويم التعليم والتدريب على أن الاختبارات اللغوية أحد أبرز مسائل القياس إن تمت وفق منهج موضوعي دون تحيز، وتوفر الهيئة العديد من الاختبارات التي تنقسـم إلى قسـمين، أولاهما متعلق باللغة العربية حيث تقيس هذه الاختبارات كفايات اللغة العربية لدى المتقدين سـواء كانوا من الناطقين بها أو من غير الناطقين بها، إضافةً إلى الاختبارات باللغة الإنجليزية، وتشمل اختبارًا لقياس الكفايات اللغوية (STEP)، واختبارًا لقياس وتحديد المستوى في اللغة الإنجليزية (EPT)، وتهدف الهيئة من الاختبارات اللغوية تقديم خدمة للمؤسـسـات التعليمية كافة لتسـهيل عملية اسـتقطاب المتعلمين أو المهنيين المؤهلين (هيئة تقويم التعليم والتدريب، 2022).

هذا وقد دأب المركز الوطني للقياس الذي يمثل الهيئة العلمية التربوية المستقلة والمتخصـصـة في تطوير المقاييس العلمية والمهنية بالوقوف على الأطر والمعايير المنوطة بتعليم اللغات والتي تبنى عليها الاختبارات المقننة عالميًا، بل وتستند عليها باعتبارها مرجعًا مهمًا في هذا المجال، ومن أبرز هذا المعايير والمرجعيات معايير المجلس الأمريكي لتعليم اللغات (the ACTFL Proficiency Guidelines 2012)، حيث قام المجلس بتطوير توصيفات ومعايير وأدوات في اللغات الأجنبية في مختلف المهارات اللغوية، إضـافةً إلى الإطار الأوروبي المشـترك للغات (Common European Framework of reference for (CEFR) Languages, 2009)، والذي وضـع بإشـراف الاتحاد الأوروبي وله سـت مسـتويات مبنية على مبدأ (ما الذي يسـتطيع المتعلم تقديمه تجاه اللغة في مسـتواه ؟) (المركز الوطني للقياس، 2021).

وقد شـكلت العديد من الدواعي والأسباب احتياجًا حقيقيًا لوضـع اختبارات مقننة يتم اسـتخدامها لقياس الكفايات اللغوية لدى المتعلمين، وقد بادرت هيئة تقويم التعليم والتدريب إلى ذلك بفعل حاجة اللغة إلى برنامج تنمية لغوية يشـمل طرائق اكتسـاب اللغة، وطرائق تعليمها وتعلمها، والأدوات العلمية العالمية ذات المنهجية والمرجعية والمصداقية التي تقيس مستويات المتعلمين في هذه الكفايات، أو تبين مدى إتقانهم للكفايات التي اكتسبوها، إضافةً إلى حرص العديد من الجهات

التربوية والتشـــغيلية قياس جودة إتقان اللغات واتخاذ بعض من هذه الاختبارات كمعيار لقبول الطلاب في برامج علمية مختلفة، كما أن هناك حاجةً لدى العديد من الأفراد والجهات المهنية إلى مؤشرات ومعايير يحتكم إليها لمعرفة مستوى إتقان أفراد معينين للمهارات الأساسية في اللغات المدروسة (المركز الوطني للقياس، 2021).

ومما سبق، ونظرًا لعمل الباحث معلمًا للغة الإنجليزية، واستشعارًا منه لضرورة القيام بدراسة علمية لتقويم الاختبارات في ضوء معايير جودة الاختبارات العالمية للغة الإنجليزية. فقد تبنى القيام بهذه الدراسة للتعرف على تقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية في ضوء معايير جودة اختبارات اللغة الإنجليزية العالمية .

مشكلة الدراسة:

من خلال اطلاع الباحث على الأدب البحثي في هذا الموضـــوع وجد ندرة في الأبحاث والدراســـات التي تناولت تقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية في ضوء معايير جودة اختبارات اللغة الإنجليزية العالمية، وذلك في ظل ارتفاع الأصوات المنادية بضرورة تحسين جودة الاختبارات.

ويرى الباحث أنه نتيجة للمستجدات التي تطرأ بشكل مستمر وما يصاحبها من تغيرات متسارعة، فإن هناك حاجةً ملحة إلى إعادة النظر في اختبارات هيئة تقويم التعليم والتدريب اللغوية وفق معايير جودة اختبارات اللغة الإنجليزية العالمية ذائعة الصيت، خاصةً وأن الباحث لم يجد أية دراسات سابقة تناولت تقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية في ضوء معايير جودة اختبارات اللغة الإنجليزية العالمية، ولذلك حدد مشكلة بحثه الرئيسة بتقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية في ضوء معايير جودة اختبارات اللغة الإنجليزية العالمية. ولذلك حددت مشكلة الدراسة بالسؤال الرئيس التالي:

"ما مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبارات هيئة تقويم التعليم والتدريب اللغوية؟"

أسئلة الدراسة:

يتفرع من السؤال الرئيسي للدراسة السؤالان التاليان:

1. ما مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار كفايات اللغة الإنجليزية STEP؟
2. ما مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار تحديد المستوى في اللغة الإنجليزية EPT؟

أهداف الدراسة :

هدفت الدراسة الحالية إلى الكشف عن مدى تحقق:

1. تقويم اختبار كفايات اللغة الإنجليزية STEP في ضوء معايير جودة الاختبارات العالمية.
2. تقويم اختبار تحديد المستوى في اللغة الإنجليزية EPT في ضوء معايير جودة الاختبارات العالمية.

أهمية الدراسة:

تتلخص أهمية هذه الدراسة فيما يلي:

أولا: الأهمية النظرية:

تتمثل الأهمية النظرية في الآتي:

1. يســـهم البحث في تقديم رؤية تطويرية لدى مشـــرفي اللغة الإنجليزية في المراحل المختلفة لتطبيق معايير جودة الاختبارات العالمية في التعليم والذي بدوره قد يساعد مطوري المناهج مستقبلا في الاستفادة منها من خلال تصميم المناهج الدراسية وفقًا لهذه المعايير.
2. يفتح البحث أمام الباحثين ومطوري المناهج في المملكة العربية السعودية آفاقًا لاستراتيجيات تدريسية تطويرية حديثة تساعد المعلمين في تكييف المناهج في ضوء الاختبارات الدولية لضمان مواكبة التطورات في عالم اللغات.

3. يتوافق موضوع البحث مع التوجهات والاستراتيجيات والأساليب الحديثة عالميًا، والتي تستهدف تطوير وتحسين تعلم اللغات، لا سيما معايير المجلس الأمريكي لتعليم اللغات، والإطار الأوروبي المشترك للغات.
4. يُتأمل من هذا البحث أن يقدم توصيفًا دقيقًا لنقاط القوة ومواطن الضعف في جودة اختبارات هيئة تقويم التعليم والتدريب اللغوية.
5. مسايرة البحث للتوجهات العالمية الحديثة التي تؤكد على ضرورة استخدام أساليب تقويمية جديدة وحديثة في العملية التعليمية.

ثانيا: الأهمية التطبيقية

تتمثل الأهمية التطبيقية للدراسة في:

1. يُتأمل من نتائج هذه الدراسة أن يلفت انتباه مصممي الاختبارات اللغوية في هيئة تقويم التعليم والتدريب نحو مراجعة هذه الاختبارات، والعمل على تحسينها وتطويرها إن كان هناك حاجة.
2. يُتأمل من هذه الدراسة أن تسهم في تجويد عمليتي التعليم والتعلم، وبالتالي تحسين مستويات متعلمي اللغة الإنجليزية.

حدود الدراسة:

اقتصرت الدراسة على الحدود التالية:

- **الحدود الموضوعية:** اقتصرت الدراسة على تقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية في ضوء معايير جودة اختبارات اللغة الإنجليزية العالمية.
- **الحدود الزمنية:** طبقت الدراسة خلال العام الدراسي 2021م- 1443هـ.
- **الحدود البشرية والمكانية:** اقتصرت الدراسة على عينة من المشرفين التربويين والمعلمين المتخصصين الملتحقين بالاختبارات اللغوية لدى هيئة تقويم التعليم والتدريب اللغوية بمنطقة مكة المكرمة.

مصطلحات الدراسة:

التقويم :

يعرف التقويم على أنه: "العملية التي يتم من خلالها جمع وتحليل البيانات ومن ثم اتخاذ القرارات الأنسب بشأنها وهو أحد مقومات العملية التعلمية، والذي من خلاله يمكن التعرف على مستوى الإنجاز الذي حققه المعلم" (خليفة، 2014: 489).

ويعرفه الباحث إجرائيًا على أنه: العملية التي يتم من خلالها جمع البيانات والمعلومات الكمية عن طريق الاستعانة بطرق القياس المختلفة بغرض التوصل إلى مجموعة من الأحكام التي تشير إلى مستوى ومدى فاعلية العملية التربوية، حيث يتم الاستناد إلى مجموعة من المعايير التي تسهم في تحقيق أهداف التقويم التربوي.

الاختبارات اللغوية:

تعرف الاختبارات اللغوية على أنها: "مجموعة من المدخلات اللغوية المفهومة التي يتم تصميمها في صيغة أسئلة تحريرية أو شفهية يتم توجيهها للطالب لكي يجيب عنها، وذلك بهدف قياس أداء الطالب فيما يتضمنه الاختبار من كفاءات نحوية أو تواصلية أو مهارات اللغة وعناصرها التي اكتسبها الطالب وهما نوعان Step,Ept": (السيد، 2016: 30).

ويعرف الباحث الاختبارات اللغوية على أنها: مجموعة من الاختبارات التي يتم تصميمها من قبل هيئة معتمدة بهدف قياس مدى اكتساب متعلمي اللغة للمهارات والكفاءات اللغوية المختلفة.

أولا: الإطار النظري

مقدمة:

لامست الفترات الأخيرة والسنوات الماضية ثورةً وتطورًا كبيرًا في مفهوم عملية التقويم وأدواته المختلفة، حيث أصبح التقويم يمتلك أهدافًا متنوعة وجديدة فتحول أسلوبه من المدرسة السلوكية التي تشير إلى وجود أهداف عالية التحديد تصاغ بسلوك قابل للقياس والملاحظة إلى أسلوب المدرسة المعرفية التي تعتمد بصورة أساسية على كل ما يجري في داخل عقل المتعلم من مختلف العمليات العقلية التي تؤثر في طبيعة سلوكه وتهتم بعمليات التفكير المختلفة وبشكل خاص عمليات التفكير العليا، كعملية بلورة ومحاكاة الأحكام ومن ثم اتخاذ القرارات ومهارات حل المشكلات باعتبار أنها مهارات عقلية تساعد الإنسان من التعامل مع كافة معطيات عصر المعلوماتية، والتقنية المتسارعة والمتطورة وكذلك تفجر المعرفة (خليفة، 2014).

ويرتكز التقويم بصورته التقليدية على استخدام الاختبارات بأشكالها وصورها المختلفة وهذه الاختبارات يتم منحها مرة واحدة فقط أو عدة مرات خلال الفصل أو العام الدراسي وذلك للحصول على معلومات تدل على مستوى تحصيل الطلبة ليتم تقديمها فيما بعد لأولياء الأمور، أو إلى أية جهة أخرى معنية بعملية التعليم، وتلك الأشكال والصور من التقويم تعتبر ذات تأثير ضعيف على عملية التعليم ولا يمكنها أن تساهم في إنتاج أجيال قادرة على إدارة وتحريك عجلة ومحور التنمية بشكل فعال، وذلك لأنه يعتمد على قياس مفاهيم ومهارات بسيطة ومحددة والتي يتم الإشارة إليها والتعبير عنها بالاستعانة بالأرقام دون إظهار أية معلومات أو بيانات ذات أهمية أو قيمة (Banse et al., 2017).

وقد عملت العديد من المؤسسات والمراكز المتخصصة في التحضير لاختبارات اللغة ومساعدة الطلبة على تحقيق أعلى مستوى فيها، ومن أهم وأبرز تلك المراكز المركز الوطني للقياس والتقويم في المملكة العربية السعودية، والذي يعمل على بناء أنظمة خاصة بالاختبارات المحوسبة والعمل على توفيرها بصورة دائمة طوال العام بدلا من العمل على تحديد أوقات معينة في السنة لتقديمها، الأمر الذي يشجع الطلبة والمتقدمين للتقدم لها في أي فترة زمنية خلال العام، فقد عمل المركز الوطني على إنشاء العديد من المقرات الخاصة بالاختبارات المحوسبة والتي تساعد في بناء وتوفير بيئة ملائمة ومناسبة للطلبة وجعلها موحدة على كافة المقرات، ومن أبرز الاختبارات التي ينظمها المركز الوطني اختبارات كفايات اللغة الإنجليزية STEP واختبارات القدرات العامة لكل من الطلاب والطالبات (الفرعين الأدبي والعلمي)، وكذلك اختبارات القدرات العامة الخاصة بالطلبة الجامعيين (أبو عيش، 2017).

المحور الأول: التقويم

يشهد عصرنا الحالي تطوراتٍ متنوعةً ومختلفةً في العديد من المجالات وعلى رأسها التعليم، والذي شهد عملية تطوير متكاملة وشاملة لمختلف الجوانب التعليمية، الأمر الذي استدعى تطوير كافة الأدوات والطرق التقويمية بصورة مستمرة وذلك بغرض الوصول إلى الغايات والأهداف المرسومة بشكل جيد، وقد شمل التطوير للعملية التعليمية كافة جوانب عملية التقويم التربوي (Murray, 2016).

ويعتبر التقويم التربوي أحد أهم وأبرز العمليات المنهجية التي تتطلب من القائمين عليها جمع البيانات والمعلومات الموضوعية والصادقة من مختلف المصادر المتنوعة والمتعددة عن طريق الاستعانة بمجموعة من الأدوات المختلفة في إطار الأهداف المرسومة والمقصودة، وذلك للوصول إلى تقديرات وصفية وكمية وأدلة يستند إليها في إصدار كافة الأحكام المتوقعة، أو العمل على اتخاذ قرارات تتعلق بالأفراد المتعلمين.

وبذلك يمكن الإشارة إلى أن التقويم التربوي يلعب دورًا مهمًا وبارزًا في توجيه وضبط سير العملية التعليمية، وبالتالي يساعد على ضبط الإدارة الصفية للفصل المدرسي أو الجامعي، ويعمل على إثراء مستوى التعلم للطلاب وتساعد في تنشيط مستوى التقدم الدراسي لديهم، وبالتالي يكون هناك تغيرٌ واضحٌ وجيد في مستوى المخرجات التعليمية (Salehi et al., 2019).

مفهوم التقويم التربوي :

يعرفه (Murray, 2016) على أنه عبارة عن عملية يتم من خلالها إعطاء وإصدار حكم على الظاهرة المراد فحصها وقياسها في إطار ما تحتويه وتتضمنه من خصائص وسمات معينة ومحددة.

بينما يعرفه (خليفة، 2014) على أنه عملية يتم من خلالها جمع وتحليل البيانات ومن ثم اتخاذ القرارات الأنسب بشأنها وهو أحد مقومات العملية التعلمية، والذي من خلاله يمكن التعرف على مستوى الإنجاز الذي حققه المعلم.

في حين يعرفه (Ambag, 2014) بأنه عملية تثمين ومحاسبة لكل الأمور التي تجري من تغييرات في التعليم والتعلم والتربية والسلوك مقارنةً بالأهداف المسطرة والموضوعة بشكل مسبق، والتي يمكن من خلالها تقدير الإنجازات المحققة من أعمال تطبيقية أو نظرية في الجوانب التربوية والتعليمية المختلفة والمتنوعة.

يعرفه الباحث على أنه العملية التي يتم من خلالها جمع البيانات والمعلومات الكمية عن طريق الاستعانة بطرق القياس المختلفة بغرض التوصل إلى مجموعة من الأحكام التي تشير إلى مستوى ومدى فاعلية العملية التربوية، حيث يتم الاستناد إلى مجموعة من المعايير التي تسهم في تحقيق أهداف التقويم التربوي.

تطور مفهوم التقويم :

تطور مفهوم التقويم التربوي حيث أنه مرَّ بمجموعة من المراحل المتتابعة والتي أسهمت في تشكيل التقويم التربوي في صورته التي نلمسها اليوم، فقد مرّ التقويم التربوي بأربع مراحل أساسية كانت كالتالي:

المرحلة الأولى :

وهي لبنة الأساس في عملية التقويم التربوي، وقد ظهر في هذه المرحلة اختبارات بينية باعتبارها مقياسًا عقليًا بالمفهوم والمصطلح المتعارف عليه، ومن ثم ظهرت اختبارات التحصيل المنقحة في مجموعة من المواد المختلفة، ومنها الاختبار المنقح في الرياضيات والاختبارات الخاصة بالمفردات اللغوية.

المرحلة الثانية :

وتتميز هذه المرحلة بانتشار المقاييس المختلفة والاختبارات المتعددة في التقويم التربوي بشكل كبير وواسع، وتتميز هذه المرحلة بتصميم عدد كبير من الاختبارات المنقحة في مختلف المواد والمهارات التعليمية ذات الأهداف المختلفة، وقد سعى المتخصصون في التربية وعلم النفس إلى تطبيقها بدرجة أبعدت توجههم عن الاتجاه المنقب والناقد (Sawyer, 2018).

المرحلة الثالثة :

وامتازت هذه المرحلة باستخدام وتفعيل أسلوب ناقد في توظيف الاختبارات فقد اتجه كل من علماء التربية وعلماء النفس إلى تحسين ماهية نوعية الاختبارات والعمل على تنقيتها وتحسين مستوى اختيارها في كل مرة يحاولون تطبيقها، ولم يعد الأمر قاصرًا على قياس مجالات ومهارات متفرقة، وحينها اتجهوا إلى تقويم التحصيل بصورة شاملة ومتكاملة.

المرحلة الرابعة:

ووصل في هذه المرحلة التقويم التربوي إلى مستوى دقيق جدا مقارنةً بالمراحل السابقة، وتنوعت أساليب التقويم اللغوي المستخدمة حيث أصبح مصطلح التقويم اللغوي يرتبط مباشرة بمصطلح التقويم القائم على الأداء والذي يعتمد بصورة أساسية على مقارنة أداء الطالب أو المتعلم بمحك أو مجال معين، حيث أصبح التقويم التربوي يهتم في هذه المرحلة بماهية المواصفات التي تم تحديدها بناءً على المستويات المعيارية (Salehi et al., 2019).

ومما سبق يتبين أن عملية التقويم التربوي قد مرت بمجموعة كبيرة من المراحل المتنوعة والمختلفة والتي أحدثت تغييرات كبيرة وملحوظة في كافة مجالات البيئة التعليمية وعناصرها بشكل كامل، الأمر الذي يشير إلى وجود مراحل متطورة ومتسلسلة ووجود أيدي بناءة عملت على تطوير هذه العملية ودفعتها لتظهر بصورتها الحالية لنا وأسهمت في تحقيق أهدافها بشكل واضح وصريح دون وجود لبس أو خلل فيها.

التقويم وعلاقته بالاختبار والقياس والتقييم:

ترتكز العلاقة بين كل من التقويم والاختبار والقياس على مجموعة من البنود والمعايير والتي تتمثل في أن القياس يعتبر سابقا لعملية التقويم وأساسا بنائيًا له، وأن عملية القياس تكون أكثر شمولا واتساعا من الاختبار، كما يشكل التقويم مصطلح أوسع وأشمل وأعم من كل من التقييم والقياس، ويعتبر الاختبار مقياسًا والعكس غير صحيح، ويشكل الاختبار عملية تقييم لجانب واحد بينما التقويم يكون أشمل وأعم منه فيكون لعدة جوانب ويعتبر عملية دائمة ومستمرة، إضافةً إلى أن الاختبار عملية تعتمد على شخص واحد وهو المعلم الذي يختار المكان والزمان اللازمين لتطبيقه في حين أن التقويم عملية تعاونية تلزم إشراك وتفعيل كافة العناصر المختلفة، ويعتبر كل من الصدق والثبات من المعايير والمتطلبات اللازمة لعملية تطبيق وسائل وأدوات القياس المختلفة (Ali & Ahmed, 2021).

وقد أضاف (Salehi et al., 2019) أن التقويم يحتاج إلى تفعيل واستخدام كافة المستويات والمعايير وكذلك المحكات كأسس بغرض الحكم على الدرجات الصادرة من عملية القياس، في حين أن القياس يعمل على تحديد قيم للظاهرة بينما عملية التقويم تصدر حكم عليها، وتعتبر النتائج محور اهتمام عمليتي التقويم والقياس، مع قصور كل منهما إلى حدود وظيفته الأساسية، كما يشكل القياس حجر الزاوية باعتباره مكونًا في عملية التقويم، ويقوم القياس بوظائف معينة ومحددة وترتكز وظائفه في الوصول إلى النتائج في حين أن التقويم يقوم بمجموعة من الوظائف والمهام المختلفة والتي تتمثل في التشخيص والعلاج واختيار الوسائل وتحديد ماهية الأهداف، إضافةً إلى أن نتائج عملية التقويم يتم الاستفادة منها في مساعدة المتدربين والمتعلمين للتقدم بمستوياتهم المختلفة.

يشير الباحث إلى أن الأنشطة التعليمية والتربوية تشتمل على مجموعة من العمليات المتكاملة والمترابطة مع بعضها البعض والتي تساعد في إدارة تلك الأنشطة بأفضل آلية ممكنة، ومن أبرز العمليات التي تتم من خلالها الأنشطة التربوية والتعليمية عمليات التقويم والتقييم والقياس والاختبار التي يكون لكل منها دور مهم وبارز مختلف عن غيرها من العمليات، وتشكل كل عملية من تلك العمليات أداةً بارزة بيد المعلم وأصحاب القرار الذين يمكنهم من خلالها إدارة الأنشطة التعليمية بصورة واضحة وبشكل جيد مما يحقق الأهداف المنشودة والمرسومة.

المبادئ التي يقوم عليها التقويم السليم :

يتميز التقويم التربوي بمجموعة من الركائز والمبادئ الرئيسة، والتي تتمثل في أن التقويم عملية إنسانية، وأنه وسيلة لا غاية، كما تعتمد عملية التقويم التربوي على التنويع في الأساليب والأدوات، ومن المهم الوعي بماهية مصادر الأخطاء المتنوعة والمختلفة المحتملة منها، والتي تتمثل في الخطأ في التخمين، والخطأ في اختيار العينة، وتحديد سمات وصفات شخصية المقوم، إضافةً إلى التركيز على أثر الهالة (Sawyer, 2018).

يبين الباحث أن عملية التقويم عبارة عن مجموعة من الركائز والمبادئ التي يقوم من خلالها أصحاب الشأن والقرار بتحديد غاياتهم وأهدافهم ورسمها والتخطيط لها بشكل يمكن تحقيقه، بالإضافة إلى أن تلك المبادئ تشكل أعمدة بنيان لعملية التقويم التربوي بحيث تساعد المعلم في تفعيل الأنشطة التعليمية المختلفة، وتدفعه للوصول إلى المطالب التعليمية المرسومة والمختلفة.

أهمية التقويم التربوي :

يعد لتقويم ذات أهمية كبيرة للقائمين والمسؤولين عن العملية التعليمية، الذين ينغمسون في مختلف المجالات التعليمية والتربوية ويشعرون بمشكلاته وتحدياته المختلفة ويسعون للوصول إلى حلول لتلك المعيقات والمشكلات، إضافةً إلى تحديد ماهية جوانب وبنود القوة والضعف والعمل على تنمية مستوى الشعور بالرضا عن طبيعة الأداء فهو يسهم في مساعدتهم على تنمية المستوى الخاص بالجوانب الوجدانية والمهارية والمعرفية، ويعتبر التقويم التربوي ذا أهمية كبيرة بالنسبة للمعلمين حيث يسهم في إلقاء الضوء على مهاراتهم التعليمية وبالتالي يساعدهم على التعرف على تحديد بنود القوة والضعف في تعيين الأهداف التعليمية، ومن ثم البحث عن أفضل السبل التي تساعد في تحقيق تلك الأهداف، حيث يشكل التقويم التربوي عنصرًا وركيزة في العملية التعليمية وتطوير محتوى المناهج التعليمية فمن خلاله يمكن التعرف على فعالة وماهية البرامج الدراسية والتعليمية (Villafuerte et al., 2017).

ويرى الباحث أن عملية التقويم التربوي من أهم وأبرز العمليات التي يجب أن تأخذ حيزًا كبيرًا من الاهتمام من قبل المتخصصين بهذا الشأن، وذلك لما له من أهمية كبيرة على مختلف الأصعدة والجوانب المتعددة والتي من أهمها وأبرزها أن التقويم يعمل على تحديد اتجاه المؤسسة التعليمية ومستوى تحقيقها لأهدافها، ويساعد التقويم التربوي على تشخيص التحديات والصعوبات التي يمكن أن تعترض طريق المتعلم والمعلم والعملية التعليمية بشكل كامل الأمر الذي يسهم في حلها

وإيجاد طرق للقضاء على مواطن الضعف فيها، ويساعد التقويم في تنمية مهارة الحكم الذاتي لدى الطالب مما يساعده في إصدار الحكم على ذاته وتنميتها وتطويرها بشكل سلس وبصورة أفضل.

أهداف عملية التقويم التربوي:

تهدف عملية التقويم التربوي إلى تحقيق مجموعة من الأهداف والغايات التربوية، ومن أهم وأبرز تلك الأهداف المساعدة في الحكم على فعالة ومدى نجاح العملية التعليمية، وتنقيح المحتوى التعليمي ومن ثم مراجعته والحكم على أساليب التدريس المتنوعة والمختلفة، ومساعدة المؤسسة التعليمية على التعرف على ما حققته من رسالة وغاية تربوية بحيث يتم تزويد المعلم بمحتوى جيد من التغذية الراجعة التي تزيد وترفع من مستوى العملية التعليمية، بالإضافة إلى الربط بين أهداف ومحتوى الخطة الدراسية وطبيعة العمل الذي يقوم به المتعلم وكذلك مساعدة المدرس في الوقوف على مستوى نجاحه في تعليم الطلبة وتربيتهم ومدى نجاحه في تحقيق الأهداف التربوية، كما يعمل التقويم التربوي على قياس المهارات اللغوية المتنوعة والمختلفة في كافة الأنشطة التعليمية وخصوصا مواد اللغات المختلفة كالمهارات السماعية، والكتابية، والتعبيرية وكذلك المهارات القرائية (Salehi et al., 2019).

بين الباحث أن التقويم التربوي وعلى وجه التحديد التقويم التربوي الحديث قد جاء ليعمل على تحقيق مجموعة من الأهداف المتنوعة والمختلفة والتي تغطي كافة الجوانب التعليمية والتربوية المتنوعة، ومن أهم وأبرز أهداف التقويم التربوي وفق ما أوضحه الباحث هي تنمية وتطوير المهارات الحياتية الحقيقية، وتفعيل مبادئ العمل التعاوني، ودمج كل من التقويم الأدائي والكتابي، والعمل على تحسين عملية رصد المستوى التعليمي للطلبة بشكل دائم ومستمر، وكذلك تشجيع وتنمية مستوى التفكير بغرض تعميم أفضل الإجابات الممكنة، وتحسين مستوى التوافق بين أنشطة التعليم ومتطلباته وكذلك نتائجه.

وظائف وأغراض عملية التقويم التربوية :

يعمل التقويم على تشخيص وعلاج العملية التعليمية وتحديد ماهية مواطن الضعف والقوة في العملية التعليمية، ويسهم في اكتشاف القدرات التي يمتلكها الطلبة وماهية ميولهم واتجاهاتهم ومن ثم العمل على تنمية وتطوير تلك المهارات والقدرات، كما أنه يساعد في تحسين وتطوير وكذلك تحديث المناهج التعليمية.

وقد أشار (Ali et al., 2021) إلى أن وظائف التقويم التربوي تتمثل في:

1. يعمل التقويم التربوي على إفادة المؤسسات والمنظمات التعليمية في التعرف على مستوى تحقيقهم لأهدافهم التي يسعون إليها وماهية نسبة الإنجاز لهم.
2. التعرف على مدى اكتساب الطلبة والمتعلمين في فترة تعلمهم لمجموعة من المهارات والقدرات المتنوعة والتي تكونت عندهم بصورة تلقائية نتيجة ممارستهم لأوجه وأشكال النشاطات التربوية المختلفة، وكذلك الحكم على مستوى الفائدة التي قد حصل عليها الطلبة من هذه الأنشطة.
3. الوقوف على الحالات الخاصة والمرضية عند الطلبة المتعلمين في الوقت المناسب والعمل على علاجها وتخفيف مواطن الضعف فيها، من خلال تفعيل الإرشاد والتوجيه النفسي من قبل المتخصصين في تلك المجالات ومعالجة حالات التخلف والضعف الدراسي.
4. تعمل على تزويد المتعلمين بنتائج عملهم بشكل مستمر أي تمنحهم التغذية الراجعة اللازمة مما يسهم في تقييم ودعم أدائهم.
5. مساعدة الإدارة المدرسية في توزيع الطلبة على مجموعة من الفصول الدراسية وعلى أنشطة تعليمية مختلفة.
6. مساعدة أصحاب وذوي الشأن والقرار في العملية التعليمية لتقييم أهداف التعليم بصورة مستمرة وفق مجموعة من المعايير المخطط لها والتي تسهم في تحقيق حاجيات المجتمع المختلفة.
7. المساعدة في عمليات تحديث وتطوير المناهج التعليمية المختلفة، لتتوافق مع التطور التقني والتكنولوجي المتسارع.
8. الإرشاد نحو اتخاذ القرارات الصائبة والصحيحة والتي تتعلق بترقية الأفراد أو العمل على تعيينهم عن طريق المعلومات التي تقوم بتوفيرها أدوات القياس.

أسس التقويم السليم :

نتيجةً للأهمية الكبيرة التي يحتلها التقويم في كافة المراحل التي تمر بها العملية التعليمية فقد سعى المختصون والخبراء في مجال عمليات التقويم إلى وضع مجموعة من الركائز والأسس والمعايير والتي يمكن من خلالها بناء وتصميم التقويم المناسب والملائم لطبيعة العملية والبيئة التعليمية، ومن أهم وأبرز أسس ومبادئ التقويم السليم ما يأتي(Banse et al., 2017):

1. الاستمرارية: يجب أن تكون عملية التقويم التربوي عبارة عن تكوين متكامل ومساير لأهداف العملية التعليمية في كافة مراحلها المتنوعة.
2. الشمولية والتعميم: وهذا يعني أن محتوى التقويم التربوي يشمل كافة الجوانب الخاصة بالعملية التعليمية وألا تكون مقتصرة على جوانب متعلقة بالمتعلم فقط.
3. التعاون: إن عملية التقويم التربوي عملية تعاونية يشارك فيها وفي إعدادها كافة أطراف وعناصر العملية التعليمية من معلمين ومتعلمين وأولياء أمور ومحتوى تدريسي، بحيث يتم اتخاذ كافة الإجراءات التي تناسب عمليات علاج جوانب الضعف والقوة المختلفة.
4. يجب أن يتم تحديد ماهية الأهداف المقصودة والمرجوة من عملية التقويم التربوي.
5. يجب أن تكون عملية التقويم التربوي عملية تقويم إنسانية في إصدار الأحكام وفي بناء الأدوات.

أوضح الباحث أن التقويم التربوي يتم بناؤه بالاعتماد على مجموعة من الأسس التي تعمل على تشكيل خط السير للمعلم أو صاحب القرار أي الإداري أو غيره في توجيه العملية التعليمية، وذلك لفحص مدى واقعيتها والعمل على توجيهها بأفضل الاستراتيجيات والأساليب الممكنة والتي تناسب طبيعة الأنشطة التعليمية والتربوية، كما أنها تساعد المعلمين وخصوصا المبتدئين منهم على تفعيل دور التقويم التربوي في تحقيق أهداف العملية التعليمية، ويجب أن يبنى التقويم التربوي على أساس صادق من الأهداف التربوية، وأن يكون شاملًا للمحتوى المتكامل للمادة التعليمية، بالإضافة إلى أنه يجب أن يشكل عملية دائمة غير متقطعة بحيث تسهم في منح المعلم نتائج متتابعة على فترات زمنية متتالية بشكل مدروس.

عناصر عملية التقويم التربوي :

ذيتكون التقويم التربوي من مجموعة من العناصر الأساسية ومنها المعلم، والهيئة التعليمية والأقسام الإدارية، والمنهاج، إضافةً إلى الملاعب والحدائق والمكتبات والمختبرات، والبرامج التعليمية، كالمسابقات والأنشطة والرياضة وغيرها من مختلف الأنشطة، والمتعلم بخصائصه المختلفة المتمثلة في آدابه وشخصيته وأخلاقه وتحصيله (Helda & Syahrani, 2022).

وأشار الباحث إلى أن العناصر التي تشملها عملية التقويم التربوي هي ركائز أساسية لا يمكن لعملية التقويم التربوي أن تتم بدونها أو بدون أحدها، لذلك يتوجب على القائمين على عملية التقويم التربوي التحقق من توافر كافة العناصر أو ما يمكن تسميتها بالمدخلات الخاصة بعملية التقويم التربوي، وتكون تلك المدخلات والعناصر عبارة عن:

- كافة السياسيات والمبادئ التعليمية كالأنظمة واللوائح والقواعد التعليمية.
- المناهج التعليمية والمواد الدراسية ومصادر التعلم المتنوعة كافة.
- الإداريين والمعلمين وكذلك المشرفين وكل من له صلة بطبيعة العملية التعليمية والتحقق من مستوى كفايتهم لمختلف الأعمال المسندة لهم.
- الظروف البيئية المحيطة بالمؤسسة التعليمية ومدى توافقها مع المتعلم.
- البرامج التعليمية والتربوية المتنوعة والمختلفة ومنها البرامج الخاصة بالطلبة الموهوبين، وبرامج التأهيل ورفع مستوى الكفاءة.
- تكنولوجيا وتقنيات التعليم المتنوعة.

خصائص التقويم التربوي الجيد:

يتمتع التقويم التربوي الجيد بمجموعة من السمات والخصائص التي تتمثل في البنود التالية، وهي:

أولا: الصدق

حيث يعتبر الصدق من أهم وأبرز السمات التي تستخدم في عملية التقويم التربوي، ويقصد بهذه الخاصية هي أن تعمل أداة التقويم على قياس السمة أو المهارة التي وضع لأجلها.

ولتحقيق هذه الخاصية في التقويم التربوي لا بد من التأكيد على اقتصار محتوى التقويم على ما تم تعليمه وتدريسه للطلبة، والتدرج في محتوى التقويم بين السهولة والصعوبة، وعدم الإفراط من استخدام المفردات الصعبة جدا أو السهلة جدا، والعمل على جعل محتوى مفردات التقويم بسيطة ومفهومة بالنسبة للطلاب، والاقتصار على المحتوى التعليمي الذي تم تدريسه في داخل إطار المقرر الخاص بالمنهج (Alshwaier et al., 2012).

ثانيا: الثبات

ويمكن الإشارة إليه على أنه عبارة عن درجة الاتساق التي يتم الحصول عليها من النتائج باستخدام أداة التقويم التربوي في حال إذا طبق التقويم التربوي على عينة تم اختيارها عدة مرات في ظروف متشابهة، وتسمى أداة التقويم التربوي بأنها ثابتة في حال لو تم تطبيق التقويم التربوي عدة مرات وتم الحصول على نفس النتائج.

الأمور التي يجب أخذها بعين الاعتبار للوصول إلى سمة الثبات للتقويم التربوي، التحقق من صدق التقويم التربوي، فالعلاقة طردية بين صدق وثبات التقويم التربوي، ومنح المتعلم الوقت الكافي والمناسب للإجابة، وتصحيح التقويم بطريقة موضوعية، إضافةً إلى العمل على كتابة كل من التعليمات والفقرات بدقة كبيرة، وتطبيق التقويم التربوي بدقة كبيرة ومتناهية، والبعد عن الأخطاء التي تحدث بسبب الأصوات الخارجية أو ضيق الوقت.

ثالثا: الموضوعية

يعتبر التقويم التربوي موضوعيًا إذا تم الحصول على نفس التقييم بغض النظر عن الشخص القائم بعملية التقييم والتصحيح.

وأهم القواعد التي تساعد في الوصول إلى مستوى جيد من الموضوعية، أن يكون الهدف والغرض من عملية التقويم التربوي واضحًا وتعليماته صريحة، ويجب أن تكون أسئلته محددة ومخصصة الإجابة، بحيث لا تكون صياغة السؤال مبهمة تعمل على فهم معنى آخر غير المعنى المقصود، إضافةً إلى وجود قواعد ومعايير ثابتة لعملية التقويم التربوي (Salehi et al., 2019).

رابعا: العملية

يجب أن يتسم أسلوب التقويم التربوي بسمة العملية وأن يتمتع بسهولة وبساطة التصميم والأخذ بعين الاعتبار الجهد والوقت اللازمين، وسهولة عملية تقدير الدرجات والعلامات، وسهولة تفسير النتائج، التكلفة المالية تكون في حدود الإمكانيات المتاحة والمتوفرة.

يرى الباحث أن المختصين في العمليات التربوية يحتاجون إلى الإلمام بما يساعدهم في الحكم على مدى صلاحية الموضوعات التي تتعلق بتخصصهم، ويعتبر التقويم التربوي من العمليات التي أحدثت تطورًا وتأثيرًا كبيرًا في العملية التعليمية، حيث مر التقويم التربوي بمجموعة من المراحل المتتابعة والتي أسهمت في تحسين وتطوير العملية التعليمية وتوجههم للوصول إلى أهدافها وغاياتها المقصودة.

أنواع التقويم التربوي:

يتخذ التقويم التربوي مجموعة من الصور والأشكال المختلفة التي تساعد في تحقيق أهداف عملية التقويم التربوي، وقد تم تصنيف التقويم التربوي تبعا لاستخدامه أو زمن إجرائه إلى مجموعة من الأصناف والأنواع والتي تتمثل في (الخروصي، 2021):

أولا: التقويم القبلي:

ويرتكز التقويم القبلي على تقويم وتحديد نتائج العملية التعليمية قبل المباشرة بها، ويتم تطبيق هذا النوع من التقييمات في بداية الحصة الدراسية أو في بداية الوحدة التعليمية أو في بداية الفصل الدراسي، وذلك بغرض تحديد مستوى الاستعداد والجهوزية لدى المتعلمين لتلقي المحتوى التعليمي، وكذلك التعرف على ماهية المداخلات الخاصة بالسلوكيات قبل البدء بإلقاء وتلقين المحتوى التعليمي للمادة الدراسية.

ثانيا: التقويم المستمر أو التكويني

ويلقب هذا النوع من التقويم بالتقويم البنائي وهو عبارة عن عملية تقويم دائمة ومستمرة، يتم ممارستها وتطبيقها أثناء تنفيذ المحتوى الخاص بالمنهج التعليمي بغرض توفير محتوى جيد من التغذية الراجعة تفيد المعلم والمتعلم وكذلك تساعد المعلم في معرفة مستوى التقدم الذي أحرزه المتعلمون لديه.

ثالثا: التقويم الختامي "النهائي":

وهو أحد أنواع التقويم الذي يتم بصورة أو بزمن معاكس للتقويم القبلي، ويتم إجراؤه بغرض فحص محتوى النتائج التي تم التوصل إليها بعد إلقاء المحتوى التعليمي، والذي من خلاله يتم نقل المتعلم من مرحلة دراسية إلى مرحلة دراسية أخرى أعلى منها.

ويشير الباحث إلى أن عملية التدريس والتعليم الجيدة تتطلب استخدام تقويم جيد ودقيق يغطي كافة جوانب العملية التعليمية والتربوية، سواء أكانت مهارات أم قدرات أم جوانب ومجالات وجدانية، وذلك قبل المباشرة بالعملية التعليمية بغرض تحديد مستوى البدء للمتعلمين، أو عند الرغبة في تحديد نقاط القوة والضعف للتمهيد لعلاج مواطن الضعف لدى المتعلمين، فيختار المعلم نوع التقويم الملائم للعملية التعليمية.

الفرق بين التقويم والتقييم :

يعنى مصطلح التقويم عمليات التعديل والإصلاح والتي يمنح من خلالها المتعلم القيمة، ويتم بناؤه بالاعتماد على مجموعة من الأسس والمعايير المحددة بغرض الوصول إلى قرار ملائم، أما التقييم فيهتم بالشكل والمظهر العام الخاص بالمتعلم أي العمل على تثمينه وجعل قيمته ذات وزن معلوم، ويعتبر بأنه جزء من عملية التقويم (أبيش وعنيشل، 2018).

أوضح الباحث أن الكثير يخلط بين مفهومي التقويم والتقييم في العملية التربوية، وأشار الباحث إلى أن أهداف ووظائف عملية التقويم تكون أعم وأشمل من عملية التقييم بحيث أن المتمعن في العملية التعليمية يمكن أن يلمس الفرق بينها من خلال تطبيق كلتا العمليتين على البرنامج التعليمي الذي يقوم به.

الخطوات الرئيسة للتقويم التربوي:

تمر عملية التقويم التربوي بمجموعة من المراحل الرئيسة والتي تتمثل في تحديد ماهية درجة السمة أو الخاصية المراد العمل على قياسها، وإجراء مقارنة بين كل من ناتج القياس وماهية القيمة المتفق عليها، إضافةً إلى تحديد مستوى القرب أو البعد بين تلك الخاصية وبين القيمة المتفق عليها، وإصدار وتحديد ما هو القرار أو الحكم على تلك الخاصية (صالح والإمام، 2020).

أوضح الباحث أن عملية التقويم التربوي السليمة ترتبط بمجموعة من البنود المتتابعة والمتسلسلة والتي يتوجب على المعلم وعلى القائمين على العملية التعليمية والتربوية متابعة خطواتها بشكل متسلسل للوصول إلى أهداف عملية التقويم التربوي ولتحقيق أقصى مستوى ممكن من النتائج المرغوبة والمتوقعة، فتبدأ عملية التقويم التربوي من رأس الهرم أي بأبسط الأمور ومن ثم تتدرج لتصل إلى أوسع مدى للعملية التربوية ومن ثم تقويمها بصورة متكاملة.

أساليب التقويم وأدوات التقويم التربوي :

تتعدد الوسائل والآليات المستخدمة في التقويم التربوي من موقف إلى موقف آخر ومن فترة زمنية إلى فترة أخرى، وتتنوع وسائل وطرق التقويم التربوي والتي يمكن من خلالها إصدار الأحكام المتعلقة بمحتوى تعليمي أو بشخص ما، ويتم إصداره في إطار ما تشير إليه عملية التقويم التربوي، ولابد من أن يتم تحديد الوسيلة الملائمة والمناسبة لاستخدامها بغرض تحقيق أهداف عملية التقويم التربوي. وذلك لأن عملية التقويم التربوي تشتمل على مجموعة من العمليات والمهمات الفرعية المتكاملة والمترابطة مع بعضها البعض، ونتيجةً للانفجار الثقافي والمعرفي والذي تزامن معه اتساع وزيادة في مهمات وأنشطة التقويم، وبالتالي أصبح لدى المختصين وذوي الشأن مجموعة من الوسائل والآليات التي تكون موزعة على مجموعة من المجالات، حيث إن كل أداة تصلح لمجال أو فئة معينة، ومن أبرز وسائل التقويم التربوي هي الاختبارات التحصيلية بكلا صنفيها ونوعيها "سواء الاختبارات المقالية أو الاختبارات الموضوعية" (خليفة، 2014).

وأشار (Banse et al., 2017; Ambag, 2014) إلى أهم وأبرز وسائل التقويم التربوي والتي تتمثل في:

1. التقويم التربوي القائم على نتائج الأداء: ويعتمد هذا التقويم على المتعلم وتوظيفه لمهاراته التي تعلمها في مختلف المواقف الحياتية التي يمر بها، أو ممارسته لعروض عملية لإيضاح وإبراز ما اكتسبه من مهارات.
2. ملف الإنجاز (portfolio) : وهو عبارة عن تجميع كامل ومنظم لما أنجزه الطالب أو المتعلم من أعمال في فترة زمنية محددة، ويتم مراجعته في إطار مجموعة من المعايير والمحكمات للحكم على مستوى أدائه.

3. التقويم الذاتي : وهو التقويم الذي يقوم به الطالب لنفسه.
4. تقويم الأقران : وهو تبادل الأدوار بين الطالب وزملائه بحيث يعمل كل منهم على تقييم أداء الآخر.
5. تقويم الأداء المبني على الملاحظة : وهو عملية يقوم من خلالها المعلم بتوجيه كافة حواسه المختلفة نحو الطالب ولمراقبته وذلك بغرض الحصول على معلومات تساعد في الحكم عليه، ومن ثم تقويم سلوكه ومهاراته.
6. التقويم القائم على المقابلات : وهي عملية التقويم التي يكون فيها لقاء بين المتعلم ومعلمه ويكون هذا اللقاء محدد في فترة سابقة، بغرض الحصول على معلومات ترتبط بأفكار واتجاهات المتعلم حول أمر أو موضوع معين.
7. التقويم المبني على خرائط المفاهيم : وهو التقويم الذي يقوم على توجيه مجموعة من المصطلحات والمفاهيم المتضمنة للطالب في داخل المحتوى التعليمي ويطلب منه الربط بين هذه المفاهيم.
8. التقويم القائم على الاختبارات : وهو الوسيلة التي يتم استخدامها في قياس ما يقوم الفرد بتحقيقه واكتسابه من معلومات ومهارات وقدرات في نهاية مشواره في الخبرة التعليمية.
9. تقويم دراسة الحالة : وهذا النوع من التقويم يستخدم عند وجود حالة أو عدد بسيط من الحالات بين الطلبة لا يمكنها الفهم بشكل جيد مقارنة بباقي الطلبة، فيعمل المعلم على جمع البيانات الخاصة بهذه الحالة ومن ثم دراستها وتحليلها بصورة منفردة والعمل على تقويمها ومعالجتها بالطريقة الأكثر ملاءمة ومناسبة.
10. المقاييس : إن وجود مخرجات سلوكية متعددة ومتنوعة استوجب العمل على استخدام مقاييس متنوعة ومختلفة.

ويشير الباحث إلى أن التقويم التربوي عبارة عن عملية متكاملة الأبعاد والعناصر وليست مجرد رقم يرصد على أوراق الطالب أو المتعلم بشكل تجريدي، فهو يساعد بالبوح بمجموعة من الأفكار والانطباعات المقصودة، لذلك فقد لجأ المختصون والخبراء في مجال التقويم التربوي والمهتمون بتطوير العملية التعليمية إلى إيجاد طرق وآليات تساعد على تحسين مخرج عملية التقويم التربوي وتساعد ذوي الشأن في اتخاذ القرارات المناسبة.

الاتجاهات الحديثة في التقويم :

لقد بدأت في فترة التسعينات العديد من الاتجاهات الحديثة لعملية تقويم تربوي للطلاب والتي نادت بأن يتجه التقويم التربوي إلى مبدأ قياس الأداء في مجموعة من المواقف الحياتية الحقيقية، حيث أطلق على هذا النوع من التقويم مسمى التقويم الأصيل وفي أحيان أخرى أطلق عليه مسمى التقويم البديل، وآخرين اصطلحوا له مفهوم تقويم الأداء (Alshwaier et al., 2012).

التقويم الأصيل أو البديل :

يعد التقويم الأصيل أو البديل أحد أبرز الاتجاهات الحديثة في التقويم التربوي فهو يعمل على قياس العديد من التطبيقات الخاصة بالمهارات والمعرفة والتي تسهم في رفع مستوى القدرة على حل المشكلات عن طريق تقييم المهام والأعمال المختلفة التي يقوم بتنفيذها الطلبة، والتي تتطلب مستوى أداء ونشاط مرتفع في مجال البحث والتحري عن مختلف المشكلات والقيام بالعديد من التجارب التي تلزم لحل هذه المشكلات، ويشمل التقويم البديل على ملاحظة ماهية نواتج تعلم الطلبة وتستخدم فيه الأدوات المتنوعة مثل صحائف المتعلم، وتقويم المشاركين، وحقائب التقويم، ومن أهم مبادئه (خاجي، 2013):

1. أن يكون التقويم التربوي دائمًا ومستمرًا ويستخدم عدة مرات في العام الدراسي وذلك ليعمل على تزويد المعلم بالتغذية الراجعة اللازمة لتحسين مستوى التعلم لدى الطلبة.
2. أن يكون التقويم التربوي صادقًا يعمل على قياس ما صمم لأجله.
3. أن يكون التقويم التربوي شاملا بحيث يقيس كل من الاتجاهات والمهارات والمفاهيم التي تهدف وتسعى عملية التعلم إلى تنميتها لدى الطالب.
4. استخدام العديد من الطرق المتنوعة والمختلفة في العملية التعليمية والتي تتوافق مع مختلف سمات وخصائص الطلبة وماهية مستويات قدراتهم، فمنهم من يتفوق في مجال ومهارة التعبير بصورة أكبر من مهارة القراءة.

مجالات وجوانب التقويم التربوي :

إن عملية التقويم التربوي من العمليات التي توصف بشموليتها في كافة العمليات التعليمية والتربوية، حيث إن التقويم التربوي يتم نسجه من خيوط العملية التربوية فهو أحد العمليات الجوهرية والحيوية فيها، وذلك يعني أن التقويم التربوي يشـمل كافة الأهداف التربوية على مختلف الجوانب والمسـتويات، ويشـمل المنهج بكافة أبعاده المختلفة، ونتيجةً لاتساع مجالات وجوانب التقويم التربوي وتعددها فيمكن الإشارة إلى أهم الجوانب فيها فيما يلي:

أولاً : التقويم التربوي للمعلم :

ويشمل هذا المجال في التقويم التربوي على المعلم بما يتمثل فيه من كفاياته التعليمية وشخصيته واتجاهاته وأفكاره نحو طلابه ونحو العملية التعليمية، ويوجد مجموعة من الدوافع التي توجه الأنظار للتقويم التربوي للمعلم منها السعي لاستخدام صحائف التقدير الذاتي أو قياس مستوى كفايته بالأثر والنتيجة التي يلمسها المعلم في تلاميذه، أو عن طريق تقويم المعلم من قبل كل من المشرفين والمتعلمين والإداريين التربويين (Alsamaani, 2014).

ثانياً : التقويم التربوي للمتعلم :

يعتبر تقويم المتعلم/ التلميذ من أهم مجالات عملية التقويم التربوي والذي يستهدف من خلال هذا التقويم الحصول على بيانات ومعلومات كمية ووصفية في جانب أو أكثر من جوانب النمو التربوي والذي حققه التلميذ، ويشغل التحصيل الدراسي حيزًا ومكانة خاصة في هذا المجال من التقويم التربوي وذلك لمعرفة مدى النمو التربوي والتعليمي الذي قد حصله بالنسبة لجوانب المعرفة المختلفة والمتنوعة والتي يشتمل عليها البرنامج التعليمي، ويعمل تقويم المتعلم على قياس مستوى قابليته العقلية وقدراته والتي تمت عن طريق تفاعلها ونشاطها الملموس مع البيئة التربوية والثقافية التي تحيط بالمتعلم، وهو ما يطلق عليه مسمى القياس العقلي والذي له قيمة ومقدار في قياس مستوى استعداد وقابلية وكفاية المتعلم لعملية التعلم، ومن الممكن أن يشتمل التقويم التربوي على قياس ميول واتجاهات الطالب المتعلم المختلفة، وبالتالي يمكن الإشارة إلى أن مجال تقويم الطالب المتعلم يشمل كافة جوانب تطوره ونموه (Khora, 2021).

ثالثاً: تقويم الإدارة التربوية

يشـتمل هذا الجانب من التقويم على كافة الأمور المتعلقة بالإدارة التعليمية والتربوية والتي تتمثل في السـلوك الإداري والنمط المتبع من قبل الإداريين وأصحاب الشأن، وكذلك الكشف عن مدى فاعلية الإدارة في الوصول إلى الأهداف التربوية المقصودة والمنشودة والمتعلقة بالمؤسسة التربوية، وتقويم كافة الممارسات والأنشطة الفنية والإدارية للمديرين في محيط المسؤوليات والوظائف وكذلك المهمات التي تحتويها الإدارة، وذلك وفق مجموعة من الأدوات والمعايير التي تم تصميمها لهذا الغرض (Salim et al., 2017).

رابعاً: تقويم المنهج والمواد التعليمية

يشتمل تقويم المواد والمنهج التعليمي على مجموعة من المجالات منها أهدافه وغاياته من حيث علاقتها وارتباطه بالفلسفة المجتمعية والأهداف العامة للتعليم والتربية ومستوى شموليتها، ومدى وضوحها وتوازنها ومراعاتها لكافة التجديدات والتغيرات التي يمكن أن تطرأ على حياة المجتمع وماهية المعرفة المقصودة، ويبحث التقويم التربوي في مدى تطور المحتوى التعليمي والتربوي ومدى توافقه مع التطورات المحيطة به، ويشجع التقويم التربوي على استخدام أساليب تعليمية حديثة تتوافق مع متطلبات العصر الحالي ومع اهتمامات وتطلعات المتعلمين (Fang, 2017).

معايير واستراتيجيات الجودة للتقويم التربوي :

اهتم المختصون في العملية التعليمية وعلى وجه التحديد في التقويم التربوي بالعمل على بناء مجموعة من المعايير التي تتحكم في مستوى الجودة للتقويم التربوي، ومن أهم وأبرز تلك المعايير أنه يجب أن تكون أدوات المستخدمة متوافقة مع مستوى الأداء الذي يبديه الطلبة مع أنماط التعلم المعنية، كما أن إجراءات وتعليمات التقويم التربوي يجب أن تكون سلسة وواضحة وصريحة بالنسبة للطلبة، إضافةً إلى توفر الصدق والثبات والموضوعية في وسائل التقويم التربوي المستخدمة، وملاءمة المعايير والتقديرات للمحتوى المقرر واتسامها بالثبات مع مرور فترات زمنية مختلفة، والتزام الموضوعية والعدالة عند رصد العلامات، ووضوح في الإجراءات المتبعة عند تدني التقويم عن الحد المطلوب (David, 2013).

ووضع الخروصي وآخرون (2021) مجموعة من المعايير والمقترحات التي تساعد في ضبط مستوى جودة التقويم التربوي، والتي يمكن الإشارة إليها في الجدول التالي:

جدول 1: مجموعة من المعايير والمقترحات التي تساعد في ضبط مستوى جودة التقويم التربوي

المعيار	الدلالة
الوضوح	ويشمل وضوح الغرض ممن التقويم، ووضوح مخرجات المتعلم المراد قياسها، ووضوح مستويات الأداء، ووضوح محتوى مفردات طرق التقويم.
الارتباطية	أن يرتبط محتوى عملية التقويم بمحتوى عملية التدريس، وأن تتلاءم طريقة التقويم المستخدمة مع الغرض من التقويم
الشفافية	وتشمل جعل الطالب شريكا في عملية التقويم بتبليغه عن الغرض من التقويم، ومستوى الأداء المتوقع، ومحتوى التقويم
الموضوعية	وترتبط بعدم تأثر نتائج التقويم بالعوامل الشخصية للقائم بعملية التقويم بحيث تعطي عملية التقويم نفس النتائج باختلاف القائم بالتقويم
العدالة	أن تتيح عملية التقويم فرص متكافئة للطلبة عند أداء التقويم بحيث تعكس نتائج التقويم مقدار ما قام به الطالب من عمل وجهد لتعلم المادة الدراسية، وأن تخلو عملية التقويم من أي شكل من أشكال التحيز
التنوع	ويتضمن استخدام طرق متعددة للتقويم تراعي الفروق الفردية بين الطلبة في المتعلم، وتمنح الطلبة الزمن المناسب لإنجاز المهام المطلوبة وفق قدراتهم المختلفة.
الشمولية	أن تشمل عملية التقويم جوانب مختلفة من شخصية الطالب (المعرفي، والمهاري، والنفسي...الخ،) ومستويات مختلفة من الأهداف التعليمية
الاستمرارية	أن يتم تطبيق التقويم بصفة مستمرة من بداية عملية التدريس وحتى نهايتها.
المصداقية	وتشير إلى صحة تفسيرات نتائج التقويم والقرارات المستمدة منها
الاعتمادية	تشير إلى الاتساق في نتائج التقويم باختلاف ظروف المواقف التقويمية
التواصل	أن يتبع عملية التقويم تقديم معلومات توجيهية للطلبة تبين نواحي القوة في الأداء ونواحي الضعف وسبل تحسينها.
التوازن	أن يكون لكل طريقة تقويم الوزن المناسب من التقدير الكلي للتحصيل الدراسي للطالب في المادة
الإجرائية	أن يتم إجراء عملية التقويم بطريقة سهلة بحيث تعطي أكبر قدر ممكن من المعلومات عن أداء الطالب بدقة أكبر وبأقل جهد
الأثر الإيجابي	أن يترتب على عملية التقويم آثار إيجابية على المعلم والطالب
الواقعية	أن تتضمن عملية التقويم مهام تقييم تكشف قدرة الطالب على تطبيق مما تعلمه في الغرفة الصفية على مواقف مرتبطة بالحياة الواقعية

أوضح الباحث أهمية وضرورة قيام المختصين في التقويم والقياس بتجويد كافة الممارسات التي يقوم بها المعلمون لعملية التقويم التربوي وذلك لما للأمر من أهمية عند إبلاغ الطلبة بماهية أساليب التصحيح والتدقيق وماهية المعايير التي يتم من خلالها الحكم على مستوى الجودة في عملية التقويم التربوي.

حقائق عن التقويم التربوي :

أشار المتخصصون والخبراء في التقويم التربوي إلى وجود مجموعة من الحقائق التي لا يمكن الطعن فيها فيما يتعلق بالتقويم التربوي، وتتمثل أبرز تلك الحقائق في(Al-Adawi, & Al-Balushi, 2016):

- يعتبر التقويم التربوي أحد أهم وأبرز العمليات التي تتم في برامج التعليم المختلفة، والتي يتوجب على المسؤولين وأصحاب القرار منحه الحيز المناسب والملائم من الاهتمام.

- إن من أبرز التوجهات والصور الحديثة للتقويم التربوي: التقويم الأصيل أو البديل أو التقويم القائم بناءً على الأداء.
- إن تطوير ورفع مستوى المنظومة التعليمية يعتمد بصورة كبيرة وأساسية على عمليات التقويم التربوي الذي يسهم في الوقوف عند مواطن الضعف فيقلصها ويمحيها وعند مواطن القوة فيرفع منها ويزيدها.
- التقويم التربوي يعتمد بصورة كبيرة على مقياس أو محك تعليمي محدد ولكن لا يشترط ارتكازه على هدف وغاية مسبقة التحديد.
- التقويم التربوي عبارة عن عملية منظمة ومترابطة يمكن من خلالها تجميع البيانات والمعلومات التي تسهم في تكوين صورة متكاملة عن الأنشطة التعليمية وتحدد توجهاتها وأساليب تفعيلها.

الخاتمة :

يعتبر التقويم التربوي النافذة التي تستخدم لقياس وفحص مستوى الأداء للعملية التعليمية، وذلك وفق مفهومها الاصطلاحي المتعارف عليه في مختلف الأنظمة التعليمية، فالتقويم التربوي يشكل عملية منظمة وممنهجة تهدف إلى جمع بيانات تتعلق وترتبط بالطالب بالاستعانة بعدة طرق متنوعة ومن ثم تفسيرها وتحليلها لمعرفة مدى تحقيق الطلبة للأهداف المقصودة، ويشكل التقويم التربوي مقياسًا حقيقيًا وواقعيًا قابلًا لعملية التطبيق على مختلف الأنشطة التعليمية والتربوية المتنوعة، بالإضافة إلى أنه يسعى إلى تعديل جميع العناصر التي تتضمنها العملية التعليمية، والتي تبدأ من حيث المنهاج والمحتوى التعليمي وتنتهي بنتائج التحصيل الدراسي.

المحور الثاني: الاختبارات ومعايير جودة الاختبارات

إن وجود التطورات المتتابعة والمتسارعة في زمننا الحالي في اختصاص تكنولوجيا المعلومات وتقنياتها تفرض علينا إنشاء وبناء استراتيجيات وأنظمة تعليمية وتربوية مرنة وقابلة لعملية التكيف مع مختلف المتطلبات العصرية، ويتم بناء مثل هذه الأنظمة والاستراتيجيات من خلال مهارات ابتكارية وإبداعية تسهم في تحقيق أفضل آلية للوصول المميز (التوبية وآخرون، 2021).

ومن أبرز اختبارات اللغة الإنجليزية الاختبار الخاص بتحديد المستوى للغة الإنجليزية والذي يتم تطبيقه على طلبة جامعة مؤتة، والذي يهدف إلى قياس مستوى القدرات الأساسية للطلبة وذلك لتحديد مدى حاجتهم لمساقات ومواد استدراكية تلزمهم لتطوير بعض المهارات الضرورية واللازمة لدراستهم الجامعية (أبوعلا والمطارنة، 2018).

وأشار Belfield et al (2012) إلى أن اختبار تحديد المستوى للغة يعتبر من أبرز الاختبارات والإجراءات الحاسمة التي يجب اعتمادها بصورة كبيرة وشاملة على مختلف الأصعدة، وذلك لأن اختبار تحديد المستوى يعتبر من الاختبارات التي تساعد الطلبة الجامعيين وتخفف عنهم بعض المصاعب التي يواجهونها بسبب قصور في اللغة.

لذلك فإن اختبار المستوى يشكل تنبيهًا وتوجيهًا للطالب لدفعه لاكتساب المهارات التي تلزمهم في مختلف المقررات الدراسية التي تتعلق بتخصصاتهم الجامعية، الأمر الذي دفع ببعض الجامعات كجامعة دلتا للعلوم والتكنولوجيا إلى إنشاء نموذج خاص بها لاختبار تحديد المستوى في اللغة الإنجليزية، وذلك نتيجةً لاختلاف المتطلبات اللغوية من مختلف التخصصات وتباينها مع بعضها البعض، وكذلك نتيجةً لارتفاع التكلفة التي تتطلبها اختبارات مستويات اللغة الإنجليزية الدولية والعالمية (العدل، 2018).

الاختبارات اللغوية :

تعتبر الاختبارات اللغوية من أهم وسائل وطرق القياس التي تعمل على تقديمها هيئة تقويم التعليم والتدريب وذلك وفق منهج موضوعي دون وجود تحيز فيه، لذلك تعمل هيئة التقويم للتعليم والتدريب على تقديم الاختبارات المتعلقة بكفايات لغوية للمتقدمين بكلتا اللغتين اللغة الإنجليزية واللغة العربية سواء كان

المتقدمون من الناطقين أو غير الناطقين بها، ومن أبرز اختباراتها اللغوية اختبار قياس الكفايات اللغوية مثل اختبار STEP أو الاختبارات الخاصة بتحديد المستوى للمتقدم في اللغة الإنجليزية مثل اختبار EPT، ويكون الهدف الرئيسي والأساسي من هذه الاختبارات والمقاييس اللغوية هو العمل على خدمة مجموعة من الجهات الأكاديمية المتنوعة والتي تشمل معاهد وجامعات ومدارس تقوم على استقطاب مهنيين أو مدرسين ذوي مؤهلات مرتفعة (Durrant, 2014).

يتم اختبار كفايات اللغة الإنجليزية STEP، ويشير اختصار STEP إلى مصطلح Standardized Test of English Proficiency والذي يعني اختبار كفايات اللغة الإنجليزية، وهو أحد أبرز الاختبارات الوطنية التي يتم تحضيرها لاختبار مستوى الكفايات لدى المتعلمين والمتقدمين في اللغة الإنجليزية، حيث يعمل هذا الاختبار على تغطية ثلاثة مهارات أساسية ورئيسية تتمثل في استخدام اللغة، والتي يمكن أن يتم قياسها بثبات وموضوعية والحصول على نتيجة ذات بعد جيد من الثبات والصدق، فيغطي الاختبار بما يحتويه من أسئلة من المفردات وتراكيبها ودلالاتها وفهم المقروء والنحو وفهم المسموع وكذلك التحليل الكتابي (Sims, 2015).

والتحضير لاختبار كفايات اللغة الإنجليزية لا يعتمد بصورة واضحة وأساسية على ماهية المعلومات التي يحصل عليها المتعلم من محتوى المقررات التعليمية أثناء دراسته لها، أو من المراجع أو الكتب المحددة والمخصصة لهذا الاختبار، فهذا الأمر يعني بأن المتقدم لهذا الاختبار لا يعتمد بصورة أساسية على التحضير والاستعداد المسبق بل يعتمد بصورة كبيرة على مدى حصيلة المتقدم في اللغة الإنجليزية، وذلك من خلال مختلف التجارب الحياتية التي تتراكم لديه نتيجة استخدام اللغة أو الدراسة، ولكن يعتبر من الأمور المهمة أن يطلع المتقدم على ماهية طريقة عرض الأسئلة وآلية وكذلك كيفية الإجابة عليها، مما يساعده على تقديم أفضل مستوى لديه (David, 2013).

يعتبر اختبار كفايات اللغة الإنجليزية من الاختبارات التي تعقد بصورة محوسبة، ويكون مفتوح بشكل دائم أي على مدار العام، ويمكن للمتقدم التسجيل في أي وقت يرغبه، ويتم تقديم الاختبار بصورة إلكترونية من خلال مراسلة الموقع الإلكتروني الرسمي الخاص بالمركز (Shelly, 2015)، يعتبر اختبار كفايات اللغة الإنجليزية صالحًا لفترة تصل إلى ثلاث سنوات من تاريخ التقدم للاختبار.

يتدافع المتقدمون إلى تقديم اختبار كفايات اللغة الإنجليزية وذلك لوجود مجموعة من الدوافع التي تختلف من شخص إلى آخر، ومن أهم وأبرز تلك الدوافع التقديم على أحد أقسام اللغة الإنجليزية، والرغبة في الالتحاق بإحدى برامج تعلم اللغة الإنجليزية وذلك وفق ما يتلاءم مع مستوى الطالب، والحصول على إعفاء من دراسة مقررات خاصة باللغة الإنجليزية وفق شروط برامج دراسية محددة، وإنهاء أحد متطلبات التقدم على برنامج الدراسات العليا، والالتحاق بإحدى الكليات العسكرية، إضافةً إلى كون المتقدم أحد معلمي اللغة الإنجليزية ويرغب في الحصول على شهادة لهذا الاختبار، والرغبة في الالتحاق ببعض الوظائف التي يكون من متطلباتها امتلاك مستوى جيد من اللغة الإنجليزية (Sidky, 2019).

ويتشكل اختبار كفايات اللغة الإنجليزية من مجموعة من العناصر والتي تشمل على(Abdel-Gawad, 2019):

- تحليل الكتابة، وتكون نسبته (10%) من إجمالي الاختبار.
- استيعاب المقروء، وتكون نسبته (40%) من إجمالي الاختبار.
- استيعاب المسموع، وتكون نسبته (20%) من إجمالي الاختبار.
- بناء الجمل والقواعد، وتكون نسبته (10%) من إجمالي الاختبار.

اختبار تحديد المستوى في اللغة **EPT** :

يأتي اختصار EPT من مصطلح English Placement Test والذي يعني اختبار تحديد المستوى في اللغة، حيث يسعى اختبار تحديد المستوى في اللغة EPT إلى تحديد ماهية المستويات المتعلقة بالكفاية اللغوية في اللغة الإنجليزية لدى الأفراد الراغبين بالالتحاق في برامج تدريبية أو علمية تكون لغة التعليم فيها هي اللغة الإنجليزية، وقد تم العمل من قبل الخبراء والمختصين على تطوير الاختبار وفق مجموعة من الممارسات التقليدية والمتعارف عليها بصورة دولية في صناعة مختلف الاختبارات الموضوعية (Salehi et al., 2019).

يعمل اختبار تحديد المستوى في اللغة EPT على قياس المستوى لدارسي اللغة الإنجليزية، وذلك عن طريق الإجابة على 80 بندا وسؤالا، وتكون تلك الأسئلة بصيغة اختيار من متعدد (Lei & Hu, 2014).

يعمل اختبار تحديد المستوى في اللغة EPT على قياس ثلاث مهارات أساسية لدى المتقدمين للاختبار، بحيث تكون تلك المهارات وعدد الأسئلة المخصصة لها كالتالي:

- فهم المقروء ويحتوي على 20 سؤالا.
- التراكيب اللغوية، وتتواجد في 40 سؤالا.
- التحليل الكتابي وتكون عبارة عن 20 سؤالا.

وهذا يعني أن اختبار تحديد المستوى في اللغة EPT يعمل على تغطية العناصر الأساسية للغة بما فيها المفردات اللغوية والتراكيب ودلالاتها وأنظمة الكتابة، وذلك عن طريق مجموعة شاملة من الأسئلة التي تغطي هذه المكونات والمهارات (Yan, 2014).

يتقدم لاختبار تحديد المستوى في اللغة EPT مجموعة من الطلبة في الكليات والمعاهد والجامعات التي تدرس موادها باللغة الإنجليزية، وذلك بغرض العمل على تأهيل المتقدمين لمراحل ومستويات دراسية وأعمال ووظائف تتطلب مستوى إجادة متميز للغة الإنجليزية (Zainal Abidin & Jamil, 2015).

التحديات والصعوبات في الاختبارات الدولية والعالمية :

يوجد العديد من التحديات والصعوبات التي تواجه الاختبارات الدولية والعالمية ومنها (Guo, 2012):

- ضعف في مستوى الاستيعاب والتجاوب من قبل أولياء الأمور مع المؤسسات التعليمية وذلك نتيجة عدم إدراكهم لأهمية العلاقة التكاملية بين كل من البيت والمؤسسة التعليمية.
- تأخر المعلمين أو غيابهم عن المؤسسة التعليمية.
- وجود بعض التدخلات من قبل أولياء الأمور في شؤون المؤسسة التعليمية في بعض الأمور التي لا تعنيهم وليست من ضمن اهتماماتهم.
- العوامل والجوانب المادية والاقتصادية: نتيجة لارتفاع سعر الاختبارات الدولية والعالمية، فإن وجود ظروف مادية واقتصادية صعبة يضعف من مستوى تقدم الطلبة لتلك الاختبارات.
- قلة وضعف في مستوى اهتمام المعلمين للوسائل والاستراتيجيات التعليمية الحديثة.
- عدم وفرة عدد كافٍ من المتعلمين المتخصصين والمهرة في اللغة بشكل يتوافق مع طبيعة تلك الاختبارات.

ويشـير الباحث مما سـبق إلى وجود مجموعة من الصـعوبات والتحديات التي يمكن أن تعرقل طريق الطلبة والمتقدمين للاختبارات الدولية والعالمية، وتلك التحديات ترتبط بكافة العوامل المختلفة والعناصـــر المكونة للعملية التعليمية، فمنها ما يتعلق بالطالب نفســـه وما يمتلكه من مهارات لغوية، ومنها ما يرتبط بالبيئة التعليمية والمحتوى التعليمي، بالإضافة إلى العديد من التحديات التي ترتبط بالمعلمين والأشخاص الذين يدرسون الطالب المحتوى التعليمي.

الفرق بين اختبارات **EPT** واختبار **STEP** :

يشير المختصـون في إعداد وتقييم الاختبارات الوطنية والدولية في اللغة الإنجليزية وخصـوصـا في اختباري EPT وSTEP إلى وجود اختلافات متعددة ومتنوعة بين كلا الاختبارين والتي يمكن الإشارة إليها في عدد من البنود المتنوعة والمختلفة، منها الفروق بين علامات وتوزيع نقاط التقويم لكل اختبار وماهية المهارات التي يعمل على قياسـها الاختبار، والفئات المسـتهدفة وصـلاحية الفترة الزمنية للاختبار، وغيرها من الأمور الأخرى، بالإضـافة إلى وجود اختلاف جوهري وأسـاسـي (Setiawan, 2021).

وهو أن كل اختبار فيهما يعمل على قياس المسـتوى الذي يحققه المتقدم بغرض تحقيق الهدف المنشـود من كل اختبار فيهما، حيث يعمل اختبار EPT على قياس طلبة المرحلة الثانوية أي المقبلين على المرحلة الجامعية، وخصـوصـا الذين يتمتعون بلغة إنجليزية ذات مسـتوى متوسـط وذلك لاجتياز بعض المتطلبات والمواد التحضـــيرية الخاصـــة باللغة الإنجليزية، في حين أن اختبار STEP هو اختبار خاص بالكفاية في اللغة الإنجليزية ويعتبر قريبا جدا من اختبارات TOFEL وILETS والذي يعمل على قياس مسـتوى الخبرة والمعرفة في اللغة الإنجليزية، وذلك عن طريق عقد اختبار موضـوعي إلكتروني "محوسـب" وموحد لكافة الجامعات السعودية (Lin & Chen, 2020).

وفيما يلي جدول يوضح الفروق بين اختباري EPT وSTEP

جدول2: الفروق بين اختباري EPT وSTEP

وجه المقارنة	اختبار STEP	اختبار EPT
المقررات التي يتم إجراء اختبار فيها	✓ فهم المحتوى المقروء ويمثل 40%. ✓ التحليل للمحتوى الكتابي وتكون نسبته 10%. ✓ فهم المحتوى المسموع ويكون بنسبة 20%. ✓ التراكيب للمحتوى النحوي ويمثل 30%.	✓ التراكيب للمحتوى اللغوي ويمثل 40 بندا وسؤالا. ✓ تحليل المحتوى الكتابي ويكون ممثل بـ 20 سؤالا. ✓ فهم المحتوى المقروء ويشكل ما نسبته 20 بندا وسؤالا.
معلومة عامة عن الاختبار	هو أحد أبرز الاختبارات الدولية الذي يتم عقده من المركز الوطني للقياس والتقويم، ويهتم هذا الاختبار بقياس مستوى الكفايات للغة الإنجليزية.	هو الاختبار الذي يتم من خلاله تحديد مستوى الطالب المتقدم في اللغة الإنجليزية ولكافة الراغبين في الالتحاق ببرامج علمية أو تدريبية تكون اللغة الإنجليزية هي لغة التعليم المعتمدة فيها.
عدد الفرص للمتقدم بالاختبار	عشر مرات للمحاولات وتكون في فترة زمنية تنحصر بثلاث سنوات.	خمس مرات للمحاولة خلال مدة أقصاها ثلاث سنوات.
الفئة المستهدفة في الاختبار	المرحلة الثانوية وكذلك المتخرجين من كلا المرحلتين الجامعية والثانوية.	الكليات والجامعات وكذلك المعاهد التي تقوم بالتدريس للمقررات الخاصة بها باللغة الإنجليزية لتأهيل المتقدمين لمراحل تعليمية ودراسية متقدمة أو بغرض وظائف ومهام تتطلب إجادة مستوى مرتفع في اللغة الإنجليزية.

وجه المقارنة	اختبار STEP	اختبار EPT
رسوم الاختبار	150 ريال سعودي	100 ريال سعودي
صلاحية الاختبار (الفترة الزمنية)	ثلاث سنوات منذ التاريخ الذي تم فيه تقديم الاختبار	ليس له فترة صلاحية معينة ومحددة.
الفترة الزمنية المستغرقة في ظهور النتيجة	ثلاثة أيام	ثلاثة أيام

معايير جودة اختبارات اللغة الإنجليزية العالمية :

يعتمد المختصون وذوو الشأن في بناء الاختبارات العالمية الخاصة باللغة الإنجليزية على توفير أعلى مستوى ممكن من الجودة الشاملة لكافة المعايير والجوانب المتنوعة والمختلفة، ومن أهم وأبرز معايير ضمان مستوى جودة عالية في اختبارات اللغة الإنجليزية الدولية ما يأتي:

- أن تكون أهداف الاختبارات الخاصة والعامة محددة وواضحة وصريحة.
- جودة المصطلحات والمفاهيم التي يتم بناء الاختبارات عليها.
- شمول الاختبارات على المكونات والعناصر العامة المتنوعة والمختلفة.
- بناء الاختبارات بالارتكاز على مجموعة من الأسس النظرية والفنية لصياغة محتوى وتكوين أسئلة الاختبار.
- تكون الاختبارات شاملة لكافة الجوانب والمجالات وكذلك المهارات التي تتمتع بها اللغة الإنجليزية.
- يتم منح كل سؤال رقم خاص به وذلك ليشكل هوية خاصة بالسؤال وهي تعتبر دليل للتعريف والتشفير للسؤال (Kling & Stæhr, 2012).

لقد بين الباحث وجود مجموعة من الاعتبارات والأساسيات التي يعتمد عليها المختصون في صياغة وإعداد الاختبارات الدولية والوطنية، وقد تم وضع تلك السياسيات والبنود وفق مجموعة من الدراسات الإجرائية التي قام بها الباحثون للوصول إلى أفضل صورة وآلية يمكن من خلالها كتابة الاختبارات الوطنية والاختبارات الدولية، ونتيجةً لضرورة وأهمية الاختبارات الدولية تم التأكيد على ضرورة إعداد مثل تلك الاختبارات من قبل أخصائيين في المجال المراد صياغة الاختبارات فيه، ويشير الباحث إلى أنه يتوجب توافر مستوى عالٍ من الخبرة لديهم وخصوصا الخبرة الكافية في طرق إعداد وصياغة الاختبارات الدولية، وذلك لأن مثل هذه الاختبارات يجب منع تعرضها لأي خطأ أيا كان بسيطا أو معقدا.

وقد قام الباحث بالرجوع إلى مجموعة من المراجع ذات الصلة لتحديد المعايير الخاصة بتقويم الاختبارات الدولية ومن هذه المراجع (Kling & Stæhr, 2012; Alshwaier et al., 2012)، وذلك من أجل تحديد أهم المعايير التي يتم الاعتماد عليها في تقويم الاختبارات اللغوية.

آليات تقويم الاختبارات الكتابية الدولية :

إن الاختبارات الدولية تمثل أهمية كبيرة لدى المختصين في الشؤون التعليمية وخاصةً الدولية منها، والذين يسعون للوصول إلى أفضل صورة وشكل للاختبارات الدولية نظرا للمكانة التي تحتلها تلك الاختبارات، لذلك وضع المختصون مجموعة من المعايير والآليات التي يمكن من خلالها تقويم الاختبارات الكتابية الدولية، ومن أهم وأبرز آليات تقويم الاختبارات الكتابية الدولية ما يأتي(Alshwaier et al., 2012):

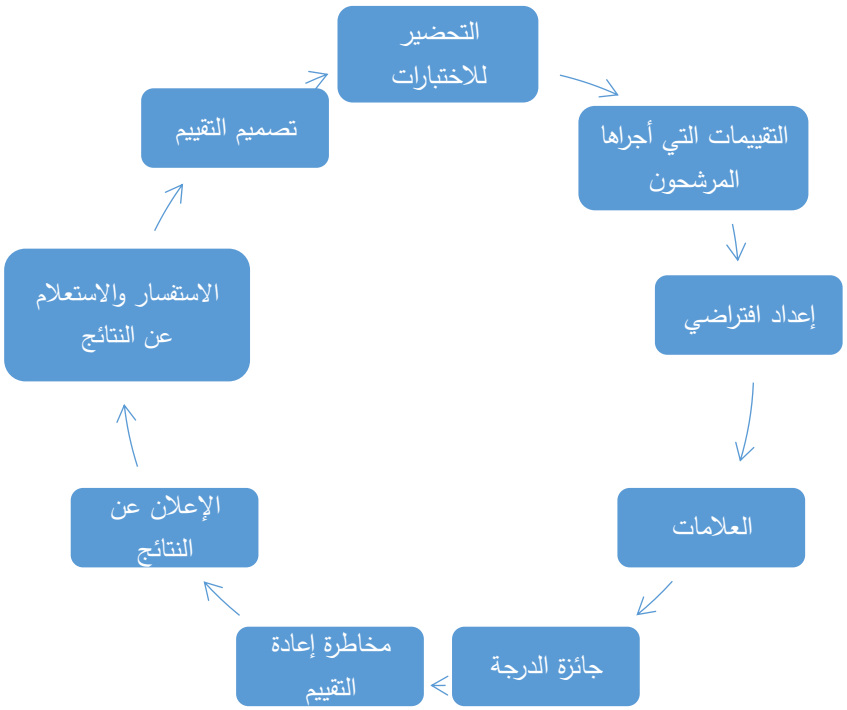
- يتم العمل على تحديد لجان خاصة للرقابة على الاختبارات الكتابية الدولية.

- نتيجة التطور التقني والتكنولوجي، يتم الآن إجراء بعض الاختبارات بصورة محوسبة أي إلكترونية، ومع ذلك تم برمجة العديد من الآليات التي يمكنها ضبط تقديم الاختبارات الدولية كأن يتم تقديم تلك الاختبارات على أجهزة محوسبة تابعة لمؤسسة أو مركز يمتلك حقوق في فتح أبوابه للطلبة أو المتقدمين لإجراء الاختبارات الدولية وفق لجان محكمة ومحددة سابقا.
- تتم عملية التقويم بصورة آلية لذا فقد صممت برمجيات لهذا الغرض تمنح الطالب أو المتقدم التقييم الملائم والمناسب له.
- يتم تقويم الاختبارات الدولية وفق مجموعة من المعايير التي تتجلى في صور مختلفة من الشفافية والنزاهة والموضوعية.
- تقوم لجان مخصصة على تقويم الاختبارات الدولية وإعطاء النتائج بمنتهى الصدق والشفافية.

أوضح الباحث أن عملية التقويم للاختبارات الكتابية الدولية تتم وفق مجموعة من الشروط والمعايير التي تضبط عملية التقدم للاختبارات وإصدار النتائج، بما يسهم في منح المركز والمؤسسة الثقة من قبل المتقدمين، ويزرع في نفوسهم الثقة بذاتهم وبما يمكنهم أن يحرزوه من خلال التقدم لمثل هذه الاختبارات الدولية.

واقع الاختبارات وتقويمها في هيئة تقويم التعليم والتدريب اللغوية :

تتميز الاختبارات الكتابية الدولية وعملية تقويمها بموضوعية عالية وذلك لأن عملية تقويم الاختبارات تعتمد بصورة أساسية على تقويم عادل لكافة المرشحين والمتقدمين لاجتياز الاختبارات الكتابية الدولية حيث تقوم عملية التقويم بقياس الأهداف التي صنعت وصممت من أجلها، وتخضع الأسئلة المكونة للاختبارات لمجموعة من المحكمين الذين يصدرون القرار بقبولها دون تعديل أو قبولها بعد إجراء التعديلات اللازمة عليها، أو عدم قبولها لعدم ملاءمتها لمعايير الاختبارات الدولية، وتتم عملية التقويم لتلك الاختبارات من خلال فريق متخصص يتم اختياره وفق معايير دولية عالية، لضمان مستوى جيد من النزاهة والشفافية، كما أن الاختبار الواحد يتم إخضاعه لعدد من المقيمين وذلك لمنح الطالب المتقدم حقه كاملا دون وجود أية عوامل خارجية وبالتالي يحصل الطالب أو المتقدم على التقويم الملائم والمناسب له ولما قدمه في تلك الاختبارات الدولية (هيئة تقويم التعليم والتدريب، 2022).



شكل 1: آليات ومراحل تقويم وتصميم الاختبارات الدولية

أوضح الباحث أن عملية التقويم للاختبارات في هيئة تقويم التعليم والتدريب اللغوية تخضع لمعايير تسهم في رفع مستوى الجودة لتلك الاختبارات، وتزيد من مستوى الإقبال والثقة من قبل المتقدمين وذلك في وجود لجنة متخصصة ونزيهة تعمل على تقييم تلك الاختبارات.

الخاتمة :

تعتبر الاختبارات الدولية وخاصة الاختبارات المتعلقة باللغة الإنجليزية من أهم الاختبارات التي تأخذ حيزًا من اهتمام وانشغال الكثير من الأفراد بمختلف مستوياتهم المجتمعية وصفاتهم المختلفة، فمنهم المعلم والطالب والجامعي والخريج، ولكل منهم غرض وغاية من تقديم الاختبارات الدولية والمتعلقة باللغة الإنجليزية، وذلك لأن اختبارات اللغة الإنجليزية الدولية بالرغم من قوتها وصعوبتها وتكاليفها إلا أنها تساعد المتقدمين لها في الحصول على العديد من الميزات التي تسهم في رفع مستواهم التعليمي أو الثقافي أو التحصيلي اللغوي، كما أن هذه الاختبارات تخضع للقوانين الدولية والمعايير العالمية من مستويات الجودة والتي تحقق أفضل عملية تقييمية ممكنة.

وتوجهت العديد من المؤسسات والمنظمات التعليمية إلى وضع شروط تتعلق بضرورة التقديم لتلك الاختبارات قبل الالتحاق بأقسامها وكلياتها ومستوياتها التعليمية المختلفة، والتي قد وجدها البعض في الظاهر بأنها عملية تعقيد على الطلبة والملتحقين ولكنها في باطنها تحمل الكثير من الغايات المقصودة والمرغوبة والتي تساعدهم في تحقيق أعلى مستويات من التقدم في المجال المعرفي والتعليمي، حيث إن حصول المتقدم وإحرازه لتلك الاختبارات بمثابة صعود درجات أخرى على السلم المعرفي والثقافي والتعليمي والاجتماعي، لذلك يتوجب على المتقدمين منح تلك الاختبارات حيزًا من الاهتمام والتركيز المناسبين.

ثانيًا: الدراسات السابقة

تناول الباحث في هذا المحور استعراض الدراسات السابقة المرتبطة بموضوع البحث من خلال الرجوع إلى المناهل التعليمية والأدبيات التربوية حيث قام بترتيب الدراسات من الأحدث إلى الأقدم وهذه نماذج من تلك الدراسات:

المحور الأول: الدراسات المتعلقة بتقويم الاختبارات :

دراسة (وحشه، 2022) هدفت هذه الدراسة إلى التعرف على أثر استراتيجية العصف الذهني في رفع مستوى أداء الطلبة في الاختبارات الدولية من وجهة نظر مديري المدارس الحكومية في محافظة عجلون، استخدمت الباحثة المنهج الوصفي الكمي التحليلي، وتم إعداد استبانة مكونة من (23) عبارة تم تقسيمها إلى ثلاثة محاور (القراءة، الرياضيات، العلوم)، تم توزيعها على عينة مكونة من (95) مديرا ومديرة، وقد أظهرت نتائج الدراسة أن دور استراتيجية العصف الذهني في رفع أداء الطلبة في الاختبارات الدولية جاء مرتفعا وبمتوسط حسابي (4.01) وعدم وجود اختلاف في مستوى وجهات نظر مديري المدارس الحكومية في محافظة عجلون في أثر استراتيجية العصف الذهني في زيادة تحصيل الطلبة في الاختبارات الدولية يعزى لمتغيري (الجنس والمؤهل العلمي).

دراسة أبو جامع (2021) هدفت هذه الدراسة إلى تقويم اختبارات الأحياء للصف الثاني عشر في ضوء عمليات العلم، وطبقا لنوعية الدراسة اتبع الباحث المنهج الوصفي التحليلي، وقام الباحث بتصميم أداة التحليل الخاصة بعمليات التعلم، وتكونت عينة الدراسة من اختبارات الثانوية العامة الموضوعية والمقالية لمبحث الأحياء خلال السنوات الثلاث 2018-2019-2020 في قطاع غزة، وتوصلت الدراسة بعد التطبيق إلى أن مهارة الاستدلال حصلت على أعلى نسبة من المهارات الأساسية، وحصلت مهارة ضبط المتغيرات على أقل نسبة، وبذلك أوصى الباحث بضرورة الانتباه في عملية وضع اختبارات الأحياء بحيث تضمن جميع مهارات التعلم.

دراسة الغزال (2020) هدفت هذه الدراسة إلى تقويم الاختبارات التحصيلية بمرحلة الدراسات العليا بقسم اللغة العربية وآدابها بكلية الآداب جامعة مصراته في ضوء معايير الجودة الشاملة، وطبقا لنوعية الدراسة اتبع الباحث المنهج الوصفي التحليلي، وقام الباحث باعتماد قائمة المعايير المعدة من قبل الغزال (2012) كأداة دراسية ملائمة لطبيعة الدراسة، وتكونت عينة الدراسة من جميع الاختبارات التحصيلية لمادة اللغة العربية خلال السنوات 2012-2018، والبالغ عددها (49) اختبارًا، وبعد استخراج قيم الثبات تبين للباحث أن معظم الاختبارات كانت مكتوبة بلغة واضحة وقاست قدرة الطلبة على ممارسة عمليات عقلية متنوعة، وأوصي الباحث بضرورة تضمين معايير الجودة الشاملة ضمن دورات التدريبية للمعلمين.

دراسة الجار (2020) هدفت هذه الدراسة إلى تقويم واقع الاختبارات الفصلية في بعض كليات جامعة شقراء كما تراها الطالبات في ضوء بعض المتغيرات، وطبقا لنوعية الدراسة اتبع الباحث المنهج الوصفي التحليلي، وتكونت عينة الدراسة طبقا لنوعية البحث على (241) طالبة تم تطبيق الاستبانة كأداة

بحثية عليهن، وخلصت الدراسة إلى أن المتوسط العام لتقدير الطالبات لممارسات أعضاء الهيئات التدريسية جاء بدرجة متوسطة، وعليه توصي الباحثة بضرورة الانتباه إلى تطوير أعضاء الهيئات التدريسية في ضوء المعطيات الحديثة والمختلفة.

دراسة البلوي (2019) هدفت هذه الدراسة إلى تقويم الاختبارات المدرسية التي تعدها معلمات العلوم في المرحلة الأساسية بمدينة تبوك من حيث تخطيطها وإخراجها وتصحيحها، وطبقا لنوعية الدراسة اتبع الباحث المنهج الوصفي التحليلي، وتكونت عينة الدراسة طبقا للدراسة من (100) معلمة من معلمات مادة العلوم، تم تصميم أداة دراسية ملائمة وهي الاستبيان لقياس مدى استيفاء الاختبارات لمعايير بناء فقرات الاختبار، وتوصلت الدراسة إلى أن درجة التزام المعلمات بمهارات التخطيط وتصحيح الاختبارات التحصيلية جاء بدرجة متوسطة، وخرج البحث بمجموعة من التوصيات من أهمها ضرورة عقد دورات تدريبية في بناء الاختبارات التحصيلية للمعلمات وتشجيعهن على الاستفادة من خبرات بعضهن البعض وتبادل الخبرات فيما بينهن.

دراسة سويلم (2018) تمحورت الدراسة حول تقويم اختبارات طلاب الشعب العلمية بكلية التربية بجامعة المنيا في ضوء معايير جودة الاختبارات، وطبقا لنوعية الدراسة اتبع الباحث المنهج الوصفي التحليلي، وقام الباحث بتصميم قائمة معايير جودة الاختبارات باعتبارها الأداة المناسبة، وطبقت الدراسة على عينة مكونة من الأوراق الاختيارية التي أعدها أعضاء هيئة التدريس وعددها (60) ورقة اختبارية، وتبين للباحث أن بعض الاختبارات التي توضح من قبل المعلمين في الشعب العلمية لا تتناسب مع المهارات الطلابية وتكون غير قادرة على قياس مهارات التفكير لدي الطلبة، وأوصي الباحث بضرورة إعداد المعلمين وإمدادهم بدورات تدريبية من شأنها المساهمة في إكساب المعلمين معايير جودة الاختبارات.

دراسة (أبيش وعنيشل، 2018) هدفت هذه الدراسة إلى تقويم أسئلة البناء اللغوي في اختبارات اللغة العربية والتي تم تطبيقها على طلبة مرحلة التعليم المتوسط السنة الثانية، وتم استخدام المنهج الوصفي التحليلي لإجراء الدراسة، وقامت الباحثتان بجمع مجموعة من النماذج والمواضيع المتعلقة بالاختبارات الخاصة بمادة اللغة العربية من مختلف المؤسسات التعليمية، وكانت الأداة المعتمدة بشكل واضح في إجراء الدراسة هي الإحصاء والتحليل، وقد شملت عينة الدراسة على عدد من المؤسسات التعليمية ذات أماكن مختلفة في الجزائر لكنها ذات منهج موحد ومتنوع في الاختبارات للفصل الأول والثاني لمادة اللغة العربية وذلك للفصل الدراسي لعام 2017- 2018، وأشارت نتائج الدراسة إلى أن معظم الأسئلة موضوعة بطريقة واضحة ومباشرة وموجهة بصورة أساسية إلى الفئة الضعيفة والتي يركز عليها المعلمون أكثر من غيرها، وذلك كونها تساعدهم في تحصيل أكبر قدر ممكن من النقاط.

دراسة (أبو عيش، 2017) هدفت الدراسة الحالية إلى التعرف على اتجاهات المتقدمات نحو الاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم (القدرات العامة، واختبار كفايات اللغة الإنجليزية STEP، واختبار القدرات العامة للجامعيين، واختبار الهيئة السعودية للتخصصات الصحية) بالطائف، والكشف عن وجود فروق دالة إحصائية بين متوسطات أبعاد مقياس الاتجاه نحو الاختبارات المحوسبة لدى عينة الدراسة الكلية من المتقدمات لتلك الاختبارات حسب مجموعات نوع الاختبار، ولتحقيق أهداف الدراسة قامت الباحثة ببناء مقياس الاتجاهات نحو الاختبارات المحوسبة شمل ثلاثة أبعاد، وتألفت عينة الدراسة من (408) مختبرة من المتقدمات للاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم، وللإجابة على أسئلة الدراسة، تم إجراء عدد من التحليلات الإحصائية وهي: استخراج المتوسطات الحسابية والانحرافات المعيارية والنسب المئوية كما تم حساب تحليل التباين أحادي الاتجاه (One way ANOVA) للفروق في المتوسطات الكلية لأبعاد مقياس الاتجاه نحو الاختبارات المحوسبة.

وقد توصلت الباحثة إلى النتائج التالية: –أن معدل عام الاتجاه نحو قلق الحاسب الآلي لدى عينة الدراسة الكلية من المتقدمات للاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم كان ضعيفا. حيث بلغت نسبته (50.3%). –أن معدل عام الاتجاه نحو الاعتقاد بفوائد الاختبار المحوسب لدى عينة الدراسة الكلية من المتقدمات للاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم كان كبيرا وبلغ (81.5%). –أن معدل عام الاتجاه نحو الاعتقاد بمساوئ الاختبار المحوسب لدى عينة الدراسة الكلية من المتقدمات للاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم كان متوسطًا حيث بلغت نسبته (59.7%). –أما بالنسبة لمعدل عام الاتجاه الكلى نحو الاختبارات المحوسبة لدى عينة الدراسة الكلية من المتقدمات للاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم كان كبيرا بنسبة (68.8%). –وجود فروق ذات دلالة إحصائية بين متوسطات الكلية لأبعاد مقياس الاتجاه نحو الاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم لدى عينة الدراسة الكلية من المتقدمات لتلك الاختبارات.

وفي ضوء ما أسفرت عنه الدراسة من نتائج فقد خرجت بعدد من التوصيات والمقترحات بزيادة مراكز الاختبارات المحوسبة في جميع مناطق المملكة وزيادة فترات انعقادها بقدر الإمكان للاستفادة من مزايا هذه الاختبارات. تفعيل استخدام الاختبارات المحوسبة في جميع مراحل التعليم وتدريب المعلمين والمعلمات على إعدادها والطلبة على استخدامها، إجراء دراسات مقارنة بين المتقدمين والمتقدمات للاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم للتعرف على اتجاهاتهم نحو هذه الاختبارات والفروق بينهم. إجراء المزيد من الدراسات حول الاختبارات المحوسبة لدى فئات أخرى من الطلبة والمقارنة بين أداء الطلبة في الاختبارات المحوسبة، والاختبارات الورقية.

دراسة طويل (2017) تمحورت الدراسة حول تقويم الاختبارات الموضوعية في كلية التمريض في جامعة تشرين في سوريا خلال الفترة 2014-2015 وطبقا لنوعية الدراسة اتبع الباحث المنهج الوصفي التحليلي، وقام الباحث بتصميم أداة تحليل المحتوي المناسبة لمثل هذا النوع من الدراسات، وتكونت عينة البحث من 988 سؤالا ذات أنواع مختلفة، وتبين للباحث بعد تطبيق الدراسة وعمل الاحصائيات الملائمة، ان أعضاء الهيئة التدريسية يتبعون شروط صياغة المتن في الأسئلة المتعددة مع الحاجة إلى الانتباه إلى الوضوح والدقة اللغوية، وبشكل عام اتبع المدرسون للشروط الملائمة لصياغة الأسئلة حصل بشكل عام على درجة قليل جدا، وأوصي الباحث بضرورة العمل على ضمان جودة الأسئلة بحيث تناسب جميع المعايير التي وضعت لأجلها.

دراسة (Alsamaani, 2014) تناولت هذه الدراسة صلاحية عملية تطوير اختبار تحصيل الطلاب في اللغة الإنجليزية كلغة أجنبية في السياق السعودي وأثرها على أداء الاختبار. اشتملت موضوعات الدراسة على 50 مدرسا مبتدئا في المرحلة الإعدادية لتعليم اللغة الإنجليزية كلغة أجنبية في اثنتي عشرة مدرسة حكومية متوسطة في منطقة القصيم. تم جمع البيانات باستخدام التقييم الذاتي القائم على المهام لمعارف ومهارات التقييم الصفي، وبروتوكول التفكير بصوت عالٍ، وتقييم استبيان احتياجات التدريب. بالإضافة إلى ذلك، تم تحليل (150) اختبار لغة وأنشطة تقييم من حيث مقياس تقييم اختبار اللغة. كشفت النتائج أن التدريب في تقييم متعلم اللغة الإنجليزية كلغة أجنبية على مستوى الخريجين بدا غير كافٍ لتلبية احتياجات معلمي اللغة الإنجليزية كلغة أجنبية المبتدئين الذين يبدو أنهم يكتسبون معظم معرفتهم بالاختبار والتقييم من خلال الخبرة أثناء العمل، وليس من خلال التعليم قبل الخدمة في على مستوى الجامعة. بالإضافة إلى ذلك، وجدت الدراسة أن الاختبارات التي يقوم بها المعلمون لا تفي بمعايير تطوير اختبار تحصيل الطلاب. علاوة على ذلك، يكشف تحليل احتياجات تقييم معلمي اللغة الإنجليزية كلغة أجنبية للمبتدئين في المدارس الإعدادية بحاجة إلى تطوير معارفهم ومهاراتهم في مجموعة واسعة من تقييمات متعلمي اللغة الإنجليزية كلغة أجنبية. تقترح الدراسة أنه يجب على معلمي اللغة الإنجليزية كلغة أجنبية أن يبدأوا عملية تغيير ممارسات التقييم الحالية أو البحث عن طرق من شأنها أن تساعد في تطوير تقنيات التقييم المحسنة. بالإضافة إلى ذلك، يوصى بإجراء مزيد من البحث لتصميم برامج التدريب أثناء الخدمة وتجربتها في مجال تقييم المتعلمين أثناء الخدمة.

دراسة (Khora, 2021) هدفت هذه الدراسة لمعرفة مدي التناغم بين أنشطة كتاب اللغة الإنجليزية للصف العاشر والأهداف المتوخاة من الطلبة بعد الانتهاء من الصف العاشر وامتحانات المعلمين في ضوء الإطار الأوروبي المشترك. ولتحقيق غرض الدراسة وللإجابة عن أسئلة البحث قامت الباحثة بتبني المنهج الوصفي التحليلي واختارت عينة مكونة من كتاب اللغة الإنجليزية الفلسطيني للصف العاشر، عينة مكونة من (55) اختبارا نصفيا ونهائيا قام بإعدادها معلمو الصف العاشر، وعينة تكونت من (10) من معلمي اللغة الإنجليزية تم اختيارهم بشكل عشوائي للإجابة على المقابلة المنظمة. قامت الباحثة باستخدام أداتين للدراسة وهما الإطار الأوروبي المشترك لتحليل أنشطة الكتاب وفقها، ومقابلة منظمة، وتم التحقق من صدق أداتي الدراسة وثباتهما. حيث تم استخدام معادلة هولستي لحساب ثبات الأداة. تم تحليل البيانات التي تم جمعها ومعالجتها من خلال التكرارات ونسبها المئوية، وبعدها تم استخدام معادلة بورتر لإيجاد التناغم ما بين الأهداف والأنشطة والاختبارات.

كشفت النتائج الرئيسية للدراسة أن معظم مهام الكتب المدرسية كانت من المستوى A2، وكانت أعلى المهام هي مهام القراءة والكتابة. امتحانات المعلم كانت أيضا على مستوى A2 مع تركيز عالٍ على مهارات القراءة والكتابة. أظهرت النتائج أن معظم الامتحانات استبعدت التحدث والاستماع. كان هناك تناغما منخفضا بين أنشطة الكتاب المدرسي للصف العاشر والأهداف المطلوب من الطالب تحقيقها بحلول نهاية الصف العاشر، كما أن هناك تناغما معتدلا بين أنشطة الكتاب المدرسي والامتحانات. كذلك حصل مستوى التطبيق في هرم بلوم المعرفي على أعلى نسبة بينما لم يحقق مستوى التقييم أي درجة. بناء على النتائج السابقة، أوصت الدارسة مصممي الكتب المدرسية بإعادة النظر في طبيعة الأنشطة المتواجدة في كتاب اللغة الإنجليزية من أجل مواءمتها مع أهداف الكتاب المدرسي، وإعادة توزيع المهارات ليصبح متوازنا لكل مهارة، وتصميم دورات تدريبية للمعلمين حول تطوير الاختبارات لتتماشى مع المعايير الأوروبية المشتركة. (CEFRL).

المحور الثاني: الدراسات المتعلقة بمعايير تقويم الاختبارات اللغوية

دراسة (المالكي وفخرو، 2017) هدفت الدراسة إلى تعرف واقع الاختبارات الوطنية في ضوء معايير جودة التقويم، ومدى تحقيقها للمعايير الخاصة بالمناهج القطرية في المواد الدراسية الأربع (اللغة العربية، اللغة الإنجليزية، الرياضيات، والعلوم)، بالصفوف من الرابع إلى السادس بالمدارس الابتدائية المستقلة في دولة قطر. ولتحقيق أهداف الدراسة تم تصميم سلم مصفوفة الاختبارات الوطنية، ويتضمن أربعة محاور، هي: "تعليمات الاختبار"، "اتفاق مفردات الاختبار مع معايير المنهج"، "اللغة"، و"التنظيم". وتكون مجتمع الدراسة من (12) ورقة اختبارية من الاختبارات الوطنية، وفي المواد الدراسية الأربع وللصفوف الثلاثة المتقدمة، التي طبقت خلال السنتين (2011-2012/2012-2013)، وقد أشارت نتائج الدراسة إلى وجود محور في مستوي غير متحقق في مادتي العلوم والرياضيات وهو محور "تعليمات الاختبار"؛ حيث وجود القصور في بند "زمن الاختبار الكلي"، و"الدرجة المعطاة لكل سؤال"، و"تناسب عدد الأسئلة مع زمن الاختبار الكلي في المواد الرئيسة الأربعة" ولجميع الصفوف الدراسية، إضافة إلى وجود قصور في معايير خاصة في بعض المواد الدراسية الرئيسة دون الأخرى، وذلك في محوري "اتفاق مفردات الاختبار مع معايير المنهج"، و"التنظيم"، وكان "محور اللغة" في مستوي مرض في المواد الأربع الرئيسة. وفي ضوء هذه النتائج تم تقديم توصيات يمكن أن تسهم في تحسين مستوي نماذج الاختبارات الوطنية بالمدارس الابتدائية.

دراسة (Salim et al., 2017) هدفت هذه الدراسة إلى التحقق من الاختبارات في ELT وتأثيراتها العكسية على طلاب جامعه السلام في كليات التربية والآداب. الهدف الرئيسي من هذه الدراسة بيان أهمية الاختبار للمعلمين والطلاب لتطوير كفاءات الأساتذة حيث يقوم بتقييم وتقويم الطلاب. اعتمد الباحث المنهج الوصفي التحليلي، وتم جمع بيانات الدراسة بواسطة أداة استبيان منظم. تكونت عينة الدراسة من (15) معلما للغة الإنجليزية كلغة أجنبية تم اختيارهم بشكل عشوائي من كليات التربية والآداب. ثم تمت معالجة البيانات وتحليلها إحصائيا. توصل الباحث إلى النتائج التالية: غياب بعض الاختبارات الشخصية، تؤثر على الطلاب، أداء الطلاب في إتقان اللغة الإنجليزية ضعيف، عدم وجود دورات ELT منتظمة للمحاضرين وليس هناك أهداف واضحة لبعض اختبارات اللغة الإنجليزية. على ضوء هذه النتائج تم التوصل إلى هذه التوصية: يعتبر الاختبار مهما لعمليات التعلم والتدريس لأنه يحفز الطلاب والمعلمين ويقودهم لتحقيق إنجازات أفضل، ضرورة الاهتمام بأهداف تدريس اللغة الإنجليزية، كما يجب التفريق بين مستويات الطلاب وتشجيعهم على أن يكونوا أفضل في اللغة الإنجليزية.

التعليق على الدراسات السابقة :

- **من حيث الهدف**

هدفت دراسة أبو جامع (2021) حول تقويم اختبارات الأحياء للصف الثاني عشر في ضوء عمليات التعلم، بينما دراسة الغزال (2020) حول تقويم الاختبارات التحصيلية بمرحلة الدراسات العليا بقسم اللغة العربية وآدابها بكلية الآداب جامعة مصراته في ضوء معايير الجودة الشاملة، ودراسة الجار (2020) حول تقويم واقع الاختبارات الفصلية في بعض كليات جامعة شقراء كما تراها الطالبات في ضوء بعض المتغيرات، ودراسة البلوي (2019) حول تقويم الاختبارات المدرسية التي تعدها معلمات العلوم في المرحلة الأساسية بمدينة تبوك من حيث تخطيطها وإخراجها وتصحيحها، ودراسة سويلم (2018) تقويم اختبارات طلاب الشعب العلمية بكلية التربية بجامعة المنيا في ضوء معايير جودة الاختبارات، ودراسة طويل (2017) تمحورت الدراسة حول تقويم الاختبارات الموضوعية في كلية التمريض خلال الفترة 2014-2015.

- **من حيث المنهج المتبع:**

تناول دراسة أبو جامع (2021) المنهج الوصفي، ودراسة الغزال (2020) المنهج الوصفي التحليلي، دراسة الجار (2020) المنهج الوصفي التحليلي، دراسة البلوي (2019) الباحث المنهج الوصفي التحليلي، دراسة سويلم (2018) المنهج الوصفي التحليلي، دراسة طويل (2017) المنهج الوصفي التحليلي. وبذلك اتفقت الدراسات مع دراسة الباحث في تناولها المنهج الوصفي التحليلي.

- اتفقت الدراسة الحالية مع الدراسات السابقة في تناول تقويم الاختبارات المختلفة في ضوء معايير متنوعة.
- اختلفت دراسة الباحث عن جميع دراسات المحور في تناولها تقويم الاختبارات العالمية الدولية في جودة اختبارات اللغة الإنجليزية العالمية.

أوجه التميز للدراسة الحالية :

تميزت الدراسة الحالية عن الدراسات السابقة فيما يلي:

1. تناولت الدراسة الحالية موضوعًا في غاية الأهمية والذي تمثل في العنوان: "تقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية في ضوء معايير جودة اختبارات اللغة الإنجليزية العالمية".
2. حاولت الدراسة دراسة موضوع تقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية في ضوء معايير جودة اختبارات اللغة الإنجليزية العالمية.

أوجه الاستفادة من الدراسات السابقة :

تمثلت أوجه الاستفادة من الدراسات السابقة في:

1. بلورة وبناء محتوى الإطار النظري للدراسة وإعداده بطريقة علمية بحثية صحيحة.
2. اختيار المنهج البحثي الملائم والمناسب لتطبيق إجراءات الدراسة.
3. استخدام الأساليب الإحصائية والتحليلية الملائمة في تناول طبيعة الدراسة البحثية.
4. تدعيم كافة نتائج الدراسة الحالية بالدراسات السابقة.
5. تعريف مصطلحات الدراسة البحثية.
6. اختيار قائمة بمعايير جودة اختبارات اللغة الإنجليزية العالمية.
7. المساهمة في تصميم أداة الاستبيان الخاصة بتقويم اختبارات هيئة تقويم التعليم والتدريب اللغوية.
8. الاستفادة من الدراسات السابقة في تحليل نتائج البحث الحالي وتفسيرها ومقارنتها مع الدراسات السابقة.

منهجية وإجراءات الدراسة

تمهيد:

هدف البحث الحالي إلى التعرف على مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار كفايات اللغة الإنجليزية STEP، ومدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

منهجية الدراسة:

من أجل تحقيق أهداف الدراسة ؛ اتبعت الدراسة الحالية المنهجية الوصفية التحليلية وهي "إحدى أشكال التحليل والتفسير العلمي المنظم لوصف ظاهرة أو مشكلة محددة وتصويرها كميا عن طريق جمع البيانات ومعلومات معينة عن ظاهرة أو مشكلة وتصنيفها وتحليلها وإخضاعها للدراسة الدقيقة" (عبد المؤمن، 2008: ص287)

مجتمع الدراسة:

المجتمع يشتمل على جميع اختبارات هيئة تقويم التعليم والتدريب اللغوية وينقسم إلى قسمين:

1. اختبار كفايات اللغة الإنجليزية STEP.
2. اختبار تحديد المستوى في اللغة الإنجليزية EPT.

عينة الدراسة:

شملت عينة الدراسة اختباري كفايات اللغة الإنجليزية STEP، وتحديد المستوى في اللغة الإنجليزية EPT، حيث إنهما الاختباران اللذان أتيح للباحث الحصول على نماذج منهما.

أداة الدراسة:

بعد أن تم الاطلاع على الأدب التربوي، والدراسات السابقة المتعلقة بموضوع البحث، قام الباحث ببناء وتطوير أداة الدراسة والتي تمثلت في بطاقة تقويم تحتوي على قائمة بمعايير اختبارات اللغة الإنجليزية العالمية بهدف التعرف على مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختباري كفايات اللغة الإنجليزية ، ومدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار تحديد المستوى في اللغة الإنجليزية STEP, EPT (ملحق رقم 1).

وصف أداة الدراسة (قائمة بمعايير اختبارات اللغة الإنجليزية العالمية):

احتوت أداة في بناء الأداة على مجموعة من المعايير والتي تمثلت في:

1. المعايير العامة.
2. معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها.
3. المعايير الخاصة بمتعلمي اللغة الإنجليزية.
4. المعايير الخاصة بمجالات الاختبار.
5. المعايير الخاصة بالسياق المجتمعي.
6. معايير الجودة الخاصة ببناء الاختبار.
7. معايير الجودة الخاصة بتصميم الاختبار.

وقد احتوت بطاقة التقويم في صورتها الأولية على جزأين رئيسين:

- **المحور الأول: "مستوى توافر المعايير العامة"** ويتكون من (6) عبارات.
- **المحور الثاني: "مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والاشراف عليها"** ويتكون من (4) عبارات.
- **المحور الثالث: "مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية"** ويتكون من (4) عبارات.
- **المحور الرابع: "مستوى توافر المعايير الخاصة بمجالات الاختبار** ويتكون من (4) عبارات.

- **المحور الخامس: "مستوى توافر المعايير الخاصة بالسياق المجتمعي "** ويتكون من (2) عبارتين فقط.
- **المحور السادس: "معايير الجودة الخاصة ببناء الاختبار "** ويتكون من (11) عبارة.
- **المحور السابع: "معايير الجودة الخاصة بتصميم الاختبار "** ويتكون من (9) عبارات.

ثم استخدم الباحث مقياس خاص لتقييم الاختبارات (ليكارت الخماسي)، وذلك وفق تدرج من (1-5)؛ للتعرف على مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار كفايات اللغة الإنجليزية STEP، ومدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار تحديد المستوى في اللغة الإنجليزية EPT. حيث يشير الرقم 1 إلى ضعف توفر المعيار في الاختبارات، أما الرقم 4 فيشير إلى توفر المعيار، وبعد إجراء التقييم تمت عملية التحليل الإحصائي للخروج بالنتائج النهائية.

صدق وثبات أداة الدراسة:

1) الصدق الظاهري

بعد أن انتهى الباحث من إعداد قائمة بمعايير اختبارات اللغة الإنجليزية العالمية وبناء فقراتها، قام بعرضها في صورتها الأولية على مجموعة من المحكمين؛ وبعد استعادة النسخ المحكمة من المحكمين ومناقشة ملاحظاتهم مع سعادة الدكتور المشرف على الرسالة وفي ضوء اقتراحات بعض المحكمين أعاد الباحث صياغة بطاقة التقويم حيث تم إعادة صياغة بعض العبارات في بطاقة التقويم وفق ما اتفق عليه أكثر من (80%) من السادة المحكمون حيث بلغ عدد العبارات في صورة الاستبيان الأولية أربعون عبارة، لتصبح بطاقة التقويم في صورتها النهائية والتي بلغت أيضاً أربعون عبارة.

2) صدق الاتساق الداخلي لأداة الدراسة:

قام الباحث بحساب صدق الاتساق الداخلي وفقا لاستجابات أفراد العينة الاستطلاعية (ن=30) بحساب معامل ارتباط بيرسون بين درجات كل عبارة والدرجة الكلية للمحور الذي تنتمي إليه العبارة من محاور بطاقة التقويم للتأكد من توافر درجة عالية من صدق الاتساق الداخلي لعبارات محاور القائمة.

3) الصدق البنائي العام لمحاور قائمة بمعايير اختبارات اللغة الإنجليزية العالمية ":

قام الباحث بالتحقق من الصـــدق البنائي لمحاور بطاقة التقويم من خلال إيجاد معاملات الارتباط بين الدرجة الكلية لكل محور والمجموع الكلي لبطاقة التقويم للتأكد من توافر درجة عالية من الصدق البنائي لمحاور القائمة.

4) ثبات أداة الدراسة (قائمة بمعايير اختبارات اللغة الإنجليزية العالمية)

قام الباحث بتطبيق اختبار ألفا كرونباخ لمحاور بطاقة التقويم للتأكد من أن القيم من معاملات الثبات تؤكد على صـلاحية بطاقة التقويم للتطبيق وإمكانية الاعتماد على نتائجها والوثوق بها.

جدول رقم (3) يوضح ثبات التحليل بواسطة الباحث والمشرف التربوي وفق التكرارات ونسبها المئوية

م	العبارة	تكرارات الباحث	تكرارات المشرف التربوي	مجموع التكرارات	مجموع الاتفاق	النسبة المئوية
المحور الأول: مستوى توافر المعايير العامة						
1	تخلو الاختبارات من قضايا التحيز والحساسية	8	8	16	16	100%
2	يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية	8	8	16	16	100%
3	يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية	8	8	16	16	100%
4	يتم تحديد مجال الاختبار بدقة	7	6	13	9	69%
5	ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار	6	4	10	6	60%
6	تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم	8	8	16	16	100%

المحور الثاني: مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها						
1	يتمتع أعضاء اللجنة بخبرة عميقة في المجالات ذات الصلة بالمعايير أو التقييمات أو كليهما	8	8	16	16	100%
2	يعكس أعضاء اللجنة التوازن المناسب في التمثيل ووجهات النظر	8	8	16	16	100%
3	أعضاء اللجنة موجهين بشكل مناسب ويتم تدريبهم	8	8	16	16	100%
4	يتوفر لأعضاء اللجنة الوقت الكافي لتنفيذ المهمة (المهام) المطلوب منهم إكمالها	8	8	16	16	100%
المحور الثالث: مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية						
1	يعتبر الاختبار التنوع بين خصائص متعلمي اللغة الإنجليزية	8	8	16	16	100%
2	إشراك الخبراء الذين لديهم دراية بنطاق التنوع الذي يمثله طلاب دارسي الإنجليزية	8	6	14	8	57%
3	هناك معايير مختلفة أو توقعات إنجاز لمختلف مجموعات دارسي الإنجليزية الفرعية	8	6	14	8	57%
4	يتم تطوير مجموعة واحدة من معايير الاختبارات الصارمة لجميع طلاب دارسي الإنجليزية	6	8	14	8	57%
المحور الرابع: مستوى توافر المعايير الخاصة بمجالات الاختبار						
1	يفحص الاختبار القواعد النظرية والبحثية والتنمية الاجتماعية والثقافية وعلم اللغة الاجتماعي	6	6	12	12	100%
2	يقيم الاختبار اتساع وعمق ومدى التعقيد والتعبير عن المهارات والمعرفة المرتبطة بها	6	8	14	8	57%
3	تحديد المهارات والمعرفة اللغوية المطلوبة في الاستماع والتحدث والقراءة والكتابة	8	8	16	16	100%
4	يركز الاختبار على الترابط بين الطرق الأربعة والمساهمة في إتقان اللغة الإنجليزية	6	8	14	8	57%

المحور الخامس: مستوى توافر المعايير الخاصة بالسياق المجتمعي						
1	عند تصميم الاختبار يراعى الإلمام بالسكان، السياسات والممارسات والبرامج السابقة والحالية	8	8	16	16	100%
2	يتم تطوير هيكل الاختبار بناء على السياق التاريخي لتعليم وتعلم اللغة الإنجليزية	8	8	16	16	100%
المحور السادس: معايير الجودة الخاصة ببناء الاختبار						
1	الصدق	8	8	16	16	100%
2	الثبات	6	8	14	8	57%
3	الصياغة الواضحة	8	8	16	16	100%
4	التنوع	8	8	16	16	100%
5	الدقة	8	8	16	16	100%
6	مصداقية	8	8	16	16	100%
7	قابلة التطبيق	8	8	16	16	100%
8	عدالة التوزيع	6	8	14	8	57%
9	سهولة الاشتراك	6	8	14	8	57%
10	الشمول	8	8	16	16	100%
11	التكامل بين مستويات المعرفة	6	8	14	8	57%
المحور السابع: معايير الجودة الخاصة بتصميم الاختبار						
1	تحديد الأهداف اللغوية التي يقيسها الاختبار	6	8	14	8	57%
2	تحديد الكفايات اللغوية التي يقيسها الاختبار	8	8	16	16	100%
3	تحديد العناصر اللغوية التي يقيسها الاختبار	8	8	16	16	100%
4	تحديد الكفايات اللغوية التي سيغطيها الاختبار	6	8	14	8	57%
5	تحديد العناصر اللغوية التي سيغطيها الاختبار	6	8	14	8	57%
6	تحديد المهارات اللغوية التي يستهدف الاختبار قياسها	8	8	16	16	100%
7	تحديد نسبة الأسئلة المقالية إلى الموضوعية في الاختبار	6	7	13	9	69%
8	تحديد عدد البدائل الخاصة بالأسئلة الموضوعية	8	7	15	12	80%
9	تحديد آلية الإجابة على أسئلة الاختبار	8	8	16	16	100%
	المجموع	**293**	**306**	**599**	**512**	**%85**

يتضح من الجدول السابق أن معامل الثبات الكلي 85% وهو معامل ثبات قوي ومرتفع ويمكن من خلاله الوثوق بنتائج المقياس.

نتائج الدراسة، مناقشتها وتفسيرها

تمهيد :

يتناول هذا الفصل الأساليب الإحصائية التي تم تناولها، ثم الإجابة عن أسئلة الدراسة، ومن ثم الوصول إلى نتائج الدراسة وعرضها وتفسيرها ومناقشتها، وذلك على النحو التالي:

الأساليب الإحصائية:

فيما يلي عرض للأساليب الإحصائية التي اعتمد عليها الباحث في الدراسة الميدانية:

1- **التكرار والنسب المئوية لوصف أفراد الدراسة،** ولحساب استجابات أفراد الدراسة اتجاه العبارات التي تتضمنها أداة الدراسة.

2- **المتوسطات الحسابية** لترتيب استجابات أفراد الدراسة.

3- **معامل ارتباط بيرسون** للتأكد من صدق الاتساق الداخلي لأداة الدراسة.

4- **معامل ثبات (ألفاكرونباخ)** لقياس ثبات أداة الدراسة.

5- **معادلة المدى:** وذلك لوصف المتوسط الحسابي للاستجابات على كل عبارة وبعد على النحو التالي:

- من 1 إلى أقل من 1.80 تمثل درجة استجابة (ضعيفة جدا).
- من 1.80 إلى أقل من 2.60 تمثل درجة استجابة (ضعيفة).
- من 2.60 إلى أقل من 3.40 تمثل درجة استجابة (متوسطة).
- من 3.40 إلى أقل من 4.20 تمثل درجة استجابة (عالية).
- من 4 إلى أقل من 5 تمثل درجة استجابة (عالية جدا).

الإجابة عن السؤال الأول :

ما مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار كفايات اللغة الإنجليزية STEP؟

للإجابة على هذا السؤال تم حساب المتوسط الحسابي والانحراف المعياري لدرجة تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار كفايات اللغة الإنجليزية STEP، والتي تضمنتها قائمة المعايير التي قام الباحث بإعدادها وقد بلغ عددها (7) معايير رئيسية، ويمكن عرض هذه النتائج الخاصة بهذه المعايير من خلال ما يلي:

أولا: مستوى توافر المعايير العامة

جدول 4: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر المعايير العامة

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	تخلو الاختبارات من قضايا التحيز والحساسية	4.00	0.00	100%
2	يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية	4.00	0.00	100%
3	يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية	4.00	0.00	100%
4	يتم تحديد مجال الاختبار بدقة	3.00	0.00	75%
5	ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار	2.50	0.71	63%

6	تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم	4.00	0.00	100%
	المتوسط العام	**3.58**	**0.12**	**90%**

يتضح من الجدول رقم (4) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر المعايير العامة"**، حيث تراوحت مستويات التقييم على هذا المحور ما بين 63%كحد أدنى للعبارة " ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار" إلى 100% لكل من العبارة " تخلو الاختبارات من قضايا التحيز والحساسية"، والعبارة " يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية"، والعبارة " يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية"، والعبارة " تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم" هذا وبلغت درجة التقييم الكلية على هذا المحور 90%.

ويبين الجدول رقم (4) أن قيم المتوسطات الحسابية لمستوى توافر المعايير العامة لجودة اختبارات اللغة الإنجليزية العالمية في اختبار كفايات اللغة الإنجليزية STEP تتراوح بين (4 – 2,5). وقد جاءت العبارة رقم (1) "تخلو الاختبارات من قضايا التحيز والحساسية" والعبارة رقم (2) "يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية" والعبارة رقم (3) "يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية" والعبارة رقم (6) "تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم" في المرتبة الأولى وبمتوسط حسابي (4,00)، وقد جاءت العبارة رقم (5) "ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار" في المرتبة الأخيرة وبمتوسط حسابي (2,5)، وقد بلغ المتوسط الحسابي العام لمحور مستوى توافر المعايير العامة (3,63)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير العامة مرتفعة في اختبار كفايات اللغة الإنجليزية STEP.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر المعايير العامة تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,71) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,12) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر المعايير العامة في اختبار كفايات اللغة الإنجليزية STEP.

ويعزو الباحث هذه النتيجة إلى أن إعداد اختبار كفايات اللغة الإنجليزية STEP يتم من خلال الاستفادة من المعايير العالمية لتصميم اختبارات تحديد المستوى في اللغة الإنجليزية، حيث تؤكد هذه المعايير على أهمية الموضوعية والحيادية وتجنب قضايا التحيز والحساسية في إعداد الاختبارات، إضافة إلى إعداد محتوى الاختبار ليركز على مستوى اكتساب اللغة الإنجليزية واللغويات التطبيقية، وذلك من خلال إشراف مجموعة من العاملين تمثل الخبرات الأساسية لتدريس وتعلم اللغة الإنجليزية وتكوين لجات متخصصة لإدارة الاختبار في كافة مراحله وخاصةً مراحل الإعداد والتنفيذ والتقييم، الأمر الذي يسهم في تحديد مجال هذه الاختبارات بدقة وتحقيق أهدافها، كما يرى الباحث أنه وعلى الرغم من المستوى الجيد لتوافر المعايير العامة، إلا أن العمل على تطوير هذه المعايير يعتبر أمرًا ضروريًا خاصةً فيما يتعلق بعمليات التقييم والرقابة.

وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة الغزال (2020) حيث أكدت أن عملية تصميم الاختبارات التحصيلية بمرحلة الدراسات العليا تتم بصورة واضحة ومناسبة لقياس قدرة الطلبة على ممارسة عمليات عقلية متنوعة.

ثانيا: مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها

جدول 5 : المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	يتمتع أعضاء اللجنة بخبرة عميقة في المجالات ذات الصلة بالمعايير أو التقييمات أو كليهما	4.00	0.00	100.0%
2	يعكس أعضاء اللجنة التوازن المناسب في التمثيل ووجهات النظر	4.00	0.00	100.0%
3	أعضاء اللجنة موجهين بشكل مناسب ويتم تدريبهم	4.00	0.00	100.0%

4	يتوفر لأعضاء اللجنة الوقت الكافي لتنفيذ المهمة (المهام) المطلوب منهم إكمالها	4.00	0.00	100.0%
	المتوسط العام	**4.00**	**0.00**	**100.0%**

يتضح من الجدول رقم (5) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها**"، حيث نالت جميع عبارات هذا المعيار على درجة تقييم 100%.

ويبين الجدول رقم (5) أن قيم المتوسطات الحسابية لمستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها جميعها متساوية حيث بلغت قيمتها (4,00) وقد بلغ المتوسط الحسابي العام لمحور مستوى توافر المعايير العامة (4,00)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير العامة مرتفعة في اختبار كفايات اللغة الإنجليزية STEP.

كما يتضح من الجدول السابق عدم وجود انحراف معياري لمستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها حيث أن قيمة الانحراف المعياري لجميع العبارات كانت (0,00)، كما بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,00) وتشير هذه القيمة إلى التطابق في استجابات أفراد عينة الدراسة حول مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها في اختبار كفايات اللغة الإنجليزية STEP.

ويعزو الباحث هذه النتيجة إلى التحول الكبير الذي يشهده قطاع التعليم في المملكة العربية السعودية والسعي المستمر إلى تطويره من مختلف جوانبه، وهو ما أسهم وبدرجة كبيرة في زيادة مستوى كفاءة المشاركين في إعداد وتصميم اختبار كفايات اللغة الإنجليزية STEP، وامتلاكهم لخبرة عالية في مختلف المجالات الخاصة بالمعايير العالمية لإعداد اختبارات اللغة الإنجليزية وتطبيقها في الاختبار في عمليات الاختبار المختلفة، كما تدل هذه النتيجة أيضًا إلى أن المسؤولين عن اختبار كفايات اللغة الإنجليزية STEP يدركون جيدًا أهمية الاختبار ويسعون إلى توفير كافة العوامل التي من شأنها أن تسهم في إنجاحه وتقديمه بصورته المناسبة للطلاب من خلال استقطاب الموارد البشرية المناسبة وتوفير كافة الإمكانات التي يحتاجونها.

وتختلف هذه النتيجة مع ما توصلت إليه نتائج الجار (2020) والتي توصلت إلى أن مستوى ممارسات أعضاء الهيئات التدريسية في إعداد الاختبارات الفصلية جاء بدرجة متوسطة مما يدلل على ضرورة تطوير أعضاء الهيئات التدريسية بما يتماشى مع المعايير الحديثة لتصميم الاختبارات، كما اختلفت أيضًا مع دراسة البلوي (2019) التي أشارت نتائجها إلى أن درجة التزام المعلمات بمهارات التخطيط وتصحيح الاختبارات التحصيلية جاء بدرجة متوسطة.

ثالثا: مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية

جدول 6 : المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	يعتبر الاختبار التنوع بين خصائص متعلمي اللغة الإنجليزية	4.00	0.00	100.0%
2	إشراك الخبراء الذين لديهم دراية بنطاق التنوع الذي يمثله طلاب دارسي الإنجليزية	3.50	0.71	88%
3	هناك معايير مختلفة أو توقعات إنجاز لمختلف مجموعات دارسي الإنجليزية الفرعية	3.50	0.71	88%
4	يتم تطوير مجموعة واحدة من معايير الاختبارات الصارمة لجميع طلاب دارسي الإنجليزية	3.50	0.71	88%
	المتوسط العام	**3.63**	**0.18**	**91%**

يتضح من الجدول رقم (6) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 88% كحد أدنى لكل من العبارات " إشراك الخبراء الذين لديهم دراية بنطاق التنوع الذي يمثله طلاب دارسي الإنجليزية"، والعبارة " هناك معايير مختلفة أو توقعات إنجاز لمختلف مجموعات دارسي الإنجليزية الفرعية"، والعبارة " يتم تطوير مجموعة واحدة من معايير الاختبارات الصارمة لجميع طلاب دارسي الإنجليزية" إلى 100% للعبارة " يعتبر الاختبار التنوع بين خصائص متعلمي اللغة الإنجليزية"، هذا وبلغت درجة التقييم الكلية على هذا المحور 91%.

ويبين الجدول رقم (6) أن قيم المتوسطات الحسابية لمستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية في اختبار كفايات اللغة الإنجليزية STEP تتراوح بين (4 – 3,5). وقد جاءت العبارة رقم (1) "يعتبر الاختبار التنوع بين خصائص متعلمي اللغة الإنجليزية" في المرتبة الأولى وبمتوسط حسابي (4,00)، بينما جاءت العبارة رقم (2) " إشراك الخبراء الذين لديهم دراية بنطاق التنوع الذي يمثله طلاب دارسي الإنجليزية " والعبارة رقم (3) " إشراك معايير مختلفة أو توقعات إنجاز لمختلف مجموعات دارسي الإنجليزية الفرعية" والعبارة رقم (4) "يتم تطوير مجموعة واحدة من معايير الاختبارات الصارمة لجميع طلاب دارسي الإنجليزية" في المرتبة الأخيرة وبمتوسط حسابي (3,5)، وقد بلغ المتوسط الحسابي العام لمحور توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية (3,63)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية مرتفعة في اختبار كفايات اللغة الإنجليزية STEP.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,71) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,18) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية في اختبار كفايات اللغة الإنجليزية STEP.

ويعزو الباحث هذه النتيجة إلى الاهتمام بتصميم اختبار كفايات اللغة الإنجليزية STEP ليتوافق مع خصائص جميع متعلمي اللغة الإنجليزية، وهي من المعايير المهمة لتصميم الاختبارات عمومًا والاختبارات اللغوية خصوصًا، كما تشير هذه النتيجة إلى أن عملية إعداد الاختبار تتم من خلال مشاركة مجموعة من الخبراء المختصين في تصميم الاختبارات اللغوية، مما يزيد من كفاءة هذا الاختبار ويتيح لجميع الطلاب المشاركة العادلة والحصول على نتائج جيدة.

وتتفق هذه النتيجة مع النتائج التي توصلت إليها دراسة أبو عيش (2017) والتي أشارت إلى التوجهات الإيجابية للطلبة للمشاركة في الاختبارات وخاصةً تلك التي تتناسب مع توجهاتهم وخصائصهم، كما اتفقت مع نتائج دراسة (Salim et al., 2017) التي أكدت على أهمية مراعاة خصائص الطلاب في الاختبارات وعمليات التقييم، بينما اختلفت مع نتائج دراسة (Alsamaani, 2014) التي بينت أن معلمي اللغة الإنجليزية كلغة أجنبية في المدارس الإعدادية المبتدئين بحاجة إلى تطوير معارفهم ومهاراتهم في مجموعة واسعة من تقييمات متعلمي اللغة الإنجليزية كلغة أجنبية، وذلك ليتمكنوا من تقييم أداء الطلاب ونتائج تعلمهم للغة الإنجليزية كلغة ثانية وفقًا لخصائصهم.

رابعا: مستوى توافر المعايير الخاصة بمجالات الاختبار

جدول 7 : المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر المعايير الخاصة بمجالات الاختبار

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	يفحص الاختبار القواعد النظرية والبحثية والتنمية الاجتماعية والثقافية وعلم اللغة الاجتماعي	3.00	0.00	75.0%
2	يقيم الاختبار اتساع وعمق ومدى التعقيد والتعبير عن المهارات والمعرفة المرتبطة بها	3.50	0.71	88%
3	تحديد المهارات والمعرفة اللغوية المطلوبة في الاستماع والتحدث والقراءة والكتابة	4.00	0.00	100.0%
4	يركز الاختبار على الترابط بين الطرق الأربعة والمساهمة في إتقان اللغة الإنجليزية	3.50	0.71	88%
	المتوسط العام	**3.50**	**0.35**	**88%**

يتضح من الجدول رقم (7) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر المعايير الخاصة بمجالات الاختبار**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 75%كحد أدني للعبارة " يفحص الاختبار القواعد النظرية والبحثية والتنمية الاجتماعية والثقافية وعلم اللغة الاجتماعي" إلى 100% للعبارة " تحديد المهارات والمعرفة اللغوية المطلوبة في الاستماع والتحدث والقراءة والكتابة"، هذا وبلغت درجة التقييم الكلية على هذا المحور 87.5%.

ويبين الجدول رقم (7) أن قيم المتوسطات الحسابية لمستوى توافر المعايير الخاصة بمجالات الاختبار في اختبار كفايات اللغة الإنجليزية STEP تتراوح بين (4 – 3). وقد جاءت العبارة رقم (3) "تحديد المهارات والمعرفة اللغوية المطلوبة في الاستماع والتحدث والقراءة والكتابة" في المرتبة الأولى وبمتوسط حسابي (4,00)، بينما جاءت العبارة رقم (2) " يقيم الاختبار اتساع وعمق ومدى التعقيد والتعبير عن المهارات والمعرفة المرتبطة بها" والعبارة رقم (4) " يركز الاختبار على الترابط بين الطرق الأربعة والمساهمة في إتقان اللغة الإنجليزية" في المرتبة الثانية وبمتوسط حسابي (3,5)، وجاءت العبارة رقم (1) "يفحص الاختبار القواعد النظرية والبحثية والتنمية الاجتماعية والثقافية وعلم اللغة الاجتماعي" في المرتبة الأخيرة وبمتوسط حسابي (3)، وقد بلغ المتوسط الحسابي العام لمحور توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية (3,5)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير الخاصة بمجالات الاختبار مرتفعة في اختبار كفايات اللغة الإنجليزية STEP.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر المعايير الخاصة بمجالات الاختبار تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,71) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,35) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر المعايير الخاصة بمجالات الاختبار في اختبار كفايات اللغة الإنجليزية STEP.

ويعزو الباحث هذه النتيجة إلى مستوى الاهتمام المتعلق بتقييم امتلاك المتعلمين المعرفة اللغوية المختلفة والتي ترتبط بالكفايات اللغوية الأساسية المرتبطة بمهارات اللغة الإنجليزية المتمثلة في الاستماع والتحدث والقراءة والكتابة، ويرجع ذلك إلى أن إتقان هذه الكفايات يعتبر من أهم الأهداف التي تساعد في تعليم اللغة الإنجليزية، إضافةً إلى التعرف على المستوى العام لإتقان اللغة الإنجليزية لدى المتعلمين، وأن التصميم الأساسي للاختبار يدور حول معرفة مدى امتلاك الطلاب لهذه الكفايات.

وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة أبو جامع (2021) والتي بينت أن اختبارات الأحياء تغطي جميع المهارات المطلوبة وبصورة متوازنة. وتختلف هذه النتيجة مع ما توصلت إليه نتائج دراسة (Alsamaani, 2014) التي وجدت أن الاختبارات التي يقوم بها المعلمون لا تفي بمعايير تطوير اختبار تحصيل الطلاب، وأن المعلمين بحاجة إلى تطوير معارفهم ومهاراتهم في مجموعة واسعة من تقييمات متعلمي اللغة الإنجليزية كلغة أجنبية، ودراسة (Khora, 2021)

التي أشـــارت إلى أن اختبارات اللغة الإنجليزية لا تغطي جميع المهارات اللغوية، حيث تركز الاختبارات على مهارات القراءة والكتابة وتســـتبعد مهارات التحدث والاستماع.

خامسا: مستوى توافر المعايير الخاصة بالسياق المجتمعي

جدول 8: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر المعايير الخاصة بالسياق المجتمعي

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	عند تصميم الاختبار يراعى الإلمام بالسكان، السياسات والممارسات والبرامج السابقة والحالية	4.00	0.00	100.0%
2	يتم تطوير هيكل الاختبار بناء على السياق التاريخي لتعليم وتعلم اللغة الإنجليزية	4.00	0.00	100.0%
	المتوسط العام	**4.00**	**0.00**	**100.0%**

يتضح من الجدول رقم (8) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر المعايير الخاصة بالسياق المجتمعي**"، حيث نالت جميع عبارات هذا المعيار على درجة تقييم 100%.

ويبين الجدول رقم (8) أن قيم المتوسـطات الحسـابية لمحور مسـتوى توافر المعايير الخاصـة بالسـياق المجتمعي في اختبار كفايات اللغة الإنجليزية STEP جاءت متساوية وبقيمة قدرها (4,00). وبذلك فقد بلغ المتوسط الحسابي العام لمحور توافر المعايير الخاصة بالسياق المجتمعي (4,00)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير الخاصة بالسياق المجتمعي مرتفعة في اختبار كفايات اللغة الإنجليزية STEP.

كما يتضح من الجدول السابق أن الانحرافات للمتوسط الحسابي العام (0,00) وهي قيمة تشير إلى التطابق التام في استجابات أفراد عينة الدراسة حول مستوى توافر المعايير الخاصة بالسياق المجتمعي في اختبار كفايات اللغة الإنجليزية STEP.

ويعزو الباحث هذه النتيجة إلى أن اختبار كفايات اللغة الإنجليزية STEP يتم تصميمه وبناؤه بناءً على الواقع العملي لتعليم وتعلم اللغة الإنجليزية في المملكة، وهو ما يجعل منه اختبارًا واقعيًا لقياس مستوى امتلاك المتعلمين لكفايات اللغة الإنجليزية المختلفة ومستوى إتقانهم لها، الأمر الذي يزيد من أهمية النتائج التي يحصل عليها المتعلمين في هذا الاختبار.

سادسا: معايير الجودة الخاصة ببناء الاختبار

جدول 9: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر معايير الجودة الخاصة ببناء الاختبار

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	الصدق	4.00	0.00	100.0%
2	الثبات	3.50	0.71	88%
3	الصياغة الواضحة	4.00	0.00	100.0%
4	التنوع	4.00	0.00	100.0%
5	الدقة	4.00	0.00	100.0%
6	مصداقية	4.00	0.00	100.0%
7	قابلة التطبيق	4.00	0.00	100.0%
8	عدالة التوزيع	3.50	0.71	88%

9	سهولة الاشتراك	3.50	0.71	88%
10	الشمول	4.00	0.00	100.0%
11	التكامل بين مستويات المعرفة	3.50	0.71	88%
	المتوسط العام	**3.82**	**0.26**	**95%**

يتضح من الجدول رقم (9) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **معايير الجودة الخاصة ببناء الاختبار**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 88%كحد أدنى لكل من عبارات "الثبات"، "عدالة التوزيع"، "سهولة الاشتراك"، "التكامل بين مستويات المعرفية" إلى 100% لكل من العبارات " الصدق"، "الصياغة الواضحة"، "التنوع"، "الدقة"، " المصداقية"، "قابلية التطبيق"، "الشمول"، هذا وبلغت درجة التقييم الكلية على هذا المحور 95%.

ويبين الجدول رقم (9) أن قيم المتوسطات الحسابية لمستوى توافر معايير الجودة الخاصة ببناء الاختبار في اختبار كفايات اللغة الإنجليزية STEP تتراوح بين (4 – 3,5). وقد جاءت العبارة رقم (1) " الصدق" والعبارة رقم (3) " الصياغة الواضحة" والعبارة رقم (4) " التنوع" والعبارة رقم (5) " الدقة" والعبارة رقم (6) " مصداقية" والعبارة رقم (7) " قابلة التطبيق" والعبارة رقم (10) " الشمول" في المرتبة الأولى وبمتوسط حسابي (4,00)، بينما جاءت العبارة رقم (2) " الثبات" والعبارة رقم (8) "عدالة التوزيع" والعبارة رقم (9) "سهولة الاشتراك" والعبارة رقم (11) "التكامل بين مستويات المعرفة" في المرتبة الثانية وبمتوسط حسابي (3,5)، وقد بلغ المتوسط الحسابي العام لمحور معايير الجودة الخاصة ببناء الاختبار (3,82)، وهو ما يشير إجمالًا إلى أن درجة توافر معايير الجودة الخاصة ببناء الاختبار مرتفعة في اختبار كفايات اللغة الإنجليزية STEP.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر معايير الجودة الخاصة ببناء الاختبار تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,71) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,26) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر معايير الجودة الخاصة ببناء الاختبار في اختبار كفايات اللغة الإنجليزية STEP.

ويعزو الباحث هذه النتيجة إلى وجود خبرة علمية ومهنية لدى المسؤولين عن الاختبار والقائمين عن بنائه وهو ما يجعله اختبارًا يتوافق مع المعايير العالمية لاختبارات اللغة الإنجليزية الدولية، بحيث تكون شاملة ومناسبة لقياس الكفايات التي وضعت لقياسها.

وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة أبيش وعنشيل (2018) التي أشارت إلى أن معظم أسئلة الاختبارات موضوعة بطريقة واضحة ومباشرة وموجهة بصورة أساسية إلى الفئة الضعيفة والتي يركز عليها المعلمون أكثر من غيرها، وذلك كونها تساعدهم في تحصيل أكبر قدر ممكن من النقاط. وتختلف هذه النتيجة مع ما توصلت إليه نتائج دراسة سويلم (2018) التي أشارت إلى أن الاختبارات التي يتم تصميمها لا تتناسب مع المهارات التي يمتلكها الطلاب وغير مناسبة لقياس العديد من المهارات المرتبطة بالتفكير.

سابعا: معايير الجودة الخاصة بتصميم الاختبار

جدول 10: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر معايير الجودة الخاصة بتصميم الاختبار

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	تحديد الأهداف اللغوية التي يقيسها الاختبار	3.50	0.71	88%
2	تحديد الكفايات اللغوية التي يقيسها الاختبار	4.00	0.00	100.0%
3	تحديد العناصر اللغوية التي يقيسها الاختبار	4.00	0.00	100.0%
4	تحديد الكفايات اللغوية التي سيغطيها الاختبار	3.50	0.71	88%
5	تحديد العناصر اللغوية التي سيغطيها الاختبار	3.50	0.71	88%
6	تحديد المهارات اللغوية التي يستهدف الاختبار قياسها	4.00	0.00	100.0%

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
7	تحديد نسبة الأسئلة المقالية إلى الموضوعية في الاختبار	3.00	0.00	75.0%
8	تحديد عدد البدائل الخاصة بالأسئلة الموضوعية	4.00	0.00	100.0%
9	تحديد آلية الإجابة على أسئلة الاختبار	4.00	0.00	100.0%
	المتوسط العام	**3.72**	**0.24**	**93%**

يتضح من الجدول رقم (10) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **معايير الجودة الخاصة بتصميم الاختبار**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 75%كحد أدني للعبارة " تحديد نسبة الأسئلة المقالية إلى الموضوعية في الاختبار" إلى 100% لكل من العبارات " تحديد الكفايات اللغوية التي يقيسها الاختبار"، " تحديد العناصر اللغوية التي يقيسها الاختبار"، " تحديد المهارات اللغوية التي يستهدف الاختبار قياسها"، " تحديد عدد البدائل الخاصة بالأسئلة الموضوعية"، " تحديد آلية الإجابة على أسئلة الاختبار"، هذا وبلغت درجة التقييم الكلية على هذا المحور 93%.

ويبين الجدول رقم (10) أن قيم المتوسطات الحسابية لمستوى توافر معايير الجودة الخاصة ببناء الاختبار في اختبار كفايات اللغة الإنجليزية STEP تتراوح بين (4 – 3). وقد جاءت العبارة رقم (2) " تحديد الكفايات اللغوية التي يقيسها الاختبار" والعبارة رقم (3) " تحديد العناصر اللغوية التي يقيسها الاختبار" والعبارة رقم (6) " تحديد المهارات اللغوية التي يستهدف الاختبار قياسها" والعبارة رقم (8) " تحديد عدد البدائل الخاصة بالأسئلة الموضوعية" والعبارة رقم (9) " تحديد آلية الإجابة على أسئلة الاختبار" في المرتبة الأولى وبمتوسط حسابي (4,00)، بينما جاءت العبارة رقم (1) " تحديد الأهداف اللغوية التي يقيسها الاختبار" والعبارة رقم (4) "تحديد الكفايات اللغوية التي سيغطيها الاختبار" والعبارة رقم (5) "تحديد العناصر اللغوية التي سيغطيها الاختبار" في المرتبة الثانية وبمتوسط حسابي (3,5)، وقد جاءت العبارة رقم (7) "تحديد نسبة الأسئلة المقالية إلى الموضوعية في الاختبار" في المرتبة الأخيرة وبمتوسط حسابي (3)، وقد بلغ المتوسط الحسابي العام لمحور معايير الجودة الخاصة بتصميم الاختبار (3,72)، وهو ما يشير إجمالًا إلى أن درجة توافر معايير الجودة الخاصة بتصميم الاختبار مرتفعة في اختبار تح كفايات اللغة الإنجليزية STEP.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر معايير الجودة الخاصة ببناء الاختبار تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,71) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,24) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر معايير الجودة الخاصة بتصميم الاختبار في اختبار كفايات اللغة الإنجليزية STEP.

ويعزو الباحث هذه النتيجة إلى أن المسؤولين عن الاختبار يتمتعون بكفاءة عالية تؤهلهم لمراعاة أفضل المعايير العالمية في تصميم الاختبار، وذلك تماشيًا مع الأهداف المرجو تحقيقها من خلال تنفيذ الاختبار ولتتمكن النتائج من عكس الصورة الصحيحة لما يمتلكه المتعلمين من كفايات لغوية.

وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة طويل (2017) التي بينت أن مستوى تصميم الاختبارات مناسب بدرجة جيدة وأن أعضاء الهيئة التدريسية يتبعون شروط صياغة المتن في الأسئلة المتعددة مع الحاجة إلى الانتباه إلى الوضوح والدقة اللغوية. وتختلف هذه النتيجة مع نتائج دراسة سويلم (2018) التي أشارت إلى أن الاختبارات التي يقدمها المعلمون غير مناسبة لقياس المهارات التي يمتلكها الطلاب، كما اختلفت مع نتائج دراسة المالكي وفخرو (2017) التي بينت وجود العديد من جوانب القصور في المعايير الخاصة بالاختبارات لمجموعة من المراحل الدراسية، كما اختلفت أيضًا مع نتائج دراسة (Salim et al., 2017) التي أشارت إلى عدم وجود أهداف واضحة لبعض اختبارات اللغة الإنجليزية.

الإجابة عن السؤال الثاني:

ما مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار تحديد المستوى في اللغة الإنجليزية EPT؟

للإجابة عن هذا السؤال تم حساب المتوسط الحسابي والانحراف المعياري لدرجة تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبار تحديد المستوى في اللغة الإنجليزية EPT، والتي تضمنتها قائمة المعايير التي قام الباحث بإعدادها وقد بلغ عددها (7) معايير رئيسة، ويمكن عرض هذه النتائج الخاصة بهذه المعايير من خلال ما يلي:

أولا: مستوى توافر المعايير العامة

جدول 11: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر المعايير العامة

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	تخلو الاختبارات من قضايا التحيز والحساسية	4.00	0.00	100.0%
2	يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية	4.00	0.00	100.0%
3	يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية	4.00	0.00	100.0%
4	يتم تحديد مجال الاختبار بدقة	3.25	0.50	81.3%
5	ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار	2.50	0.58	62.5%
6	تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم	4.00	0.00	100.0%
	المتوسط العام	**3.63**	**0.16**	**90.6%**

يتضح من الجدول رقم (11) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر المعايير العامة**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 62.5%كحد أدنى للعبارة " ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار" إلى 100% لكل من العبارة " تخلو الاختبارات من قضايا التحيز والحساسية"، والعبارة " يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية"، والعبارة " يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية"، والعبارة " تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم" هذا وبلغت درجة التقييم الكلية على هذا المحور 90.6%.

ويبين الجدول رقم (11) أن قيم المتوسطات الحسابية لمستوى توافر المعايير العامة لجودة اختبارات اللغة الإنجليزية العالمية في اختبار تحديد المستوى في اللغة الإنجليزية EPT تتراوح بين (4 – 2,5). وقد جاءت العبارة رقم (1) "تخلو الاختبارات من قضايا التحيز والحساسية" والعبارة رقم (2) "يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية" والعبارة رقم (3) "يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية" والعبارة رقم (6) "تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم" في المرتبة الأولى وبمتوسط حسابي (4,00)، وقد جاءت العبارة رقم (5) "ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار" في المرتبة الأخيرة وبمتوسط حسابي (2,5)، وقد بلغ المتوسط الحسابي العام لمحور مستوى توافر المعايير العامة (3,63)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير العامة مرتفعة في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر المعايير العامة تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,58) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,16) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر المعايير العامة في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

ويعزو الباحث هذه النتيجة إلى أن إعداد اختبار تحديد المستوى في اللغة الإنجليزية EPT يتم من خلال الاستفادة من المعايير العالمية لتصميم اختبارات تحديد المستوى في اللغة الإنجليزية، حيث تؤكد هذه المعايير على أهمية الموضوعية والحيادية وتجنب قضايا التحيز والحساسية في إعداد الاختبارات، إضافة إلى إعداد محتوى الاختبار ليركز على مستوى اكتساب اللغة الإنجليزية واللغويات التطبيقية، وذلك من خلال إشراف مجموعة من العاملين تمثل الخبرات الأساسية لتدريس وتعلم اللغة الإنجليزية وتكوين لجات متخصصة لإدارة الاختبار في كافة مراحله وخاصةً مراحل الإعداد والتنفيذ والتقييم، الأمر الذي يسهم في تحديد مجال هذه الاختبارات بدقة وتحقيق أهدافها، كما يرى الباحث أنه وعلى الرغم من المستوى الجيد لتوافر المعايير العامة، إلا أن العمل على تطوير هذه المعايير يعتبر أمرًا ضروريًا خاصةً فيما يتعلق بعمليات التقييم والرقابة.

وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة الغزال (2020) حيث أكدت أن عملية تصميم الاختبارات التحصيلية بمرحلة الدراسات العليا تتم بصورة واضحة ومناسبة لقياس قدرة الطلبة على ممارسة عمليات عقلية متنوعة.

ثانيا: مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها

جدول 12: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	يتمتع أعضاء اللجنة بخبرة عميقة في المجالات ذات الصلة بالمعايير أو التقييمات أو كليهما	4.00	0.00	100.0%
2	يعكس أعضاء اللجنة التوازن المناسب في التمثيل ووجهات النظر	4.00	0.00	100.0%
3	أعضاء اللجنة موجهين بشكل مناسب ويتم تدريبهم	4.00	0.00	100.0%
4	يتوفر لأعضاء اللجنة الوقت الكافي لتنفيذ المهمة (المهام) المطلوب منهم إكمالها	4.00	0.00	100.0%
	المتوسط العام	**4.00**	**0.00**	**100.0%**

يتضح من الجدول رقم (12) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها**"، حيث نالت جميع عبارات هذا المعيار على درجة تقييم 100%.

ويبين الجدول رقم (12) أن قيم المتوسطات الحسابية لمستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها جميعها متساوية حيث بلغت قيمتها (4,00) وقد بلغ المتوسط الحسابي العام لمحور مستوى توافر المعايير العامة (4,00)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير العامة مرتفعة في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

كما يتضح من الجدول السابق عدم وجود انحراف معياري لمستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها حيث إن قيمة الانحراف المعياري لجميع العبارات كانت (0,00)، كما بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,00) وتشير هذه القيمة إلى التطابق في استجابات أفراد عينة الدراسة حول مستوى توافر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

ويعزو الباحث هذه النتيجة إلى التحول الكبير الذي يشهده قطاع التعليم في المملكة العربية السعودية والسعي المستمر إلى تطويره من مختلف جوانبه، وهو ما أسهم وبدرجة كبيرة في زيادة مستوى كفاءة المشاركين في إعداد وتصميم اختبار تحديد المستوى في اللغة الإنجليزية EPT، وامتلاكهم لخبرة عالية في مختلف المجالات الخاصة بالمعايير العالمية لإعداد اختبارات اللغة الإنجليزية وتطبيقها في الاختبار في عمليات الاختبار المختلفة، كما تدل هذه النتيجة أيضًا إلى أن المسؤولين

عن اختبار تحديد المستوى في اللغة الإنجليزية EPT يدركون جيدًا أهمية الاختبار ويسعون إلى توفير كافة العوامل التي من شأنها أن تسهم في إنجاحه وتقديمه بصورته المناسبة للطلاب من خلال استقطاب الموارد البشرية المناسبة وتوفير كافة الإمكانات التي يحتاجونها.

وتختلف هذه النتيجة مع ما توصلت إليه نتائج الجار (2020) والتي توصلت إلى أن مستوى ممارسات أعضاء الهيئات التدريسية في إعداد الاختبارات الفصلية جاء بدرجة متوسطة مما يدلل على ضرورة تطوير أعضاء الهيئات التدريسية بما يتماشى مع المعايير الحديثة لتصميم الاختبارات، كما اختلفت أيضًا مع دراسة البلوي (2019) التي أشارت نتائجها إلى أن درجة التزام المعلمات بمهارات التخطيط وتصحيح الاختبارات التحصيلية جاء بدرجة متوسطة.

ثالثا: مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية

جدول 13: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	يعتبر الاختبار التنوع بين خصائص متعلمي اللغة الإنجليزية	4.00	0.00	100.0%
2	إشراك الخبراء الذين لديهم دراية بنطاق التنوع الذي يمثله طلاب دارسي الإنجليزية	3.50	0.58	87.5%
3	هناك معايير مختلفة أو توقعات إنجاز لمختلف مجموعات دارسي الإنجليزية الفرعية	3.50	0.58	87.5%
4	يتم تطوير مجموعة واحدة من معايير الاختبارات الصارمة لجميع طلاب دارسي الإنجليزية	3.50	0.58	87.5%
	المتوسط العام	**3.63**	**0.14**	**90.6%**

يتضح من الجدول رقم (13) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 87.5% كحد أدنى لكل من العبارات " إشراك الخبراء الذين لديهم دراية بنطاق التنوع الذي يمثله طلاب دارسي الإنجليزية"، والعبارة " هناك معايير مختلفة أو توقعات إنجاز لمختلف مجموعات دارسي الإنجليزية الفرعية"، والعبارة " يتم تطوير مجموعة واحدة من معايير الاختبارات الصارمة لجميع طلاب دارسي الإنجليزية" إلى 100% للعبارة " يعتبر الاختبار التنوع بين خصائص متعلمي اللغة الإنجليزية"، هذا وبلغت درجة التقييم الكلية على هذا المحور 90.6%.

ويبين الجدول رقم (13) أن قيم المتوسطات الحسابية لمستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية في اختبار تحديد المستوى في اللغة الإنجليزية EPT تتراوح بين (4 – 3,5). وقد جاءت العبارة رقم (1) "يعتبر الاختبار التنوع بين خصائص متعلمي اللغة الإنجليزية" في المرتبة الأولى وبمتوسط حسابي (4,00)، بينما جاءت العبارة رقم (2) " إشراك الخبراء الذين لديهم دراية بنطاق التنوع الذي يمثله طلاب دارسي الإنجليزية " والعبارة رقم (3) " إشراك معايير مختلفة أو توقعات إنجاز لمختلف مجموعات دارسي الإنجليزية الفرعية" والعبارة رقم (4) "يتم تطوير مجموعة واحدة من معايير الاختبارات الصارمة لجميع طلاب دارسي الإنجليزية" في المرتبة الأخيرة وبمتوسط حسابي (3,5)، وقد بلغ المتوسط الحسابي العام لمحور توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية (3,63)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية مرتفعة في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,58) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,14) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

ويعزو الباحث هذه النتيجة إلى الاهتمام بتصميم اختبار تحديد المستوى في اللغة الإنجليزية EPT ليتوافق مع خصائص جميع متعلمي اللغة الإنجليزية، وهي من المعايير المهمة لتصميم الاختبارات عمومًا والاختبارات اللغوية خصوصًا، كما تشير هذه النتيجة إلى أن عملية إعداد الاختبار تتم من خلال مشاركة مجموعة من الخبراء المختصين في تصميم الاختبارات اللغوية، مما يزيد من كفاءة هذا الاختبار ويتيح لجميع الطلاب المشاركة العادلة والحصول على نتائج جيدة.

وتتفق هذه النتيجة مع النتائج التي توصلت إليها دراسة أبو عيش (2017) والتي أشارت إلى التوجهات الإيجابية للطلبة للمشاركة في الاختبارات وخاصةً تلك التي تتناسب مع توجهاتهم وخصائصهم، كما اتفقت مع نتائج دراسة (Salim et al., 2017) التي أكدت على أهمية مراعاة خصائص الطلاب في الاختبارات وعمليات التقييم، بينما اختلفت مع نتائج دراسة (Alsamaani, 2014) التي بينت أن معلمي اللغة الإنجليزية كلغة أجنبية في المدارس الإعدادية المبتدئين بحاجة إلى تطوير معارفهم ومهاراتهم في مجموعة واسعة من تقييمات متعلمي اللغة الإنجليزية كلغة أجنبية، وذلك ليتمكنوا من تقييم أداء الطلاب ونتائج تعلمهم للغة الإنجليزية كلغة ثانية وفقًا لخصائصهم.

رابعا: مستوى توافر المعايير الخاصة بمجالات الاختبار

جدول 14: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر المعايير الخاصة بمجالات الاختبار

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	يفحص الاختبار القواعد النظرية والبحثية والتنمية الاجتماعية والثقافية وعلم اللغة الاجتماعي	3.00	0.00	75.0%
2	يقيم الاختبار اتساع وعمق ومدى التعقيد والتعبير عن المهارات والمعرفة المرتبطة بها	3.50	0.58	87.5%
3	تحديد المهارات والمعرفة اللغوية المطلوبة في الاستماع والتحدث والقراءة والكتابة	4.00	0.00	100.0%
4	يركز الاختبار على الترابط بين الطرق الأربعة والمساهمة في إتقان اللغة الإنجليزية	3.50	0.58	87.5%
	المتوسط العام	**3.50**	**0.29**	**87.5%**

يتضح من الجدول رقم (14) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر المعايير الخاصة بمجالات الاختبار**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 75%كحد أدني للعبارة " يفحص الاختبار القواعد النظرية والبحثية والتنمية الاجتماعية والثقافية وعلم اللغة الاجتماعي" إلى 100% للعبارة " تحديد المهارات والمعرفة اللغوية المطلوبة في الاستماع والتحدث والقراءة والكتابة"، هذا وبلغت درجة التقييم الكلية على هذا المحور 87.5%.

ويبين الجدول رقم (14) أن قيم المتوسطات الحسابية لمستوى توافر المعايير الخاصة بمجالات الاختبار في اختبار تحديد المستوى في اللغة الإنجليزية EPT تتراوح بين (4 – 3). وقد جاءت العبارة رقم (3) "تحديد المهارات والمعرفة اللغوية المطلوبة في الاستماع والتحدث والقراءة والكتابة" في المرتبة الأولى وبمتوسط حسابي (4,00)، بينما جاءت العبارة رقم (2) " يقيم الاختبار اتساع وعمق ومدى التعقيد والتعبير عن المهارات والمعرفة المرتبطة بها" والعبارة رقم (4) " يركز الاختبار على الترابط بين الطرق الأربعة والمساهمة في إتقان اللغة الإنجليزية" في المرتبة الثانية وبمتوسط حسابي (3,5)، وجاءت العبارة رقم (1) "يفحص الاختبار القواعد النظرية والبحثية والتنمية الاجتماعية والثقافية وعلم اللغة الاجتماعي" في المرتبة الأخيرة وبمتوسط حسابي (3)، وقد بلغ المتوسط الحسابي العام لمحور توافر المعايير الخاصة بمتعلمي اللغة الإنجليزية (3,5)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير الخاصة بمجالات الاختبار مرتفعة في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر المعايير الخاصة بمجالات الاختبار تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,58) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,29) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر المعايير الخاصة بمجالات الاختبار في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

ويعزو الباحث هذه النتيجة إلى مستوى الاهتمام المتعلق بتقييم امتلاك المتعلمين المعرفة اللغوية المختلفة والتي ترتبط بالمهارات اللغوية الأساسية المتمثلة في الاستماع والتحدث والقراءة والكتابة، ويرجع ذلك إلى أن إتقان هذه المهارات يعتبر من أهم أهداف تعليم اللغة الإنجليزية، إضافةً إلى التعرف على المستوى العام لإتقان اللغة الإنجليزية لدى المتعلمين، وأن التصميم الأساسي للاختبار يدور حول معرفة مدى إتقان الطلاب لهذه المهارات.

وتختلف هذه النتيجة مع ما توصلت إليه نتائج دراسة (Alsamaani, 2014) التي وجدت أن الاختبارات التي يقوم بها المعلمون لا تفي بمعايير تطوير اختبار تحصيل الطلاب، وأن المعلمين بحاجة إلى تطوير معارفهم ومهاراتهم في مجموعة واسعة من تقييمات متعلمي اللغة الإنجليزية كلغة أجنبية، ودراسة (Khora, 2021) التي أشارت إلى أن اختبارات اللغة الإنجليزية لا تغطي جميع المهارات اللغوية، حيث تركز الاختبارات على مهارات القراءة والكتابة وتستبعد مهارات التحدث والاستماع.

خامسا: مستوى توافر المعايير الخاصة بالسياق المجتمعي

جدول 15: المتوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر المعايير الخاصة بالسياق المجتمعي

م	العبارات	المتوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	عند تصميم الاختبار يراعى الإلمام بالسكان، السياسات والممارسات والبرامج السابقة والحالية	4.00	0.00	100.0%
2	يتم تطوير هيكل الاختبار بناء على السياق التاريخي لتعليم وتعلم اللغة الإنجليزية	4.00	0.00	100.0%
	المتوسط العام	**4.00**	**0.00**	**100.0%**

يتضح من الجدول رقم (15) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **مستوى توافر المعايير الخاصة بالسياق المجتمعي**"، حيث نالت جميع عبارات هذا المعيار على درجة تقييم 100%.

ويبين الجدول رقم (15) أن قيم المتوسطات الحسابية لمحور مستوى توافر المعايير الخاصة بالسياق المجتمعي في اختبار تحديد المستوى في اللغة الإنجليزية EPT جاءت متساوية وبقيمة قدرها (4,00). وبذلك فقد بلغ المتوسط الحسابي العام لمحور توافر المعايير الخاصة بالسياق المجتمعي (4,00)، وهو ما يشير إجمالًا إلى أن درجة توافر المعايير الخاصة بالسياق المجتمعي مرتفعة في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

كما يتضح من الجدول السابق أن الانحرافات للمتوسط الحسابي العام (0,00) وهي قيمة تشير إلى التطابق التام في استجابات أفراد عينة الدراسة حول مستوى توافر المعايير الخاصة بالسياق المجتمعي في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

ويعزو الباحث هذه النتيجة إلى أن اختبار تحديد المستوى في اللغة الإنجليزية EPT يتم تصميمه وبناؤه بناءً على الواقع العملي لتعليم وتعلم اللغة الإنجليزية في المملكة، وهو ما يجعل منه امتحانًا واقعيًا لقياس مستوى امتلاك المتعلمين لمهارات اللغة الإنجليزية المختلفة ومستوى إتقانهم لها، الأمر الذي يزيد من أهمية النتائج التي يحصل عليها المتعلمون في هذا الاختبار.

سادسا: معايير الجودة الخاصة ببناء الاختبار

جدول 16: الوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر معايير الجودة الخاصة ببناء الاختبار

م	العبارات	الوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	الصدق	4.00	0.00	100.0%
2	الثبات	3.50	0.58	87.5%
3	الصياغة الواضحة	4.00	0.00	100.0%
4	التنوع	4.00	0.00	100.0%
5	الدقة	4.00	0.00	100.0%
6	مصداقية	4.00	0.00	100.0%
7	قابلة التطبيق	4.00	0.00	100.0%
8	عدالة التوزيع	3.50	0.58	87.5%
9	سهولة الاشتراك	3.50	0.58	87.5%
10	الشمول	4.00	0.00	100.0%
11	التكامل بين مستويات المعرفة	3.50	0.58	87.5%
	المتوسط العام	**3.82**	**0.21**	**95.5%**

يتضح من الجدول رقم (16) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **معايير الجودة الخاصة ببناء الاختبار**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 87.5%كحد أدني لكل من عبارات "الثبات"، "عدالة التوزيع"، "سهولة الاشتراك"، "التكامل بين مستويات المعرفية" إلى 100% لكل من العبارات " الصدق"، "الصياغة الواضحة"، "التنوع"، "الدقة"، " المصداقية"، "قابلية التطبيق"، "الشمول"، هذا وبلغت درجة التقييم الكلية على هذا المحور 95.5%.

ويبين الجدول رقم (16) أن قيم المتوسطات الحسابية لمستوى توافر معايير الجودة الخاصة ببناء الاختبار في اختبار تحديد المستوى في اللغة الإنجليزية EPT تتراوح بين (4 – 3,5). وقد جاءت العبارة رقم (1) " الصدق" والعبارة رقم (3) " الصياغة الواضحة" والعبارة رقم (4) " التنوع" والعبارة رقم (5) " الدقة" والعبارة رقم (6) " مصداقية" والعبارة رقم (7) " قابلة التطبيق" والعبارة رقم (10) " الشمول" في المرتبة الأولى وبمتوسط حسابي (4,00)، بينما جاءت العبارة رقم (2) " الثبات" والعبارة رقم (8) "عدالة التوزيع" والعبارة رقم (9) "سهولة الاشتراك" والعبارة رقم (11) "التكامل بين مستويات المعرفة" في المرتبة الثانية وبمتوسط حسابي (3,5)، وقد بلغ المتوسط الحسابي العام لمحور معايير الجودة الخاصة ببناء الاختبار (3,82)، وهو ما يشير إجمالًا إلى أن درجة توافر معايير الجودة الخاصة ببناء الاختبار مرتفعة في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر معايير الجودة الخاصة ببناء الاختبار تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,58) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,21) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر معايير الجودة الخاصة ببناء الاختبار في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

ويعزو الباحث هذه النتيجة إلى وجود خبرة علمية ومهنية لدى المسؤولين عن الاختبار والقائمين عن بنائه وهو ما يجعله اختبارًا يتوافق مع المعايير العالمية لاختبارات اللغة الإنجليزية الدولية، بحيث تكون شاملة ومناسبة لقياس المهارات والمفاهيم التي وضعت لقياسها.

وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة أبيش وعنشيل (2018) التي أشارت إلى أن معظم أسئلة الاختبارات موضوعة بطريقة واضحة ومباشرة وموجهة بصورة أساسية إلى الفئة الضعيفة والتي يركز عليها المعلمون أكثر من غيرها، وذلك كونها تساعدهم في تحصيل أكبر قدر ممكن من النقاط. وتختلف هذه النتيجة مع ما توصلت إليه نتائج دراسة سويلم (2018) التي أشارت إلى أن الاختبارات التي يتم تصميمها لا تتناسب مع المهارات التي يمتلكها الطلاب وغير مناسبة لقياس العديد من المهارات المرتبطة بالتفكير.

سابعا: معايير الجودة الخاصة بتصميم الاختبار

جدول 17: الوسط الحسابي والانحراف المعياري والوزن النسبي لعبارات مستوى توافر معايير الجودة الخاصة بتصميم الاختبار

م	العبارات	الوسط الحسابي	الانحراف المعياري	الوزن النسبي
1	تحديد الأهداف اللغوية التي يقيسها الاختبار	3.50	0.58	87.5%
2	تحديد الكفايات اللغوية التي يقيسها الاختبار	4.00	0.00	100.0%
3	تحديد العناصر اللغوية التي يقيسها الاختبار	4.00	0.00	100.0%
4	تحديد الكفايات اللغوية التي سيغطيها الاختبار	3.50	0.58	87.5%
5	تحديد العناصر اللغوية التي سيغطيها الاختبار	3.50	0.58	87.5%
6	تحديد المهارات اللغوية التي يستهدف الاختبار قياسها	4.00	0.00	100.0%
7	تحديد نسبة الأسئلة المقالية إلى الموضوعية في الاختبار	3.00	0.00	75.0%
8	تحديد عدد البدائل الخاصة بالأسئلة الموضوعية	4.00	0.00	100.0%
9	تحديد آلية الإجابة على أسئلة الاختبار	4.00	0.00	100.0%
	المتوسط العام	**3.72**	**0.19**	**93.1%**

يتضح من الجدول رقم (17) المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " **معايير الجودة الخاصة بتصميم الاختبار**"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 75%كحد أدني للعبارة " تحديد نسبة الأسئلة المقالية إلى الموضوعية في الاختبار" إلى 100% لكل من العبارات " تحديد الكفايات اللغوية التي يقيسها الاختبار"، " تحديد العناصر اللغوية التي يقيسها الاختبار"، " تحديد المهارات اللغوية التي يستهدف الاختبار قياسها"، " تحديد عدد البدائل الخاصة بالأسئلة الموضوعية"، " تحديد آلية الإجابة عن أسئلة الاختبار"، هذا وبلغت درجة التقييم الكلية على هذا المحور 93.1%.

ويبين الجدول رقم (17) أن قيم المتوسطات الحسابية لمستوى توافر معايير الجودة الخاصة ببناء الاختبار في اختبار تحديد المستوى في اللغة الإنجليزية EPT تتراوح بين (4 – 3). وقد جاءت العبارة رقم (2) " تحديد الكفايات اللغوية التي يقيسها الاختبار" والعبارة رقم (3) " تحديد العناصر اللغوية التي يقيسها الاختبار" والعبارة رقم (6) " تحديد المهارات اللغوية التي يستهدف الاختبار قياسها" والعبارة رقم (8) " تحديد عدد البدائل الخاصة بالأسئلة الموضوعية" والعبارة رقم (9) " تحديد آلية الإجابة عن أسئلة الاختبار" في المرتبة الأولى وبمتوسط حسابي (4,00)، بينما جاءت العبارة رقم (1) " تحديد الأهداف اللغوية التي يقيسها الاختبار" والعبارة رقم (4) "تحديد الكفايات اللغوية التي سيغطيها الاختبار" والعبارة رقم (5) "تحديد العناصر اللغوية التي سيغطيها الاختبار" في المرتبة الثانية وبمتوسط حسابي (3,5)، وقد جاءت العبارة رقم (7) "تحديد نسبة الأسئلة المقالية إلى الموضوعية في الاختبار" في المرتبة الأخيرة وبمتوسط حسابي (3)، وقد بلغ

المتوسط الحسابي العام لمحور معايير الجودة الخاصة بتصميم الاختبار (3,72)، وهو ما يشير إجمالًا إلى أن درجة توافر معايير الجودة الخاصة بتصميم الاختبار مرتفعة في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

كما يتضح من الجدول السابق أن الانحرافات المعيارية لمستوى توافر معايير الجودة الخاصة ببناء الاختبار تضمنت قيمًا صغيرة، حيث تراوحت بين (0,00 – 0,58) وهو ما يدل على وجود تجانس عالٍ في استجابات أفراد عينة الدراسة على عبارات هذا المحور، وقد بلغت قيمة الانحراف المعياري للمتوسط الحسابي العام (0,19) وهي قيمة صغيرة تشير إلى التقارب في استجابات أفراد عينة الدراسة حول مستوى توافر معايير الجودة الخاصة بتصميم الاختبار في اختبار تحديد المستوى في اللغة الإنجليزية EPT.

ويعزو الباحث هذه النتيجة إلى أن المسؤولين عن الاختبار يتمتعون بكفاءة عالية تؤهلهم لمراعاة أفضل المعايير العالمية في تصميم الاختبار، وذلك تماشيًا مع الأهداف المرجو تحقيقها من خلال تنفيذ الاختبار ولتتمكن النتائج من عكس الصورة الصحيحة لما يمتلكه المتعلمين من مهارات لغوية.

وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة طويل (2017) التي بينت أن مستوى تصميم الاختبارات مناسب بدرجة جيدة وأن أعضاء الهيئة التدريسية يتبعون شروط صياغة المتن في الأسئلة المتعددة مع الحاجة إلى الانتباه إلى الوضوح والدقة اللغوية. وتختلف هذه النتيجة مع نتائج دراسة سويلم (2018) التي أشارت إلى أن الاختبارات التي يقدمها المعلمون غير مناسبة لقياس المهارات التي يمتلكها الطلاب، كما اختلفت مع نتائج دراسة المالكي وفخرو (2017) التي بينت وجود العديد من جوانب القصور في المعايير الخاصة بالاختبارات لمجموعة من المراحل الدراسية، كما اختلفت أيضًا مع نتائج دراسة (Salim et al., 2017) التي أشارت إلى عدم وجود أهداف واضحة لبعض اختبارات اللغة الإنجليزية.

الإجابة عن السؤال الرئيسي للدراسة :

ما مدى تحقق معايير جودة اختبارات اللغة الإنجليزية العالمية في اختبارات هيئة تقويم التعليم والتدريب اللغوية؟"

يتضح من خلال الإجابة عن السؤال الأول أن المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " مستوى توافر المعايير العامة"، حيث قام كلا من الباحث والمشرف بالتقييم، وذلك لسهولة التواصل بين الباحث والمشرف، وتم الأخذ بعين الاعتبار نموذجين لكل اختبار، وتراوحت مستويات التقييم على هذا المحور ما بين 63%كحد أدني للعبارة " ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار" إلى 100% لكل من العبارة " تخلو الاختبارات من قضايا التحيز والحساسية"، والعبارة " يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية"، والعبارة " يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية"، والعبارة " تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم" هذا وبلغت درجة التقييم الكلية على هذا المحور 90%.

يتضح من خلال الإجابة عن السؤال الأول أن المتوسطات الحسابية والانحرافات المعيارية والأوزان النسبية لتقييم الباحث والمشرف التربوي على محور " مستوى توافر المعايير العامة"، حيث تراوحت مستويات التقييم على هذا المحور ما بين 62.5%كحد أدني للعبارة " ضمان الجودة في التقييم والأعمال الرقابية لتنفيذ الاختبار" إلى 100% لكل من العبارة " تخلو الاختبارات من قضايا التحيز والحساسية"، والعبارة " يقوم على إعداد معايير الاختبارات مجموعات تمثل الخبرات الأساسية"، والعبارة " يركز الاختبار على اكتساب اللغة الإنجليزية واللغويات التطبيقية"، والعبارة " تكوين لجان متخصصة بإعداد المعايير وتنفيذ الاختبار وتطوير التقييم" هذا وبلغت درجة التقييم الكلية على هذا المحور 90.6%.

ويشير تحقق الأجزاء إلى أن الكل المتكون منها متحقق بالضرورة وفقًا للمعايير المحددة، وبذلك يتضح تحقق جودة اختبارات اللغة الإنجليزية العالمية في اختبارات هيئة تقويم التعليم والتدريب اللغوية.

ملخص نتائج الدراسة، التوصيات والمقترحات

تمهيد:

يتناول هذا الفصل ملخصًا للنتائج التي توصلت إليها الدراسة الحالية، ومن ثم تقديم التوصيات والمقترحات، وذلك على النحو التالي:

أولًا: ملخص النتائج

بعد أن قام الباحث بتطبيق إجراءات الدراسة توصل الباحث إلى مجموعة من النتائج التي تم استعراضها بصورة مفصلة في الفصل السابق، ويمكن تلخيص النتائج التي توصلت إليها الدراسة فيما يلي:

1. تتوفر المعايير العامة لجودة اختبارات اللغة الإنجليزية العالمية بدرجة كبيرة في كل من اختبار كفايات اللغة الإنجليزية STEP واختبار تحديد المستوى في اللغة الإنجليزية EPT، وذلك بمتوسط حسابي (3,63) ودرجة تقييم كلية 90% لاختبار كفايات اللغة الإنجليزية STEP، ومتوسط حسابي (3,63) ودرجة تقييم كلية 90.6% لاختبار تحديد المستوى في اللغة الإنجليزية EPT.
2. تتوفر معايير تكوين لجان إعداد معايير الاختبارات وتقييمها والإشراف عليها بدرجة كبيرة في كل من اختبار كفايات اللغة الإنجليزية STEP واختبار تحديد المستوى في اللغة الإنجليزية EPT، وذلك بمتوسط حسابي (4,00) ودرجة تقييم كلية 100% لاختبار كفايات اللغة الإنجليزية STEP، ومتوسط حسابي (4,00) ودرجة تقييم كلية 100% لاختبار تحديد المستوى في اللغة الإنجليزية EPT.
3. تتوفر المعايير الخاصة بمتعلمي اللغة الإنجليزية بدرجة كبيرة في كل من اختبار كفايات اللغة الإنجليزية STEP واختبار تحديد المستوى في اللغة الإنجليزية EPT، وذلك بمتوسط حسابي (3,63) ودرجة تقييم كلية 91% لاختبار كفايات اللغة الإنجليزية STEP، ومتوسط حسابي (3,63) ودرجة تقييم كلية 90,6% لاختبار تحديد المستوى في اللغة الإنجليزية EPT.
4. تتوفر المعايير الخاصة بمجالات الاختبار بدرجة كبيرة في كل من اختبار كفايات اللغة الإنجليزية STEP واختبار تحديد المستوى في اللغة الإنجليزية EPT، وذلك بمتوسط حسابي (3,5) ودرجة تقييم كلية 88% لاختبار كفايات اللغة الإنجليزية STEP، ومتوسط حسابي (3,5) ودرجة تقييم كلية 87.5% لاختبار تحديد المستوى في اللغة الإنجليزية EPT.
5. تتوفر المعايير الخاصة بالسياق المجتمعي بدرجة كبيرة في كل من اختبار كفايات اللغة الإنجليزية STEP واختبار تحديد المستوى في اللغة الإنجليزية EPT، وذلك بمتوسط حسابي (4,00) ودرجة تقييم كلية 100% لاختبار كفايات اللغة الإنجليزية STEP، ومتوسط حسابي (4,00) ودرجة تقييم كلية 100% لاختبار تحديد المستوى في اللغة الإنجليزية EPT.
6. تتوفر معايير الجودة الخاصة ببناء الاختبار بدرجة كبيرة في كل من اختبار كفايات اللغة الإنجليزية STEP واختبار تحديد المستوى في اللغة الإنجليزية EPT، وذلك بمتوسط حسابي (3,82) ودرجة تقييم كلية 95% لاختبار كفايات اللغة الإنجليزية STEP، ومتوسط حسابي (3,82) ودرجة تقييم كلية 95.5% لاختبار تحديد المستوى في اللغة الإنجليزية EPT.
7. تتوفر معايير الجودة الخاصة بتصميم الاختبار بدرجة كبيرة في كل من اختبار كفايات اللغة الإنجليزية STEP واختبار تحديد المستوى في اللغة الإنجليزية EPT، وذلك بمتوسط حسابي (3,72) ودرجة تقييم كلية 93% لاختبار كفايات اللغة الإنجليزية STEP، ومتوسط حسابي (3,72) ودرجة تقييم كلية 93.1% لاختبار تحديد المستوى في اللغة الإنجليزية EPT.

ثانيًا: التوصيات

في ضوء النتائج التي توصلت إليها الدراسة الحالية يوصي الباحث بما يلي:

1. استمرار متابعة أعمال القائمين على تصميم وبناء كل من اختبار كفايات اللغة الإنجليزية STEP واختبار تحديد المستوى في اللغة الإنجليزية EPT، وتدريبهم وتطويرهم من أجل الحفاظ على أفضل مستوى ممكن من معايير الجودة فيه.
2. العمل على التحديث المستمر لمعايير الجودة التي يتم الاعتماد عليها في تصميم وبناء وتنفيذ الاختبارات لتتوافق مع المعايير الدولية وما يطرأ عليها من تطورات وتغيرات.
3. تطوير تدريس اللغة الإنجليزية للمرحلة الابتدائية وحتى المرحلة الجامعية في المملكة العربية السعودية لتتوافق أهدافه ومخرجاته مع المتوقع من المتعلمين من امتلاك المهارات والكفايات اللغوية الأساسية للغة الإنجليزية.
4. الاستفادة من التقنيات التعليمية الحديثة في تصميم وتنفيذ الاختبارات، مع العمل على تحويلها إلى اختبارات محوسبة.
5. التأكيد على أن الاختبار متضمن لجميع المهارات والمعارف والكفايات التي يجب أن يمتلكها متعلمي اللغة الإنجليزية وبما يتوافق مع الاختبارات الدولية.

ثالثًا: المقترحات

في ضوء النتائج والتوصيات السابقة يقترح الباحث الدراسات التالية:

1. دراسة تتناول معايير الجودة التي يتم تطبيقها في اختبارات اللغة العربية لغير الناطقين بها.
2. دراسة حول التحديات التي تواجه المتعلمين في التحضير لاختبارات اللغة الإنجليزية واجتيازها بنجاح.
3. دراسة حول أثر جودة تعليم اللغة الإنجليزية في مدارس التعليم العام في المملكة على نتائج المتعلمين في اختبارات اللغة الإنجليزية المحلية والعالمية.

أولا: المراجع العربية

أبو جامع، ليندا. (2021م). تقويم أسئلة اختبارات الأحياء للصف الثاني عشر في ضوء عمليات العلم خلال السنوات من 2018-2020م. *مجلة الجامعة الإسلامية للدراسات التربوية والنفسية، 29* (4)، 78-106.

أبو عيش، بسينة رشاد علي. (2017). اتجاهات المتقدمات نحو الاختبارات المحوسبة التابعة للمركز الوطني للقياس والتقويم بمركز قياس للاختبارات المحوسبة بالطائف. *مجلة كلية التربية*، كلية التربية، جامعة بنها، *28* (109)، 451- 496.

أبيش، فتحية وعنيشل، خديجة. (2018). *تقويم أسئلة البناء اللغوي في اختبارات اللغة العربية: مرحلة التعليم المتوسط السنة الثانية عينة: دراسة وصفية تحليلية*. رسالة ماجستير، كلية الآداب واللغات، جامعة قاصدي مرباح، ورقلة، الجزائر.

أوعلا، بوشري محمد والمطارنة، أحمد جبريل عثمان. (2018). الأداء التفاضلي لفقرات اختبار تحديد المستوى في مادة اللغة الإنجليزية المطبق على طلبة جامعة مؤتة. *مجلة العلوم التربوية والنفسية*، مركز النشر العلمي، جامعة البحرين، *19* (2)، 449- 475.

البلوي، سماح. (2019) حول تقويم الاختبارات المدرسية التي تعدها معلمات العلوم في المرحلة الأساسية بمدينة تبوك من حيث تخطيطها وإخراجها وتصحيحها. *مجلة العلوم التربوية والنفسية، 3* (25)، 48-26.

البوعبيد، عبد الله. (2016). مدى تحقيق اختبار القدرات العامة للمرحلة الثانوية لمبدأ تكافؤ الفرص التعليمية من وجهة نظر طلاب السنة التحضيرية بجامعة الملك فيصل. *مجلة كلية التربية*، كلية التربية، جامعة بنها، *27*(107)، 107-139.

التوبية، منى بنت سعيد بن صالح ولاشين، محمد عبد الحميد وحمد، عمر هاشم إسماعيل. (2021). التحديات التي تواجه مديري المدارس بسلطنة عمان في تحسين مستوى أداء الطلبة في الاختبارات الدولية "TIMSS" وآليات التغلب عليها. *المجلة الدولية للدراسات التربوية والنفسية*، مركز رفاد للدراسات والأبحاث، *9* (1)، 279- 292.

الجار، ثريا. (2020). واقع الاختبارات الفصلية في بعض كليات جامعة شقراء كما تراها الطالبات في ضوء بعض المتغيرات "دراسة تقويمية". *مجلة البحوث التربوية والنفسية*، *17* (65)، 106-141

جبارة، كوثر. (2017). مدى التزام المحاضرات بالقواعد الأساسية في بناء الاختبارات الموضوعية في كليتي التربية والآداب في جامعة حائل. *المجلة العربية لضمان جودة التعليم الجامعي، 10* (32)، 55-75.

خاجي، ثاني حسين. (2013). إسهامات استراتيجيات التقويم البديل (Evaluation Alternative) في تحسين جودة التعميم. *مجلة كلية التربية الأساسية*، جامعة بابل، ع (14)، 352- 363.

الخروصي، حسين بن علي والريامي، حمد بن جمعة والحوسني، سالم بن خلفان. (2021). معايير جودة التقويم التربوي: مراجعة تحليلية نظرية. *المجلة العربية للعلوم التربوية والنفسية*، المؤسسة العربية للتربية والعلوم والآداب، ع (21)، 145- 159.

خليفة، عبد السلام الشيباني. (2014). الاتجاهات المعاصرة في التقويم التربوي ودورها في تطوير العملية التعليمية. *فكر وإبداع*، رابط الأدب الحديث، ج (86)، 485- 506.

طاشكندى، سارة. (2018). اللغة الإنجليزية أحد التحديات التي تواجه تعليم المرأة السعودية. *المؤتمر السنوي العربي الثالث عشر – الدولي العاشر: التعليم العالي النوعي في مصر والوطن العربي في ضوء إستراتيجيات التنمية المستدامة، 1* (1)، 115-136.

الفلاحات، بلسم. (2020). *درجة فهم معلمي اللغة الإنجليزية وممارستهم لمعيار التقويم بوصفه أحد معايير الاتحاد العالمي لتعليم اللغة الإنجليزية لغة أجنبية*. رسالة ماجستير، كلية الدراسات العليا، جامعة العلوم الإسلامية العالمية، الأردن.

الرباط، بهيرة. (2015). *المناهج الدراسية رؤية استشرافية*. الرياض: دار الزهراء.

زكي، أسامة. (2016). *الاختبارات اللغوية: مقاربة منهجية تطبيقية*. الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية.

سويلم، همام. (2018) تقويم اختبارات طلاب الشعب العلمية بكلية التربية بجامعة المنيا في ضوء معايير جودة الاختبارات. *مجلة البحث في التربية وعلم النفس، 33* (3)، 272-304

صالح، ثريا حسن محمد والإمام، آمال موسى عباس. (2020). معايير تقويم الكفاءة اللغوية في اختبارات اللغة العربية للناطقين بغيرها. المؤتمر الدولي للغة العربية وآدابها 2020. *إثراء المعرفة للمؤتمرات والأبحاث*، 109- 116.

طعيمة، رشدي، والشعيبي، محمد. (2006). *تعليم القراءة والأدب*. القاهرة: جار الفكر العربي.

طويل، علاء. (2017). تقويم الاختبارات الموضوعية في كلية التمريض – جامعة تشرين. *مجلة جامعة تشرين للبحوث والدراسات العلمية، 39* (1)، 35-51

عبد المؤمن، علي معمر. (2008). *البحث في العلوم الاجتماعية*. منشورات جامعة 7 أكتوبر إدارة المطبوعات والنشر، ليبيا.

السيد، أسامة. (2016). *الاختبارات اللغوية مقاربة منهجية تطبيقية*. مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية.

العدل، عبد الرحمن السيد. (2018). تقييم صدق وثبات اختبار القبول وقياس إدراك الطلاب: نموذج جامعة الدلتا للعلوم. *مجلة كلية التربية في العلوم التربوية*، كلية التربية، جامعة عين شمس، *42*(1)، 14- 40.

العساف، صالح. (2012). *المدخل إلى البحث في العلوم السلوكية*. ط2. الرياض: دار الزهراء.

الغزال، محمد. (2020). تقويم الاختبارات التحصيلية بمرحلة الدراسات العليا بقسم اللغة العربية وآدابها بكلية الآداب جامعة مصراتة في ضوء معايير الجودة الشاملة، *مجلة كلية الآداب جامعة مصراته، 15* ، 31-55

المالكي، بدرية سعيد وفخرو، عائشة أحمد. (2017). دراسة تحليلية لمدى توافق الاختبارات الوطنية مع معايير الاختبار الجيد للمدارس المستقلة في المرحلة الابتدائية بدولة قطر. *مجلة دراسات الخليج والجزيرة العربية، مجلس النشر العلمي*، جامعة الكويت، س (43)، ع (166)، 21- 61.

المركز الوطني للقياس. (2021). *اختبار اللغة العربية المقنّن لغير الناطقين بها*. الرياض: المركز الوطني للقياس.

هيئة تقويم التعليم والتدريب. (2022). الاختبارات اللغوية. مسترجع بتاريخ 2022/03/25 من الرابط التالي: https://etec.gov.sa/AR/PRODUCTSANDSERVICES/QIYAS/LINGUAL/Pages/default.aspx.

وحشه، دلال عبد الله الطلب (2022). دور استراتيجية العصف الذهني في رفع مستوى أداء الطلبة في الاختبارات الدولية من وجهة نظر مديري المدارس الحكومية في محافظة عجلون. *مجلة العلوم التربوية والنفسية، 6* (4)، 92- 107.

ثانيا: المراجع الأجنبية

Abdel-Gawad, A.S.(2019). *The Effect of Flipped Classroom Instruction on Developing EFL Writing Skills of Secondary Stage Students*. master thesis, faculty of Education: Helwan university.

Al-Adawi, S. S., & Al-Balushi, A. A. (2016). Investigating Content and Face Validity of English Language Placement Test Designed by Colleges of Applied Sciences. *English Language Teaching*, *9*(1), 107-121.

Ali, O. A., & Ahmed, M. F. (2021). Calendar Criteria For Selection Of The Educational Supervisors In Secondary Schools From The Point Of View Of Teachers. *Academy Of Entrepreneurship Journal, 27*(3), 1-19.

Alsamaani, Abdulaziz. (2014). Evaluating Classroom Assessment Techniques of Novice Saudi EFL Teachers. *Journal of Arabic and Human Sciences, 7* (2), 63- 81.

Alshwaier, A., Youssef, A., & Emam, A. (2012). A new trend for e-learning in KSA using educational clouds. *Advanced Computing, 3*(1), 81.

Ambag, S. C. (2014). Adaptation of the College of Education Students and Faculty Regarding Academic Calendar Shift. *UNP Research Journal, 23*, 99-112.

Banse, H. W., Palacios, N. A., Merritt, E. G., & Rimm-Kaufman, S. E. (2017). Scaffolding English language learners' mathematical talk in the context of Calendar Math. *The Journal of Educational Research, 110*(2), 199-208.

Belfield, C. R., & Crosta, P. M. (2012). Predicting Success in College: The Importance of Placement Tests and High School Transcripts. *CCRC Working*. Paper No. 42. Community College Research Center, Columbia University.

David, R. J. (2013). *National Alliance for secondary education and transition.* Transition toolkit for systems improvement, Institute on community Integration University of Minnesota.

Durrant, J. (2014). *The Reliability and Validity of the Ten Test and Exploring a New Visual Version*. Master thesis, McMa ter University, Hamilton, Ontario.

Fang, F. G. (2017). ENGLISH AS A LINGUA FRANCA: IMPLICATIONS FOR ï¿ ¼PEDAGOGY AND ASSESSMENT. *Teflin Journal, 28*(1), 57-70.

Guo, S. C. (2012). Using Authentic Materials for Extensive Reading to Promote English Proficiency. *English Language Teaching, 5*(8), 196-206.

Helda, H., & Syahrani, S. (2022). National Standards of Education in Contents Standards and Education Process Standards in Indonesia. *Indonesian Journal of Education (INJOE), 3*(2), 257-269.

Khora, Haneen Mohammed (2021). Evaluation of English for Palestine Textbooks and English Language Examination In Light of Common European Framework of References for Languages "CEFRL". *Journal of Educational and Psychological Sciences, 5* (24), 152- 175.

Kling, J., & Stæhr, L. S. (2012). *The development of the test of oral English proficiency for academic staff (TOEPAS).* University of Copenhagen.

Lei, J., & Hu, G. (2014). Is English-medium instruction effective in improving Chinese undergraduate students' English competence?.. *International Review of Applied Linguistics in Language Teaching, 52*(2), 99-126.

Lin, Y. M., & Y. Chen, M. (2020). Understanding writing quality change: a longitudinal study of repeaters of a high-stakes standardized English proficiency test. *Language Testing, 37*(4), 523-549.

Maxwell, T. W. (2012). Assessment in Higher Education in the Professions: Action Research as an Autherntic Assesmetn Tast. *Teaching in Higher Education,17*(6), 686-696. Retrived from: https://eric.ed.gov/?id=EJ991135

Murray, A. T. (2016). Assessing the impacts of traditional school year calendar start dates. *Socio-Economic Planning Sciences, 54*, 28-36.

Salehi, M., Nasirzadeh, A., Niazkar, H. R., & Moghaddam, K. B. (2019). Designing an Annual Educational Calendar of Medical Parasitology Based on Training Weeks. *Disease and Diagnosis, 8*(2), 90-94.

Salim, D. M. A. B., Mohamed, I. M. A., & Abulgasim, Y. S. H. (2017). The Impacts of Testing in Improving English Language Learners Proficiency. *West Kordofan University Journal of Sciences and Humanities*, West Kordofan University, No. (14), 49- 76.

Sawyer Jr, D. T. (2018). *Balanced calendar versus traditional calendar: Measuring the difference in reading academic achievement among kindergarten through third-grade students who have been identified as reading below or well-below grade level*. University of Missouri-Kansas City.

Setiawan, S. A. (2021). *Developing Listening Exercise for English Proficiency Test in The Form of Mobile Application for Eleventh Grade Students of SMA Alizzah Batu*. University Of Islam Malang Graduateprogramenglish Language Teaching Studyprogram, 1- 12.

Shelly, S.A(2015). *Effectiveness of Extensive Reading in Second Language Learning: A Study on the Development of Linguistic Skills in Reference to Reading*. doctoral dissertation, Department of English and Humanities Of BRAC University.

Sidky, G.(2019). The Flipped Classroom Enhances Students' Speaking Skills. *Journal of Studies in Curriculum and Teaching Methods,243*,22-49.

Sims, J. M. (2015). A valid and reliable English proficiency exam: A model from a university language program in Taiwan. *English as a Global Language Education (EaGLE) Journal, 2*(1), 91-93.

Villafuerte, M. Q., Juanillo, E. L., & Hilario, F. D. (2017). Climatic insights on academic calendar shift in the Philippines. *Philippine Journal of Science, 146*(3), 267-276.

Yan, X. (2014). An examination of rater performance on a local oral English proficiency test: A mixed-methods approach. *Language Testing, 31*(4), 501-527.

Zainal Abidin, S. A., & Jamil, A. (2015). Toward an English proficiency test for postgraduates in Malaysia. *SAGE Open, 5*(3), 2158244015597725