

# اختبار اللغة العربية المقنن لغير الناطقين بها:

بناؤه وآلية إنتاجه





بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



# المحتويات

٦	مقدمة
٧	دراسة الحاجة
٨	عمل اللجنة المكلفة بالإعداد للاختبار:
٨	- مسح الأدبيات
٨	- اختيار إطار الاختبار المرجعي
٩	- تعريف بالإطار الأوروبي المشترك للغات
١٢	- قرار اللجنة
١٢	طبيعة الاختبار
١٣	هدف الاختبار
١٣	مكوّنات الاختبار ونسبها
١٣	وصف المكونات:
١٣	- فهم المقروء
١٣	- الكتابة
١٥	- فهم المسموع
١٥	- التحدّث
١٥	إنتاج الاختبار
١٦	كتابة أسئلة الاختبار
١٧	تحكيم الأسئلة
١٧	اللغة المستهدفة
١٧	الإجراءات الميدانية ودراسات الصدق والثبات:
١٧	- التحليل الإحصائي
١٨	- المنهج:
١٨	• اختيار العينة
١٨	• الأداة
١٨	- التحليل والدراسات
٢٠	- الدراسات التي أجراها المركز على الاختبار
٢١	- الهوامش

## • مقدمة:

اللغة العربية اليوم بحاجة مأساة إلى برنامج تنمية متكامل. ولا بد للجهات المعنيّة بقضاياها أن تجمع الجهود لتنميتها، كل جهة في مجال اختصاصها. وقد جمع المركز الوطني للقياس، وهو هيئة علمية تربوية غير ربحية مستقلة متخصصة في تطوير المقاييس العلمية والمهنية، قد جمع عزمه على تطوير مقياس لخدمة اللغة العربية في مجال اختصاصه باسم «اختبار اللغة العربية المقتن لغير الناطقين بها». وقد درج المركز على إصدار مواد علمية تُعرّف بالمقاييس التي ينتجها، قد تكون هذه المواد مطبوعة أو إلكترونية، وقد تكون نشرات أو رسائل أو كتيبات أو أدلة. وكل مستفيد من مقاييس المركز سيجد ما يناسب حاجته منها، وهذا الإصدار هو أحدها، ويحتوي على البناء النظري للاختبار من حيث دراسة الحاجة إليه، والإعداد له، وطبيعته وهدفه ومكوناته. ويحتوي -أيضاً- على آلية إنتاج الاختبار من حيث كتابة أسئلته وتحكيمها ولغته المستهدفة، وإجراءاته الميدانية، ودراسات صدقه وثباته.



## • دراسة الحاجة:

اجتمع عدد من الدواعي والأسباب للنظر في الحاجة إلى وضع اختبار مقنن يقيس كفايات متعلمي اللغة العربية من غير أبنائها، منها:

١. حاجة اللغة العربية الماسة إلى برنامج تنمية لغوية، يشمل طرائق اكتساب اللغة، وطرائق تعليمها وتعلمها، وتطوير مناهجها ومقرراتها المبنية على المدونات اللغوية وقوائم المفردات الشائعة، وتحديد الكفايات التي ينبغي لمتعلميها أن يتقنوها، والمهارات التي تساعدهم على إتقان هذه الكفايات، والأدوات العلمية العالمية ذات المنهجية والمرجعية والمصدقية التي تقيس مستويات المتعلمين في هذه الكفايات أو تبين مدى إتقانهم للكفايات التي اكتسبوها<sup>(١)</sup>.

٢. وجود الجهات المتخصصة من الجامعات والمعاهد والمراكز ذات القدرة على الشروع في عمل متضافر ومتكامل يُعنى بوضع إطار مرجعي شامل ينظر إلى تعليم اللغة وتعلمها وقياس درجة إتقان ذلك على أنها عمليات يكمل بعضها بعضاً.

٣. حرص المركز على أن يشارك في مجال اختصاصه خدمةً للغة العربية ببناء اختبار يقيس مستوى كفايات متعلميها.

٤. حاجة الجامعات وتقدم بعضها بطلب إلى المركز لتطوير اختبار مقنن يقيس كفايات متعلمي اللغة العربية من غير أهلها يمكن استخدامه معياراً لقبول الطلاب غير العرب في برامج علمية تكون اللغة العربية هي وسيلة التدريس فيها.

٥. حاجة كثير من الناس والجهات المهنية إلى مؤشر يُحتكم إليه لمعرفة مستوى إتقان أفراد معينين للمهارات الأساسية في اللغة العربية.

٦. تنامي الإقبال على دراسة اللغة العربية عالمياً؛ لما للبيئة العربية من تأثير في المجالات العلمية والثقافية والتجارية والسياحية.

٧. توحيد الجهود المتفرقة في مجال اختبارات كفايات اللغة العربية بتوحيد المنهج وأسلوب التطبيق، مع الاستفادة من خبرات جميع المتخصصين.

ولذا كَوّن المركز لجنة علمية لدراسة الحاجة إلى اختبار مقنن يقيس مستوى متعلمي اللغة العربية من غير الناطقين بها في المهارات اللغوية المنشودة، وكان أعضاؤها من المتخصصين في المجالات الآتية: (١) القياس التربوي (٢) الاختبارات اللغوية (٣) اللغة العربية (٤) علم اللغة التطبيقي (٥) تعليم اللغة العربية لغير الناطقين بها. وكانت مهماتها: (١) الوقوف على التجارب السابقة في الاختبارات اللغوية لمتعلمي اللغات الأخرى، (٢) الوقوف على التجارب السابقة في وضع اختبارات لمتعلمي اللغة العربية من غير أهلها، (٣) تحديد الأطر والمعايير العالمية للاختبارات اللغوية المحددة للكفايات اللغوية وكيفية قياسها، (٤) الوصول إلى قرار بشأن الحاجة إلى اختبار مقنن يقيس مستوى كفاية متعلمي اللغة العربية من غير أهلها في المهارات المنشودة.

## • عمل اللجنة المكلفة بالإعداد للاختبار:

### - مسح الأدبيات:

ليس بمستغرب على اللجنة - وهي الجهة المتخصصة - أن تبدأ بالوقوف على الأطر والمعايير التي تعلم أن الاختبارات المقتننة عالمياً تبني عليها، بل وتتخذها مرجعاً. ولا شك أن أهم المرجعيات التي ارتأت اللجنة أن تنظر فيها بعمق هما:

(١) معايير المجلس الأمريكي<sup>(١)</sup>:

(ACTFL) the ACTFL Proficiency Guidelines 2012

(٢) الإطار الأوروبي المشترك للغات<sup>(٢)</sup>:

(CEFR) Common European Framework of reference for Languages, 2009

وقد بلغت التجارب السابقة في الاختبارات اللغوية لمتعلمي اللغات الأخرى شأواً عظيماً لم تحتج اللجنة إلى أن تقف عنده كثيراً، غير أن من المفيد تلخيص بحثها في النقطتين التاليتين:

(١) أن اللغات العالمية مخدومة باختبارات مطوّرة ومقتننة، ولنأخذ على ذلك مثلاً اللغة الإنجليزية، ففي الولايات المتحدة الأمريكية طوّرت هيئة خدمات الاختبارات التربوية (ETS) اختبار (توفل) المعروف عالمياً:

Test of English as a Foreign Language (TOEFL) ولجامعة (كامبردج) البريطانية قسم متخصص في الاختبارات اللغوية المقتننة يعرف باسم (ESOL)، وقد طوّر اختبار (آيلتس):

International English Language Testing System (IELTS) الذي لا يقل شأنه عالمياً عن سابقه الأمريكي، وطوّرت مؤسسة بيرسون (PEARSON) عددًا من اختبارات اللغة الإنجليزية .

(٢) أما اللغة العربية - وإن كانت إحدى اللغات العالمية - فلم تحظ بما حظيت به قريناتها من العناية بالاختبارات المقتننة. نعم، هناك تجارب بعضها جيد، لكنها لا تكفي.

### - اختيار إطار الاختبار المرجعي:

اللغة في الاستعمال البشري ظاهرة شديدة التعقيد، تتطلب عددًا كبيراً من المهارات / الكفايات المختلفة. ومن المهم في كل اختبار يعتمد قياس الكفايات اللغوية أن يبدأ ببناء أنموذج واضح لهذه الكفايات ولعلاقة بعضها ببعض؛ لأن مهمة هذا الأنموذج هي تعرّف السمات الرئيسة للكفاية اللغوية لدى الإنسان، فهو نقطة ابتداء لمعرفة الكفاية التي يريد الاختبار قياسها، والمساعدة في التأكد من أن نتيجة الاختبار مفيدة ويمكن تفسيرها. والكفاية ليست سمة سهلة القياس، غير أننا محظوظون بوجود نماذج جيدة عند البحث في هذه المسألة.





انظر مثلاً: (Canale & Swain, 1981) و (Weir, C. J. 1990 و Bachmann, L. F. 2005)، ولا يشترط في هذا النموذج أن ينطلق من زعم أو دعوى قوية عن كيفية انتظام وبناء هذه الكفايات في عقل الإنسان.

وبعد مقارنة عدد من الأطر في مجال قياس الكفايات اللغوية ارتأى المركز أن يهتدي بالنموذج الإطار الأوروبي المشترك وفلسفته، ويسعى المركز جاهداً إلى تطبيق معاييرها بتفصيلاتها واستكمال عمليات الربط.

## - تعريف بالإطار الأوروبي المشترك للغات:

CEFR (Common European Framework of reference for Languages)

وُضع هذا الإطار بإشراف مجلس الاتحاد الأوروبي (Council of Europe) وله ستة مستويات مبنية على مبدأ: ما الذي يستطيع المتعلم إنجازه باللغة في مستواه؟

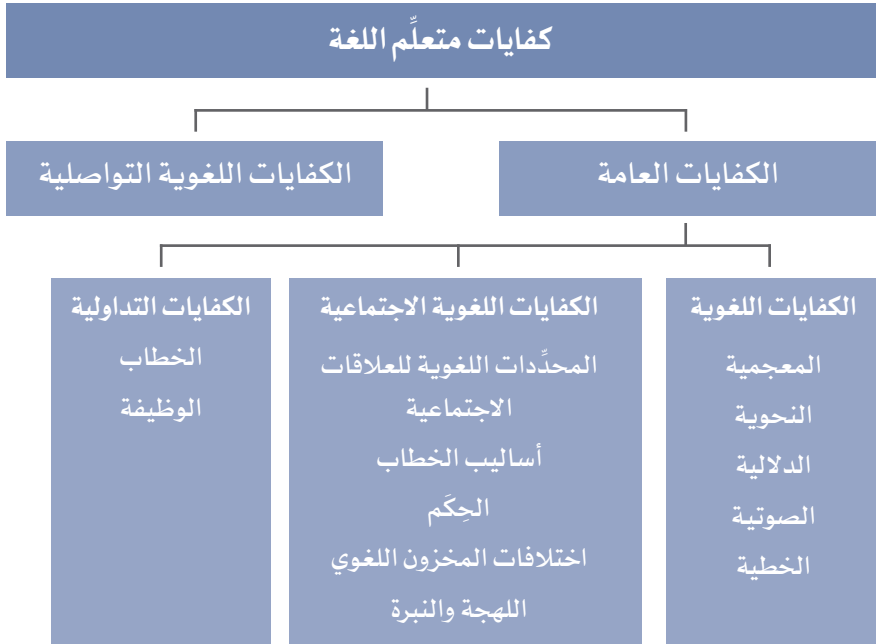
“can do statements”

هذه المستويات هي: A1, A2, B1, B2, C1, C2 وتسمى (global scale)، وهي بالترتيب من اليسار إلى اليمين (A1) ويرمز إلى أدنى مستويات المتعلم في الكفايات اللغوية و (C2) الذي يرمز إلى أعلى مستويات المتعلم في الكفايات اللغوية. وتقدم المستويات وصفاً دقيقاً لما يستطيع المتعلم فعله بلغته في مستواه.

أما أنموذجه في تشخيص السمات الأساسية للكفاية اللغوية فتُعبّر عنه الفقرة الآتية<sup>(٤)</sup>:

\* تتطور أفعال الإنسان الكلامية إلى كفايات عامة، وإلى كفايات تواصلية على وجه الخصوص. ويُعوّل المتكلم على الكفايات التي يملكها في سياقات معيّنة وتحت قيود محدّدة ليشترك في نشاطات معيّنة فيها عمليات لغوية؛ لكي ينشئ أو يتلقّى نصوصاً معيّنة في موضوعات ومجالات حياتية مختلفة. وهو بهذا يُوظّف أدواته وآلياته اللغوية التي تتناسب مع مهمات يحتاج إلى إنجازها. وتؤدي مراقبة المتكلم لأفعاله الكلامية إلى تعزيز كفاياته اللغوية أو تعديلها.

\* تُعرّف هذه الفقرة بالعناصر الرئيسة لأنموذج الإطار الأوروبي، والكلمات المغلّطة فيها تُعبّر عن مفاهيمه المفتاحية التي تُبيّن بناء الهرمي، كما في الشكلين الآتيين:



## C2 المستخدم البارع:

يفهم بسهولة كل شيء يسمعه أو يقرؤه؛ فيستطيع مثلاً أن يلخص معلومات مأخوذة من مصادر مكتوبة أو منطوقة، ويستطيع أن يعيد بناء الأحداث والحجج في عرض متماسك: يستطيع المستخدم البارع في هذا المستوى أن يعبر عن نفسه بارتجال وعضوية، وبطلاقة ودقة مع القدرة على إدراك الفروق الدلالية بين الكلمات والأساليب حتى في الأحوال الدقيقة والمعقدة.

## C1 المستخدم البارع:

يفهم بسهولة عدداً كبيراً من النصوص الصعبة والطويلة ويفهم المعاني الضمنية فيها: يستطيع المستخدم البارع في هذا المستوى أن يعبر عن نفسه بارتجال وعضوية، دون أن يلاحظ عليه تلكؤ واضح في البحث عن الكلمات والتعبير. كما يستطيع أن يستخدم اللغة بطلاقة وفاعلية في المجالات الاجتماعية والأكاديمية والمهنية، ويستطيع كذلك أن ينتج نصوصاً واضحة ومتينة الصياغة ومفصلة في موضوعات معقدة، ويظهر قدرة على التحكم في استعمال قوالب الكتابة التنظيمية وأدوات الربط.

## B2 المستخدم القادر:

يفهم الأفكار الرئيسية في النصوص الصعبة ذات الموضوعات المجردة وغير المجردة، وهذا يشمل المناقشات الفنية في مجال الاختصاص: يستطيع المستخدم القادر في هذا المستوى أن يتخاطب بقدر من الطلاقة والعضوية مع غيره، ويمكنه بهذا القدر من اللغة أن يتخاطب تخاطباً منتظماً مع أبناء اللغة المتعلمة دون كبير مشقة لدى الطرفين. كما يستطيع كذلك أن ينتج نصوصاً واضحة وتفصيلية في مجالات متعددة من الموضوعات، ويبسط وجهة نظره في قضية ما، ويذكر المنافع والمضار لاختيارات متنوعة.

## B1 المستخدم القادر:

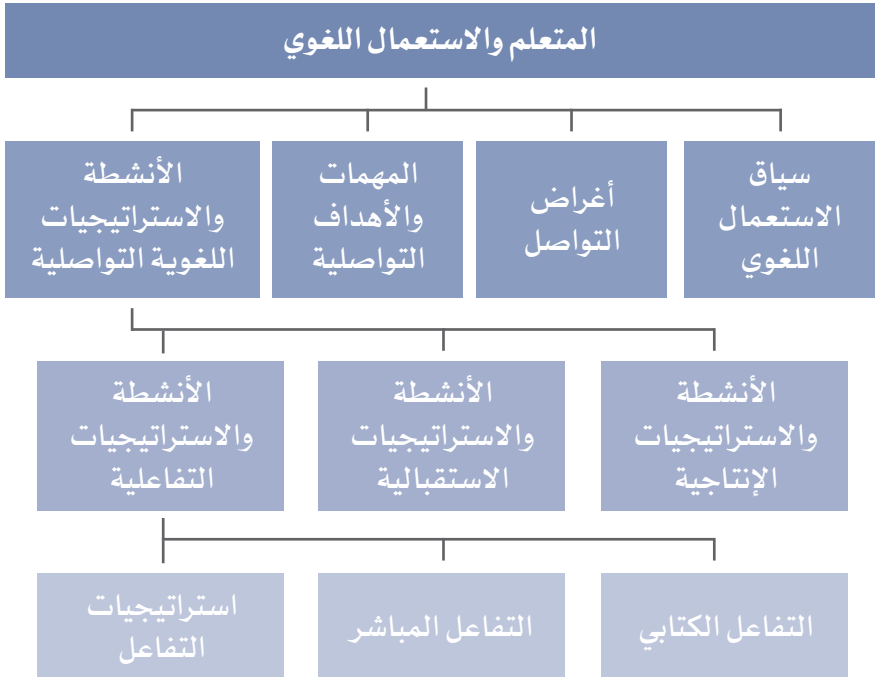
يفهم النقاط الرئيسية في نصوص مألوفة ذات موضوعات متكررة في العمل والمدرسة ومكان النزهة وزمانها... : يستطيع المستخدم القادر في هذا المستوى أن يتعامل مع معظم الحالات التي يمكن أن تطرأ في مكان وهو يتجول في أي مكان من أماكن اللغة المتعلمة. كما أنه يستطيع أن ينتج نصوصاً سهلة متصلة في مجالات من الموضوعات المألوفة أو الممتعة له شخصياً. ويستطيع كذلك أن يصف بعض التجارب والأحداث والأمنيات والطموحات، ويقدم باختزال الأسباب والتفسيرات للأراء والخطط.

## A2 المستخدم المبتدئ:

يفهم الجمل والعبارات المتكررة التي لها علاقة باحتياجاته القريبة جداً، كالمعلومات الشخصية والعائلية الأساسية والتسوق والمعلومات الجغرافية المحلية والوظيفة: يستطيع المستخدم المبتدئ في هذا المستوى أن يتخاطب في المهمات البسيطة والروتينية التي تتطلب تبادل معلومات حول قضايا مألوفة ونمطية. كما يستطيع كذلك أن يصف بلغة بسيطة بعض سمات سيرته ومحيطه المباشر والموضوعات التي تقع ضمن احتياجاته المباشرة.

## A1 المستخدم المبتدئ:

يفهم ويستخدم العبارات المألوفة والمتكررة يومياً، والتعبيرات الأساسية التي يقصد بها تحقيق احتياجاته المباشرة والمحسوسة: يستطيع المستخدم المبتدئ في هذا المستوى أن يقدم نفسه للآخرين، ويستطيع أن يسأل ويحجب عن الأسئلة المتعلقة بالتفاصيل الشخصية، مثل مكان السكنى والأشخاص الذين يعرفهم والأشياء التي يمتلكها. كما يستطيع أن يتواصل مع الآخرين بلغة بسيطة بشرط أن يتكلموا ببطء ووضوح ولديهم استعداد لمساعدته.



#### - قرار اللجنة:

كان قرار اللجنة بعد أن استعرضت التجارب السابقة أن هناك حاجة لاختبار يقيس جميع كفايات اللغة العربية لدى غير الناطقين بها تتوافر فيه الخصائص الآتية:

١. أن يكون اختباراً شاملاً يقيس جميع كفايات اللغة العربية لدى غير الناطقين بها، التي من أهمها أن المهارات اللغوية تمثل كائناً واحداً وهو اللغة التي لا ينبغي في الظروف الاعتيادية أن تطغى فيها مهارة على أخرى وأن الكفاية تقاس فيها منطقياً وطبيعياً على حد سواء، وهذا ما لا يتوافر في مجمل التجارب السابقة، فمنها ما يختبر المنطوق دون غيره ومنها ما يختبر المقروء دون غيره.
٢. أن يكون عالمياً، وبهذا يكون متناسباً مع مكانة اللغة العربية بوصفها إحدى اللغات المتداولة في المحافل العالمية، كما يتناسب مع عراقتها التاريخية وراثتها اللغوية والأدبي والحضاري. والتجارب السابقة تجارب لصيقة بمؤسسات علمية محدّدة، وبعضها موضوع لتقديم خدمات مجتمعية محلية.
٣. أن يتولى تطويره ومتابعته وتطبيقه وصيانته ودراسة ثباته وصدق بنائه هيئة متخصصة علمياً وإدارياً ولوجستياً؛ ذلك أن الاختبارات المقننة لم تعد اليوم عملاً تستطيع أية هيئة أن تجعله ضمن مهماتها الفرعية؛ لما لنتائجها من خطورة على الأفراد والمؤسسات العلمية والمهنية والتعليم برُمته والمجتمع كله؛ ولهذا كله حدّدت اللجنة اسم الاختبار، وطبيعته، وهدفه، ومكوّناته، وأوزانها المقترحة على النحو الآتي:

#### • طبيعة الاختبار:

هو اختبار مقنن لقياس كفايات المختبرين في فهم المسموع، وفهم المقروء، وسلامة التراكيب النحوية، والكتابة الحرة. وقد رُوّعت في مراحل تطويره الشروط والمعايير المتعارف عليها عالمياً في صناعة المقاييس والاختبارات المقننة.



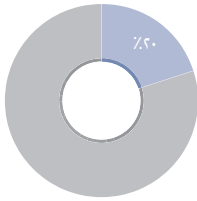
## • هدف الاختبار:

يهدف إلى تقويم مستوى إتقان اللغة العربية لدى أشخاص من غير الناطقين بها؛ يُمكنهم من الالتحاق بمؤسسات أكاديمية تُقدّم برامجها باللغة العربية.

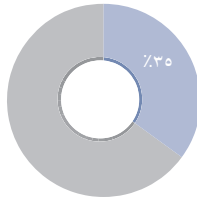
## • مكوّنات الاختبار ونسبها:

تختلف مكوّنات الاختبار ونسبها تبعاً لهدفه، فإذا كان يراد منه اختبار كفايات معيّنة في مهارات محدّدة، حُدّدت المكوّنات ونسبها بحسب هدف الاختبار. والأصل أن يُبنى اختبار عام يقيس الكفايات العامة في المهارات اللغوية الأربع المتعارف عليها، وهي: فهم المقروء، وفهم المسموع، والكتابة، والتحدّث بنسبٍ متساوية. غير أنه عند بناء هذا الاختبار كان لدى اللجنة العلمية اعتباران:

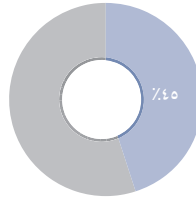
1. أن الطلب المعروض أمامها مقدّم من هيئات علمية تحتاجه لأغراض القبول في برامجها العلمية الأكاديمية،
2. أن بعض المهارات اللغوية تحتاج في قياسها قياساً موضوعياً وعلمياً إلى احتياجات خاصة إدارية ومنهجية ولوجستية، فُحصرت مكوّنات الاختبار في الكفايات الآتية:



فهم المسموع



التراكيب اللغوية  
والتعبير الكتابي



فهم المقروء

## • وصف المكوّنات:

### - فهم المقروء:

تُبنى أسئلة استيعاب المقروء على نصوص حقيقية ذات أطوال مختلفة. فالطويل منها يكون عدد مفرداته في حدود (٣٠٠) كلمة، ويقسّم عادة إلى فقرات مرقمة يستعان بأرقامها عند السؤال عنها لتجنب المفضوح قراءة النص كله عند الإجابة عن السؤال الواحد وهذا يضيع عليه كثيراً من الوقت. والمتوسط منها يكون في فقرتين مرقمتين أيضاً، والقصير منها يكون في فقرة واحدة أو حتى في جملة طويلة أو جملتين إحداهما مرتبة على الأخرى من حيث المعنى. وتغطي الأسئلة مجالات معرفية متعدّدة، منها: البحث عن معلومة، أو تحديد عنوان لفقرة أو للنص كله، أو تحديد الفكرة الرئيسة لفقرة أو للنص كله، أو تحديد معنى كلمة أو تركيب حسب مقتضى النص، أو البحث عن علاقات بين الجمل أو الفقرات كالسببية أو التمثيل أو الشرح والتوضيح. ومنها: البحث عن الربط المنطقي الناشئ عن إحالات الضمائر أو الإشارة أو الصلة، والاستنتاج والاستدلال المنطقي أو توقع نتيجة لنص غير مكتمل.

### - الكتابة:

تتكون أسئلة الكتابة من نوعين: (١) نوع (الاختيار من متعدد) وفيه يجب المفضوح عن أسئلة تركيبية تسأل عن سلامة الاستعمال اللغوي لبعض العبارات أو الجمل القصيرة؛ ولذا تكثر فيها أسئلة المطابقة والمخالفة في التعريف والتنكير والعدد والتذكير والتأنيث والإعراب. ويسأل فيها عن ترتيب الكلام في الجملة وعن الإسناد والإضافة والإتباع والتراكيب العددية. وتقيس هذه البنود عمليات عقلية معيّنة كالفهم والتطبيق. (٢) أما النوع الثاني فيتطلب كتابةً مقاليةً لا يقل عدد كلماتها

عن (٢٠٠) كلمة فيها رأي أو فكرة أو قضية جدلية. ويطلب من المختبر في هذا النوع أن يتخذ رأياً معيناً، موافقاً أو معارضاً، في قضية من القضايا مع الحرص في كل الأحوال على تأييد رأيه بأمثلة واضحة وأفكار تفصيلية محددة وحجج وبراهين مقنعة. وكان من أوجه التحدي الحقيقية في هذا الجزء المهم من قياس مهارة الكتابة ضمان اتساق أسلوب تقييمه، سواء أتم ذلك ألياً أم بشرياً لضمان عدالة الدرجة المعطاة. وإلى أن يُطوّر محرّك حاسوبي يُطمأن إلى تصحيحه النصوص العربية تصحيحاً دقيقاً يوازي على الأقل المحاولات الدولية لخدمة لغة عالمية كاللغة الإنجليزية، فإن المركز استثمر طاقاته في تطوير التصحيح البشري. فأعدّ أولاً مسودة لمعايير (rubrics) وأبعاد (dimensions) يمكن أن تطور إلى أن تصل إلى درجة يمكن الاحتكام إلى نتائجها باطمئنان، ثم عقّد ورشة عمل دعا إليها (١٢) أستاذاً جامعياً منهم المتخصص في علوم اللغة العربية، ومنهم المتخصص في علم اللغة التطبيقي، ومنهم المتخصص في تعليم اللغة العربية لغير الناطقين بها. وبدأت الورشة بعرض لأجزاء الاختبار وضرورة ضمان عدالة درجته بما في ذلك ما سيتدربون عليه، ثم أعطي المتدربون في أول تمرين من تمارين الورشة ورقة من أوراق الطلاب الذين أخذوا هذا الاختبار « اختبار اللغة العربية المقنن لغير الناطقين بها»، وطلب من كل أستاذ تصحيحها انفرادياً على الطريقة التي يصحّح بها كل أستاذ أوراق طلابه الحقيقيين في الفصل. وبعد اكتمال التصحيح عُرضت أمامهم الدرجات التي أعطها كل أستاذ لهذه الورقة. وهنا أدرك الأساتذة تفاوت الدرجات مع اجتماع كلمتهم على معظم الأخطاء التي ارتكبتها الطالب، وهنا أيضاً استرعى انتباههم إلى أن التخصص وحده لا يكفي، وأن كون الأستاذ يدرّس لسنوات طوال لا يكفي، وإنما الحاجة ماسة لتطوير معايير تصحيح دقيقة، وهذا وحده كذلك لا يكفي، إذ لا بد من جمع كلمة المصححين دائماً على هذه المعايير كما سيتبين من التمرين الذي يليه. بدأ عرض المعايير وتطويرها بالتعديل أو الدمج أو التغيير أو الإضافة. ومع كل مرحلة كانت تعطى ورقة واحدة ليصححها الأساتذة مع السماح لهم بالتشاور الجماعي والعمل المشترك، ثم قُسموا كذلك إلى أربع مجموعات، وكانت كل مجموعة تعطى ورقة للتصحيح المشترك داخل اللجنة، وبعد الانتهاء منها تعرض درجة كل مجموعة، وكان يُنظر دائماً إلى نقاط الاجتماع والتفاوت، فما اجتمعوا عليه عُزّرت معاييره وأُبقيت، وما تفاوتوا فيه أُعيد النظر في معاييره فعدّل أو استُبدل. وهكذا إلى أن أكملوا سبع ورقات. ثم وُزعت ثلاث ورقات للتصحيح الانفرادي ونُظر في تقييم كل مصحّح على شاشة العرض العامة أمام جميع المشاركين فيعرف كل مصحّح مدى قربه وبعده من المصححين الآخرين. وهكذا اكتملت الورشة على مدى أربعة أيام ثم أُخترت عشوائياً خمسون ورقة ووزعت على المصححين الاثني عشر، بمعنى أن كل مصحّح يصحّح انفرادياً حزمة الأوراق الخمسين، وفعلاً صحّحت الأوراق وفق المعايير والأبعاد المقررة في ورشة العمل، ثم أعطيت كل حزمة بعد تصحيح المصحّح الأول لمصحّح آخر؛ لدراسة ثبات معايير التصحيح لدى المصحّحين جميعاً، ثم أُجريت عليها دراسة ثبات المصححين وارتباطهم ببعض؛ فتراوح معدل الارتباط في الورشة الأولى بين (٠.٦) و (٠.٧). ولم تكن هذه النتيجة تمثّل ما يطمح إليه المركز، فقرر عقد ورشة أخرى دعا لها (١٢) مرشحاً من مدرسي اللغة العربية في التعليم العام وتعليم اللغة العربية لغير الناطقين بها، لكن أحد المرشحين اعتذر قبل انعقاد الورشة بيوم واحد فاستقر العدد على (١١) مرشحاً. وعقدت ورشة عمل ثانية واتبعت فيها خطوات الورشة الأولى نفسها مع بعض التعديلات والإضافات في معايير التصحيح. وبعد دراسة الارتباط بين المصححين في الورشة الثانية ارتفعت درجته فتراوحت بين (٠.٩٢) و (٠.٩٥)؛ وعندها أُعتمدت درجة التعبير الكتابي، وضمّت إلى الدرجة الكلية للمختبر. ويُقدّم في هذا القسم (قسم الكتابة) خمسة وأربعون سؤالاً في النموذج الواحد للاختبار الواحد، وتسعون سؤالاً في النماذج الخمسة للاختبار، هذا عدا سؤال التعبير الحر.



## - فهم المسموع:

يُقدّم للمفحوص أنواع من التسجيلات الصوتية القصيرة جدًا والقصيرة والمتوسطة، وقليلًا ما يُسأل عن مقاطع صوتية طويلة. ومن هذه التسجيلات ما هو مفرد (monologue) ومنها ما هو حوار (dialogue). فأما التسجيل المفرد فقد يكون إعلانًا أو تصريحًا أو جزءًا من نشرة أخبار أو مقطعًا من محاضرة، وأما الحوار فتداول كلامي مباشر بين شخصين في مكتب أو سوق أو مطعم، أو حوار هاتفي أو غير ذلك. ويُسمح للمفحوص أن يأخذ رؤوس أقلام عن أرقام أو تواريخ أو أماكن أو أزمنة محدّدة، مع تنبيهه إلى أن ذلك لا ينبغي أن يكون على حساب سماع المقطع الصوتي سماعًا كاملاً. وتوجّه إليه أسئلة في هذه المقاطع عن عمليات عقلية أولية كالسؤال عن مكان أو تاريخ محدّد أو وصف اتجاه. أما السؤال عن عمليات عقلية عالية كالتحليل والتركيب فيقول في هذا النوع. وتُسجّل هذه المقاطع الصوتية في مختبرات عالية التقنية لضمان الجودة، وما يُسجّل منها لأول مرة يقرؤه متخصصون في اللغة العربية من أبناؤها وتتوافر في أصواتهم فنيًا أعلى خصائص الوضوح، سواء أكانوا من الذكور أم من الإناث. وقد تطمّن التسجيلات ببعض المؤثرات الصوتية كجرس الهاتف أو الباب أو صوت محرك سيارة أو صوت باب يفتح أو يغلق، أو إقلاع طائرة أو هبوطها، أو سماع إنذار حريق أو طوارئ، وهكذا. ويُقدّم في هذا القسم واحد وأربعون سؤالاً في النماذج الخمسة للاختبار؛ لأن هذا القسم يُقدّم لجميع المفحوصين في وقت واحد.

## - التحدّث:

هذه المهارة هي إحدى مهارات اللغة الأربع الأساسية، لكن يُلاحظ في توزيع المكونات ونسبها أن هذه المهارة لم تُدرج ضمن المهارات اللغوية، ليس لأنها غير مهمة، أو أنها ليست موضع اهتمام المركز، بل العكس هو الصحيح، وهذا التوزيع مؤقّت إلى أن يضمن المركز عدالة الدرجة التي تُمنح عند تقويم هذه المهارة ولا يحصل هذا إلا بالتأكد من مناسبة طريقة الاختبار ونوع الأسئلة والأهم من ذلك كله طريقة التقويم، كما فعل بالضبط مع مهارة التعبير الكتابي الحر. وقد اضطلع المركز بدراسة مهارة الكلام/التحدّث في بعض الاختبارات العالمية عند ربطه ببعض اختباره اللغوية، فوجد أن قراءات هذا المكوّن من الاختبار تعطي ارتباطًا سالبًا مع المهارات الثلاث الأخرى في اختبار المركز. ولما رُبط أداء المختبرين في مهارة التحدّث بأدائهم في المهارات الثلاث الأخرى في الاختبار العالمي نفسه، وُجد أن الارتباط بينها سالب أيضًا. وهذا يؤكّد ما يراه المركز من أن درجة هذا الجزء من الاختبار لا تُدرج ضمن الدرجة الكلية إلا بعد التأكد من أن شروط عدالة الدرجة متحقّقة.

## • إنتاج الاختبار:

يتكوّن الاختبار من مكوّنين رئيسيين: (١) التعليمات، ومنها تعليمات عامة كالإذن بفتح كتيب الاختبار، ونهاية كل قسم، وعدد البنود فيه، ووقته، وتعليمات تظليل الإجابات. ومنها تعليمات تفصيلية كتعليمات كل مكوّن مثل تعليمات الإجابة عن أسئلة المسموع أو المقروء أو الكتابة أو التراكيب النحوية... وهذه التعليمات التفصيلية تُجرّب مقروئيتها في النسخة التجريبية من الاختبار للتأكد من وضوحها وعدم تضمّنها أي شكل من أشكال الغموض. (٢) البنود، توزّع بترتيب وعدد معين في أقسام، وترتّب بكيفية معيّنة حسب الصعوبة. ويشتمل النموذج الواحد من الاختبار على (١٠٠) بند أساسي من نوع الاختيار من متعدّد تُرصد لأخذ الاختبار درجتها، ويشتمل على سؤال واحد للتعبير الكتابي الحر. وهناك بنود تجريبية موزعة بالتساوي على نماذج الاختبار



الخمسة لكي يستفاد منها في بناء الاختبارات المستقبلية بعد تحليلها ومعرفة الخصائص السيكومترية (IIF) لكل بند منها. وتوزع جميع البنود على ستة أقسام، مدة كل قسم (٢٥) دقيقة، فيكون مجموع زمن أداء الاختبار (١٥٠) دقيقة. وللإختبار الواحد خمسة نماذج (A, B, C, D, E) لها وظيفتان أساسيتان: السماح بتوزيع أخذي الاختبار عند تطبيقه في قاعة الاختبار فيكون لدى كل مفحوص نموذج يختلف عن المجاور له، وإعطاء فرصة لتطبيق عدد أكبر من البنود التجريبية.

### • كتابة أسئلة الاختبار:

يتبع المركز سياسة صارمة في تطوير اختباره وتطبيقها وتحليلها ودراسة شواهد صدقها وثباتها وصيانتها، ويشترط في منتجاته توافر مواصفات جودة عالية، كما يشترط في العاملين على منتجاته أخلاقيات عمل مهنية، واستكتاب الأسئلة واختيار كتبها من ضمن المنتجات والعاملين عليها؛ ولذلك تبدأ عملية وضع الاختبار بعد تحديد هدفه ومكوناته، باختيار كتابة الأسئلة واصطفائهم، عن طريق الدراية الشخصية من المركز أو عن طريق المهنيين ذوي الدراية بشركاء المهنة، عن طريق الإدارة والإشراف أو المجال المعرفي المتخصص. ويدعى عدد مناسب للمهام المطلوب تحقيقها، وفي كل ورشة لمهام اختبار اللغة العربية لغير الناطقين بها يدعى من (١٢ - ١٥) مرشحاً لكتابة الأسئلة، وقد يزداد العدد مبدئياً إلى (٢٠) تحسباً للطوارئ أو الاعتذارات والهدف من هذا العدد هو تغطية المهام المطلوبة والحصول على عدد يتيسر تدريجه، فكلما زاد العدد قلت فاعلية التدريب المباشر. تجدد الورشة قبل انعقادها بستة أسابيع إلى ثمانية، وتُهيأ لها حقائب التدريب التي تشمل الاحتياجات العينية كالأوراق والأفلام والحقائب والشارات، والاحتياجات الإدارية كنماذج الحضور والتعهدات الخطية والتعليمات، والاحتياجات الفنية العلمية كالعروض والاستبانات والنماذج والإرشادات. وتقدم للمرشحين في اليومين الأولين مفاهيم عامة عن صناعة الاختبارات المقننة كفنيات الأسئلة وخصائصها الجيدة، وتحليل البنود وكيفية تحقيق الاختبار لسمات الثبات والصدق، ويقدم هذه المفاهيم متخصصون في القياس والاختبارات. ثم يقدم لهم هدف الاختبار الذي سيشاركون في كتابته مادته ومكوناته وأوزان هذه المكونات. بعدها يبدأ العمل التدريبي المباشر، فيقسمون إلى مجموعات صغيرة يشرف على كل واحدة متخصص في مجال الاختبار، مع حضور متخصصين في القياس والتقويم والاختبارات. وإذا كانت الورشة قد سبقتها ورشة أولى على الأقل في مجال الاختبار المعني دُعي بعض الكتاب القدماء وبخاصة من الجيدين منهم لتواصل التجربة بين الكتاب قديمهم وحديثهم. تعالج كل مجموعة مكوناً من مكونات الاختبار ويطلب منهم كتابة أسئلة لمناقشتها علنياً مع زملائهم تحت إشراف المتخصصين. وإذا انتهت مجموعة من التدريب على أسئلة أحد المكونات انتقلت إلى مكان المجموعة الثانية للتدريب على المكون الذي لم تتدرب عليه، وتبادلها المجموعة الأخرى الانتقال حتى تغطي جميع مكونات الاختبار ويكون لدى كل متدرب إمام جيد بكل مكون من مكونات الاختبار. بعدها يكلف كل متدرب بكتابة عدد معين من الأسئلة يتراوح بين (٥٠ - ١٠٠). وإذا كان المتدرب من القدماء استكتب في المكون الذي أوصت لجنة التحكيم باستكتابها فيه مع إطلاعه على توجيهاتها. وقبل أن تدبل جذوة التدريب ومفاهيم الورشة ماتزال حية في الأذهان يطلب من الكتبة تقديم ما كلفوا به في غضون خمسة أسابيع، ولكتاب الأسئلة أن يحضر أسئلته بنفسه إن كان قريباً من مقر المركز، أو يرسلها عبر ناقل متعهد ومتعاقد مع المركز لضمان سرية الأسئلة وخصوصيتها، وتوجه إلى رئيس قطاع الاختبارات والقياس. تُستلم الإرساليات وتحفظ نسختها الأصلية بعد تصويرها ثلاث نسخ.





## • تحكيم الأسئلة:

يُعدى ذوو الخبرة من المحكمين لاجتماع تُعرض فيه إجراءات التحكيم وبياناته واستماراته وقوائم المصطلحات المتداولة، وتُشكل منهم لجان التحكيم، ثلاثة في كل لجنة، أحدهم من المختصين في مجال السؤال، والثاني من المختصين في القياس، والثالث من مجال يجمع بين الاختصاصين مهمته الموازنة بين المتخصصين وتمثيل دور الطالب أو ولي الأمر أو المجتمع. يختار أعضاء اللجنة رئيسًا لهم من بينهم لإدارة الجلسة وتدوين الملحوظات، ويسجل كل عضواً رأيه وتعليقاته على نسخته على أن تكون نسخة الرئيس هي النسخة المرجعية، ويدون عليها بالقلم الأحمر ما اتفق عليه من تعديلات على السؤال على ألا تتجاوز مدة المناقشة لكل سؤال سبع دقائق. وتدوّن اللجنة رقمياً كل ملحوظاتها بما في ذلك الحكم على السؤال بالقبول دون تعديل أو بتعديل طفيف أو جذري أو مرفوض مع تبين سبب الرفض. والسؤال المعدل جذرياً ينقل إلى نموذج خاص بذلك. كما أنها تبين وجهة نظرها في مناسبة السؤال لمجاله ومدى صعوبته وإبداعيته إذا كان كذلك، وخلوّه أو عدم خلوّه من التحيز ضد الجنس أو المنطقة أو نوعية التعليم. وبعد انتهاء اللجنة من تحكيم أسئلة كاتب ما، تملأ له استمارة تقويم تتضمن مجموع ما قُبِل أو عدّل أو رُفض من أسئلته في كل مكوّن مع تقويمها لجوانب قوته وضعفه، فيُستفاد من ذلك في إعطاء كل كاتب تغذية راجعة عن أدائه، وعن سبل تعزيزه، وعن المكوّن الذي لا يظهر أن للكاتب فيه مهارة؛ فلا يكلف بالكتابة فيه في المرات القادمة. تدخل جميع بيانات التحكيم في الحاسوب وتبدأ من هناك عملية تجهيز الأسئلة للتجريب. فإذا ما جُهرت للتجريب حُكمت للمرة الثانية لضمان اجتيازها مرحلة التجريب بنجاح؛ إذ يتجنّب فيها ما عُرف من الصوارف التي تصرف أنظار الطلاب إلى الاختيارات غير الصحيحة. ويُنأكَد من وجود إجابة صحيحة لكل سؤال.

## • اللغة المستهدفة:

يقيس هذا الاختبار كفاية المختبرين في مهارات القراءة والاستماع والكتابة في اللغة العربية الفصيحة المعاصرة المستعملة عربياً في وسائل الإعلام والهيئات الرسمية والمؤسسات الأكاديمية وسبل التواصل الأخرى، وهي أيضاً المستوى المتعارف عليه عالمياً في المحافل الدولية.

## • الإجراءات الميدانية ودراسات الصدق والثبات:

### - التحليل الإحصائي:

توزع الأسئلة التجريبية في كتيب الاختبار الحقيقي بطريقة موزونة وممثلة لمكونات الاختبار ومجالاته المعرفية وصعوبته. ولا يوجد شيء يميّز الأسئلة التجريبية عن غيرها؛ ولذا فإن المبحوثين يجيبون عن الأسئلة في جميع كتيب الاختبار بمستوى واحد من الجدية، ويجيبون عليها كما يجيبون على الأسئلة الحقيقية مع أنهم يعلمون سلفاً بأن هناك أسئلة تجريبية غير محتسبة الدرجة. وبعد التطبيق تصحح الأسئلة التجريبية وتبدأ قراءتها وتحليلها إحصائياً، والسؤال الذي يحقق خصائص السؤال الجيد يدخل في بنك الأسئلة ويصبح سؤالاً حقيقياً، والذي لا تُحقّق خصائصه شروط السؤال الجيد فإنه يُنظر فيه فإن أمكن استصلاحه كان ذلك وأُعيد إلى التجريب مرة أخرى وإلا رُفض.

## - المنهج:

### اختيار العينة:

طُبِّقَت العينة التجريبية في أربعة معاهد لتعليم اللغة العربية في أربع جامعات سعودية، هي: جامعة الملك سعود (الرياض)، وجامعة أم القرى (مكة المكرمة)، وجامعة الإمام محمد بن سعود الإسلامية (الرياض)، والجامعة الإسلامية (المدينة المنورة). وشارك في هذه العينة حوالي (١٠٠٠) طالب يمثلون أكثر من (٥٠) جنسية وتتفاوت مستوياتهم.

### الأداة:

أُسْتُخِرَ من التجريب السابق اختباران (٧٠٥١، ٧٠٥٢) وطُبِّقَا في أربع جامعات في (جاكرتا) لتَحْرِى شواهد الصدق. شارك في هذا التطبيق أكثر من (١٢٦٠) طالبًا وطالبة. أنهى أغلب هذه العينة برنامج تعلم اللغة العربية وانخرطوا في الدراسة الجامعية، واللغة العربية هي وسيلة التدريس فيها. وبقيت العينة هم طلاب تتفاوت مستوياتهم من متوسط إلى متقدم في برنامج تعليم اللغة العربية.

## - التحليل والدراسات:

أجريت على المقياس عدة دراسات للتأكد من صدقه وثباته؛ فكانت درجة ثباته (٠,٩٠)، وهذا يعني أنه ذو ثبات عال. أما ثبات معايير تصحيح التعبير الكتابي (الكتابة الحرة أو ما يُسمَّى free writing) فقد نُظِرَ إليها من جانبين: أحدهما يتعلق بالتباين المُفسَّر وفق نظرية التعميم أو ما يُسمَّى (G-Theory) حيث بلغت قيمته (٠,٩٧)، أي أن (٩٧٪) من الفروق بين درجات الطلاب هي فروق في قدرتهم اللغوية وليس شيئاً آخر، أما الآخر فيتعلق بمدى إمكانية الاعتماد على مصحح واحد أو أكثر؛ فأظهرت الدراسة أن الثبات النسبي والمطلق كان عند درجة (٠,٩٦٩) حين يُحْتَكَم إلى مصحح واحد، ويصبح عند درجة (٠,٩٨٤) حين يُحْتَكَم إلى مصححين، وعند درجة (٠,٩٨٩) حين يُحْتَكَم إلى ثلاثة مصححين، وعند درجة (٠,٩٩) عندما يُحْتَكَم إلى أربعة مصححين. ويمكن ملاحظة أن الثبات لا يرتفع كثيرًا حين يضاف مصححان آخران، ويُستخلص من هذا أنه لو لم يُحْتَكَم إلا إلى مصحح واحد لكان الثبات عاليًا جدًا.

أما شواهد الصدق فقد كانت متعددة، ففي إحداها فُحِصَت حساسية الاختبار على التمييز بين المستويات المختلفة من الكفاية اللغوية لدى غير الناطقين بالعربية، وكانت العينة البالغ عددها (٧٩٤) طالبًا هم من طلبة وطالبات الجامعات الأندونيسية الدارسين في قسمي الدراسات الإسلامية واللغة العربية وبعض الدارسين في معاهد تعليم اللغة العربية. وكان السؤال الأول لهذه الدراسة عن الفروق بين الطلبة الجامعيين وطلبة أنها الثانوية العامة وانخرطوا في معاهد تعليم اللغة العربية على أبعاد الاختبار الثلاثة (فهم المسموع، وفهم المقروء، والكتابة)، وأظهر التحليل الإحصائي أن هناك فروقًا دالة إحصائية في المستوى اللغوي بين طلبة وطالبات الجامعة وطلبة وطالبات المرحلة الثانوية لصالح طلبة الجامعة. وحين طرح سؤال عن قدرة الأبعاد المختلفة على التمييز بين المستويات المختلفة لعينة الدراسة واستُخِدم التحليل التمايزي (Discriminant analysis) أظهرت نتائج التحليل أن أبعاد المقياس الثلاثة قادرة على التمييز بين طلبة الجامعة وطلبة الثانوية العامة، إلا أن حساسية فهم المقروء التمييزية أعلى، يليه فهم المسموع، ثم الكتابة.

وأجريت دراسة أخرى على العينة نفسها للتعرف على البنية العاملة للمقياس وفق التحليل العاملي التوكيدي للتأكد من أن الاختبار يقيس بُعدًا واحدًا، وأي مكوناته يتحلل باستقلالية يمكن معها إظهار درجاتها للطلبة عند إصدار التقارير؛ فاتضح أن المهمن والمسيطر على إجابات الطلبة والطلبات هو بعد واحد وهو القدرة اللغوية.

كما أن البنود المستخدمة في قياس أجزاء الاختبار الثلاثة (فهم المسموع، وفهم المقروء، والكتابة) تقيس ما وضعت من أجله، وأنه يمكن تقدير كفاية الطالب أو الطالبة اللغوية من خلال الدرجة الكلية في الاختبار، وتقدير مهاراته اللغوية في كل بعد من أبعاد الاختبار على حدة من خلال درجات أحدهم في ذلك البعد.

وفي دراسة ثالثة أُضيفت شواهد صدق البناء الداخلي وفق نموذج (راش) لتأكيد إمكانية تفسير درجات المقياس لما وُضعت من أجله، ومدى ثبات الدرجة المعطاة، وقد أظهرت نتائج البحث أن ثبات درجة الشخص هو (٠,٩٢) وهي قيمة عالية وتدل على ثبات الدرجات التي يحصل عليها الشخص حين يعيد الاختبار أكثر من مرة في مدد متقاربة. وقد أظهرت النتائج أيضاً أن الاختبار أحادي البعد، كما أظهرت الدراسة أن الاختبار مناسب لشرائح قدرات الطلاب على الرغم من وجود بعض الأسئلة السهلة. وأجريت دراستان عن التصحيح الكتابي لتقويم دقة المصححين، ومدى إمكانية الاعتماد على واحد أو أكثر. فاختيرت عشوائياً أوراق ثلاثة مصححين من (٤٩) ورقة مصححة، ثم رصدت تقديراتهم لدرجات الطلاب على أبعاد التصحيح الكتابي، وقد استخدم نموذج (راش) للجوانب (الوجوه) المتعددة (Many-facet Rasch model) للتعرف على سلوكيات المصححين في إعطاء الدرجة ومدى التشابه والاختلاف بينهم، فأظهرت النتائج أن قدرات الطلاب يمكن تصنيفها إلى ستة مستويات عند حذف القيم المتطرفة وبثبات قدره (٠,٩٤)، أو إلى أربعة مستويات بثبات قدره (٠,٨٦) عند إدراج القيم المتطرفة. وفي كلتا الحالتين فإن التصنيف يكون دقيقاً نظراً لقيم الثبات المرتفع. أما عن سلوكيات المصححين، فقد أظهرت النتائج عدم تمييز المصححين بعضهم من بعض، ولم تظهر سلوكيات معينة للمصححين في أثناء إعطاء الدرجات، أي أن فهم المصححين لمعايير التصحيح واستخدامهم لها متشابه بدرجة عالية.

وأبرزت أحدث دراسة (الحربي، ٢٠١٥) أن لهذا الاختبار قدرة على اكتشاف المجموعات الكامنة/الخفية (التي لم يتعرف عليها صانعو الاختبار سلفاً)، التي من الممكن أن تتعرف على سمات خفيه لأخذي الاختبار تفيد صانعي البرامج التعليمية لتدريس اللغة العربية لغير الناطقين بها مستقبلاً في التصنيف أو التسكين أو الاختيار. وتحليل نتائج أسلوب تحليل المجموعات الكامنة في محاور اختبار اللغة العربية لغير الناطقين بها، وُجد أن هناك ثلاث مجموعات كامنة مهمة، يمكن الاستفادة منها بالتعرف على بعض خصائص هذا الاختبار، حيث ظهرت المجموعة الكامنة الأولى وكان نسبة المسكنين منها من المختبرين (٢٨,٢٪)، وقد تبين أن هؤلاء المختبرين من الأفراد ذوي المستويات المنخفضة في جميع محاور الاختبار، والمجموعة الثانية وكانت نسبة المسكنين فيها من المختبرين (٣١,٠٪) ويقع فيها المختبرون ذوو المستويات المتوسطة، والمجموعة الثالثة (٤٠,٤٠٪) ويقع فيها المختبرون ذوو المستويات المرتفعة. وبالنظر إلى محاور الاختبار في ضوء المستويات الثلاثة للمختبرين (المجموعات الكامنة) المستقاة من بيانات هذه الدراسة، وجد الباحث أن أعلى مستوى أداء حدث في محور فهم المقروء في المجموعات الثلاث (المرتفعة، والمتوسطة، والمنخفضة) مقارنة بمستوى الأداء في محوري فهم المسموع والكتابة. ووجد أيضاً أن أكبر الفروق في المتوسطات بين مستويات الأداء الثلاثة كان في محور فهم المقروء، يلي ذلك في مستوى الأداء والفروق محور الكتابة، ثم فهم المسموع. وبدراسة مدى تأثير العلاقة بين متغيرات الجنس والمرحلة الأكاديمية والخلفية اللغوية مع المجموعات الكامنة المقدره من بيانات هذه الدراسة، وُجد أن الأكثر وجوداً في المجموعة الكامنة الثالثة (المختبرين ذوي الأداء المرتفع) هم الإناث ومن ذوي اللغة الواحدة، وفي المستوى الأكاديمي وجد الباحث أن الدارسين في المرحلة الجامعية من عينة هذه الدراسة هم أبعد مجموعة من أن يكونوا ضمن منخفضي المستوى أو متوسطيه. أما المجموعة الأولى (ذوو المستويات المنخفضة) فوجد بها الذكور وذوو اللغة المتعددة، ومجموعة معينة من الدارسين في المرحلة الثانوية.

- الدراسات التي أجراها المركز على هذا الاختبار:

١. د. أمجد العويضة، (٢٠١٣)، شواهد صدق البناء وفق منهج (راش).
٢. د. أمجد العويضة، (٢٠١٥)، استخدام نظرية استجابة البند في تحديد تحيُّز البنود والاختبار: اختبار اللغة العربية المُقنَّن لغير الناطقين بها مثلاً

Detecting Test and Item Bias in Light of IRT Model:  
Evidence from STAPSOL.

٣. د. خليل الحربي، (٢٠١٣)، تحليل بنود الاختبار وشواهد الثبات.
٤. د. خليل الحربي، (٢٠١٣)، شواهد صدق البناء وفق برنامج (m-plus).
٥. د. خليل الحربي، (٢٠١٣)، ربط درجة الاختبار بمستوى الطالب التعليمي.
٦. د. خليل الحربي، استخدام التحليل الهرمي الكامن في تحديد مستويات الطلاب اللغوية: اختبار اللغة العربية المُقنَّن لغير الناطقين بها.

Employing Latent Classes Analysis in Identifying  
Language Proficiency Levels in Arabic Learners: the Case  
of STAPSOL.

٧. د. سيف القحطاني، (٢٠١٣)، تصحيح التعبير الكتابي وفق نظرية (G+ theory).
٨. د. سيف القحطاني، (٢٠١٣)، تصحيح التعبير الكتابي وفق نظرية الجوانب المتعددة (Multifaceted Theory).
٩. د. سيف القحطاني، (٢٠١٥)، أحد تطبيقات نظرية التعميم في الكشف عن الخصائص السيكمترية لاختبار اللغة العربية المقنن لغير الناطقين بها (STAPSOL).
١٠. د. عبدالرحمن الشمrani، (٢٠١٣)، إطار الاختبار النظري وتحديد خريطته.
١١. د. عبد الرحمن الشمrani، (٢٠١٣)، تصحيح التعبير الكتابي: الإطار النظري وتحديد المعايير (Rubrics).
١٢. د. عبدالله القاطعي، (٢٠١٥)، اختبار اللغة العربية المُقنَّن لغير الناطقين بها: أين نحن، وإلى أين الاتجاه؟



- الهوامش:

1. انظر مثلاً: Common European Framework of Reference for Languages, P.1. (2007)
2. ACTFL Proficiency Guidelines (2012) – Arabic annotations and samples.
3. Common European Framework of reference Forlanguage,2009 .
4. Association of Language Testers in Europe, Manual for Test (4) Development and Examining. [www.coe.int/lang](http://www.coe.int/lang)





## رؤية المركز

أن نكون مرجعاً عالمياً في القياس والتقويم.

## رسالة المركز

تقديم حلول شاملة متكاملة لقياس المعارف والمهارات والقدرات وتقويمها؛ بمنهجية علمية؛  
إسهاماً في تحقيق العدالة والجودة وتلبيةً للاحتياجات التنموية.

١. ينبغي الاطلاع على شروط القبول في الجهات التي تشترط اختبارات المركز.
٢. ينبغي التسجيل في اليوم والمكان المناسب لك خلال فترة التسجيل المحددة.
٣. يجب الحضور للاختبار في الموعد الذي سجلت به.
٤. لا يمكنك دخول الاختبار بدون أصل اثبات الشخصية الموضح في هذا الكتيب.
٥. سترسل نتائج الاختبار إلكترونياً إلى جميع الجهات التي تشترط الاختبار، لذلك لا يتطلب منك احضار شهادة ورقية.



للتذكير

📞 9200 33 555

📠 +966 11 490 9077

@ faq@qiyas.org

✉ ص.ب 68566 الرياض 11537

المركز الوطني للقياس

National Center for Assessment



📱 🌐 📧 📞 Qiyasonline | www.qiyas.sa