الفصل السادس

تأثير اختبارات اللغة المحوسبة
The Impact of CALT

يقول بينيت الباحث في مجال القياس وأحد المؤيدين المتحمسين لاستخدام التقنية عن التأثير التحويلي الكبير الذي تحدثه هذه التقنية في التقييم التربوي واسع النطاق:

> "ستحدث التقنية الجديدة (نوعاً) من التحول (في التقييم) من خلال السماح لنا بإعداد اختبارات أكثر رسوخًا من حيث استقرار المفاهيم المتعلقة بما يحتاج المرء لمعرفته وتمكينه من النجاح في مجال ما؛ وذلك بجعل تقييم الأداء أمرًا عمليًا ومنتظمًا باستخدام المحاكاة المحوسبة، وإنتاج البنود آليًا وأتمتة تصحيح المقال، وأيضًا بتغيير أساليب تقديم الاختبارات واسعة النطاق وأغراض استخدامها. (Bennett, 1999a، ص ١١).

وتعود حماسة بينيت، بشأن القوة التحويلية للتقنية في مجال التقييم، إلى جذور مبكرة، وقد أشار الباحثون في مجال القياس التربوي إلى كثير من هذه الأفكار منذ سنوات (على سبيل المثال Bejar, 1985 ، وCole, 1993 ، وBejar & Braun, 1994). ومع أن بينيت وغيره من الباحثين المتحمسين لا يشيرون على وجه الخصوص إلى اختبارات اللغة الثانية، إلا أنهم طالما تناولوا اختبارات الكتابة في مناقشاتهم. وباستقراء هذه التوقعات، لا يستطيع معلمو اللغات الثانية ومعدو الاختبارات والباحثون إغفال ما إذا كانت تقييماتنا تمثل جزءًا من الطفرة التي يشهدها مجال التقييم. وإذا كان الأمر كذلك،

فإننا لا نعرف ما إذا كانت هذه الطفرة قد وقعت بالفعل أم أننا نعيشها حاليًا، أم أننا على موعد معها مستقبلاً.

سنعرض في هذا الفصل فكرة مفادها أن هذه الطفرة في مجال تقييم اللغة الثانية — على الرغم من التغيرات والتطورات التي طرأت بفضل استخدام التقنية — التي تصورها بينيت لم تحدث بعد.

ومع أن التقييم المؤتمت (الآلي) وتقديم الاختبارات بالحاسوب قد أصبحا حقيقة واقعة وملموسة، فإننا لم نستطع تقديم الدليل على أن تقييم الأداء قد أصبح أمرًا عمليًا من خلال استخدام المحاكاة على نطاق واسع، أو إنتاج بنود موثوقة، أو إحداث تغييرات ملموسة في أغراض الاختبار من خلال الاعتماد على التقنية. إن مثل هذه التغييرات الهائلة بحاجة إلى الدعم والانتشار عبر التطورات المفاهيمية في فهمنا للغة واستخدامها وأساليب تعلمها، ومنذ ما يربو على خمس عشرة سنة خلت، أشار آلدرسون إلى هذا الأمر عند تناوله للاختبارات الصفية ذات الطابع الفردي، حيث قال:

> "تثير احتمالات التشخيص والعلاج مشكلة مهمة في وجه علماء اللغة التطبيقيين ومدرسي اللغة؛ فما يحد من نطاق إعداد هذه الاختبارات لا يعود إلى إمكانيات الأجهزة أو تعقيد مهمة البرمجة، وإنما يعود إلى فهمنا القاصر لطبيعة تعلم اللغة وكيفية استخدامها ... فالتحدي الماثل (لاختبارات اللغة المحوسبة) يجابه علماء اللغة وعلماء اللغة التطبيقيين من حيث تقديم معلومات مناسبة بشأن طبيعة نظم التفريع، وكذا الإشارات، والقرائن، والتعقيبات التي تساعد المتعلمين، بأكثر مما يطال مبرمج الحاسوب في تطوير البرامج المناسبة". (Alderson, 1990، ص ٢٥).

كان آلدرسون يركز على اختبارات اللغة المحوسبة لتكون في خدمة التعلم داخل الصفوف الدراسية، لكن كان بالإمكان توجيه التعليقات المشابهة لما ناله هذا المجال إلى أنواع أخرى من الاختبارات أيضًا. وليس هناك الآن سوى أثارة من أدلة تؤهل لطرح

مفاده إحراز المزيد من التقدم في بناء أنواع بعينها من المعارف التي قد تدعم هذه الطفرة. والواقع أن الدراسة الحديثة لمشروع D1ALANG (Alderson, 2005) — عن اختبار تشخيصي محوسب واسع النطاق — ينتهي للنتيجة نفسها ؛ ذلك أن المشروع استُحدث على أساس مستويات الإطار الأوروبي المشترك للغات (Common European Framework)، ولكن نظرًا لعدم تطرق بحوث اكتساب اللغة الثانية لتحصيل القواعد والمفردات في ضوء هذه المستويات، فإن التشخيص المنتظر لصيغ ووظائف لغوية محددة ليس له أساس واضح من حيث النظرية ولا البحث. في هذه الحالة، ثمة محاولة لتوسيع استخدامات الاختبارات بصفة أساسية من خلال تزويد المتعلمين بمعلومات تشخيصية من شأنها توفير توجيه دقيق بشأن طبيعة المادة المدروسة ؛ إلا أن المعرفة المتخصصة في اكتساب اللغة الثانية، لا تكفي لتوفير الأساس لهذا التغيير الجذري.

مع أن هذه الطفرة لما تحدث بعد، فإن فكرة مسؤولية التقنية عن إحداث تغييرات في اختبارات اللغة — فيما يتعلق بممارسة اختبارات اللغة حالياً — تحظى ببعض المصداقية. وعلى كل، فكثير من اختبارات اللغة أصبحت محوسبة، كما أن الكتب المعنية بكل قضية من قضايا تقييم اللغة، تتوقع دورًا كبيرًا للتقنية مستقبلاً. فظاهريًا، يبدو أن كثيراً من الطلاب متأثرون باختبارات اللغة المحوسبة. من هنا نرى — وحق لنا ذلك — أن هذه التغييرات تعد تحولاً وتطوراً ؛ ذلك أنها لم تُحدث تقدمًا وتغييرًا في طريقة الاختبار أو طرق استخدام الاختبارات، بل خطا الباحثون خطوات ملموسة إزاء التعامل مع بعض المشكلات الملحة في مجال تقييم المهارات اللغوية باستخدام اختبارات اللغة المحوسبة. وعوضًا عن وصف طفرة تميزها أنواع جديدة من الاختبارات والأدوار المنتظرة منها، فقد كتبنا عن التطورات الكبرى والقضايا التي نشأت جراء تطور جوانب الاختبارات ومدى تحقيق الصدق فيها. وفي هذا الفصل، سننظر في بعض ملابسات هذه التطورات

فيما يتعلق باللغويات التطبيقية، كما سنطرح رؤيتنا لبعض الاتجاهات المستقبلية، وسنناقش ما قد يترتب على هذه التطورات الجديدة.

التطورات التي حدثت في التقييم اللغوي جراء استخدام اختبارات اللغة المحوسبة

عرضنا في الفصول السابقة وصفًا لاختبارات اللغة المحوسبة المستخدمة حاليًا، والبحوث التي عنيت بزيادة فهم هذه الاختبارات وتحسينها، إضافة إلى الطرائق المستخدمة في إعدادها والقضايا المتعلقة بتحقيق عنصر الصدق فيها. فقد تناول الفصل الأول دور التقنية المتعاظم في الممارسات اليومية للتقييم، وانتهينا إلى أن المعرفة وفهم القضايا ذات الصلة بالتقنية، تعد من القضايا الجوهرية لمدرسي اللغات، والمتخصصين في إعداد الاختبارات، والباحثين في مجال اختبار اللغة.

تطرق الفصل الثاني إلى طرائق استخدام اختبارات اللغة المحوسبة في كثير من برامج الاختبارات ومواقفها، مما يزيد من الخيارات أمام معدي الاختبارات لبناء مزيد من المهام الاختبارية، وقدمنا وصفًا للطرق التي تؤثر من خلالها التقنية في خصائص طرائق الاختبار، بما في ذلك معطيات الزمان، والمكان، ومعايير الأداء الخاصة بالاختبار، وخصائص المدخلات والاستجابات، فضلاً عن التفاعل الذي يحدث بين الاستجابات والمدخلات والاستجابات وخصائص التقييم. وقد أوضحنا بأن أكثر الاختلافات بروزاً توجد في خصائص المدخلات والاستجابات والتفاعل الذي حدث بينهما، وبالتالي التفاعل الذي يحدث بينهما وبين عملية التقييم. أما التطورات في مجال تقديم الاختبارات والوصول إليها، فكانت واضحة من خلال أمثلة المدخلات السياقية الثرية وتنوع أساليب الاستجابات وخاصية التكيف التي يوفرها الحاسوب، والتصحيح الآلي الذي أضحى ممكنًا بفضل الحاسوب. ومع ذلك، لم نتمكن من الحديث عن التغييرات المترتبة على الطفرة العلمية مثل تقييم الأداء الذي أصبح فعالاً بفضل استخدام المحاكاة.

تناول الفصل الثالث المشكلات المحتملة التي قد تنجم عن استخدام اختبارات اللغة المحوسبة فيما يتعلق بطريقة تأثيرها في معامل الصدق. وأشرنا في هذا الصدد إلى ستة مخاوف تطرأ غالبًا باعتبارها مخاوف محتملة ذات صلة باختبارات اللغة المحوسبة، وهي: اختلاف معدلات الأداء في الاختبارات، وأنواع جديدة من المهام، وأوجه من القصور جرّاء الاختيار التكيّفي للبنود، وعدم الدقة عند تصحيح الاستجابات آلياً، والإخلال بعنصر الأمان والسرية، والتأثير السلبي للاختبار. كما أوضحنا أن دراسات قليلة قد حاولت الوصول إلى فهم أفضل لمعنى نتائج الاختبارات بالنظر في اختبارات اللغة المحوسبة وما يترتب على استخدام هذه الاختبارات. فضلاً عن ذلك، أشرنا في الفصل ذاته إلى أن المهددات المحتملة التي تؤثر في عنصر الصدق، قد دخلت ضمن إطار الشكوك حول مدى إضعاف التقنية الممارسات الحالية في الاختبارات أو مدى إنصافها، أي أن أكثر المناقشات شيوعًا بشأن اختبارات اللغة المحوسبة يحيط بها إطار من الشك في كيفية إضعاف التقنية للممارسات الحالية بدلاً من وجهة النظر الابتكارية الرامية إلى اكتشاف كيفية إسهام التقنية في إحداث طفرة تعمل على تطوير الإفادة الكلية من التقييم.

تناولنا في الفصل الرابع، كيفية تعامل معدي اختبارات اللغة المحوسبة مستقبلاً مع أدوات التوليف مثل نظام WebCT، مشيرين في ذلك إلى أن الأدوات المطلوبة في إعداد الاختبارات إنما تعتمد على الغرض من التقييم، والقضايا العملية ذات الصلة مثل عناصر: الزمان، والمكان، والتكلفة، والخبرة. تكمن حقيقة تقييم المهارات اللغوية، في أن قلة الموارد المالية والخبرات ذات الصلة لتطوير أدوات توليف خاصة بتقييم هذه المهارات، يحد من احتمالات الابتكارات الجديدة. وترتب على ذلك عدم تطوير نظم التوليف المثالية بهدف إجراء التقييم اللغوي، ولكنها تشكل مجالاً ثرياً للبحث والتقصي.

في حين طرح الفصل الخامس فكرة ضرورة تقييم الاختبارات المحوسبة، في ضوء معايير تتفق مع تلك المستخدمة لتقييم الاختبارات الأخرى. بيد أن القضايا المتصلة

بالتقنية، تستلزم تسليط الضوء عليها. ومن ثم، يجب وضع هذه القضايا المتعلقة بالتقنية التي حددها الباحثون في مجال اختبارات اللغة المحوسبة، ضمن إطار أوسع لتقييم الاختبارات – بالتركيز على الجوانب التي يستفاد منها في الاختبارات – والتي تبين العوامل المهمة في تقييم مهارات اللغة مثل الموثوقية باعتبارها عنصرًا رئيسًا في إعداد الاختبارات وتقييمها. كما استكشفنا في الفصل ذاته طرحاً تفسيرياً حدد بدوره هيكلاً قد يدخل في نطاقه اختبارات معينة وأموراً تخص قضايا واجهات المستخدمين. إلا أن هذه المناقشة كانت استهلالاً بالضرورة في ضوء القدر المحدود من البحوث التي أجريت على عنصر الصدق التجريبي في اختبارات اللغة المحوسبة من وجهات نظر معاصرة.

خلاصة القول، إن واقع اختبارات اللغة المحوسبة الآن ليس في حال يؤهله لأن نطلق عليه طفرة في مجال تقييم مهارات اللغة؛ إذ تختلف خصائص بعينها في طرائق اختبارات اللغة المحوسبة اختلافًا جذرياً عن الاختبارات المشتملة على وسائل أخرى لتقديم الاختبارات والاستجابات. لكن التقنية لم تُعِد صياغة دور التقييم في منظومة التعليم والتعلم من الألف إلى الياء. وبذلك نظرنا إلى اختبارات اللغة المحوسبة على أنها قفزة في مجال التقييم، توسع دائرة العمل في الاختبارات أكثر من كونها طفرة تغير أوجه ارتباط التقييم بتعليم اللغات والبحوث ذات الصلة. وقد تحدث طفرة في هذا الصدد مستقبلاً، لكن في الوقت الحالي، وفي ظل الدور المحوري لتقييم مهارات اللغة في اللغويات التطبيقية، فإن التغييرات المنبثقة عن التقنية تتقاطع عند نقاط مهمة مع مجالات أخرى من اللغويات التطبيقية.

اختبارات اللغة المحوسبة في اللغويات التطبيقية

يمثل إعداد اختبارات اللغة المحوسبة، وكذا استخدامها وتقييمها تحديًا، ويفتح آفاقاً رحبة لخيال اللغويين التطبيقيين، نظرًا للخيارات الجديدة التي أفرزتها الاختبارات

من خلال التقنية. ومن بين هذه القضايا، الحاجة إلى الاهتمام بطبيعة القدرات اللغوية اللازمة في التفاعلات والاتصالات التقنية؛ ومن هنا تبرز الحاجة إلى إعادة النظر في مكونات الاختبارات. وثمة قضية ثانية مفادها أن دقة المعلومات التي تجمع عن المتعلمين، وبالإمكان تحليلها على أساس استجابات الطلاب الكتابية، تدفع مصممي الاختبارات على إمعان النظر في أسلوب التعاطي مع مثل هذه القدرات. فعلى سبيل المثال: هل يجب على معدي الاختبارات إعادة النظر في كيفية إفادة بحوث اكتساب اللغة الثانية في إعداد الاختبارات التي تقدم معلومات أكثر تفصيلاً من الاختبارات، التي يعتمد فيها التقييم على العنصر البشري؟ أما القضية الثالثة فهي مرونة التقنية من حيث: اختيار زمان الاختبار، ومكانه، والاحتفاظ بالسجلات، ومدى قدرة هذه العوامل مجتمعة على توفير فرص جيدة للارتقاء بالتعليم من خلال التقييم. تلكم القضايا الثلاث، التي ظلت حاضرة على امتداد هذا العمل، تستحق مزيداً من المناقشة.

القدرة اللغوية والاستخدام

تؤكد الأعمال التي تناولت اختبارات اللغة المحوسبة على حقيقة مفادها، أن مكونات اللغة التي يُعوّل عليها في تفسير النتائج، تتطلب النظر فيها من منظور السياق الذي تستخدم فيه هذه اللغة. ومن هنا قد يتحدث علماء اللغويات التطبيقية عن القدرة اللغوية باعتبارها القدرة على اختيار الموارد اللغوية المناسبة واستخدامها في مواقف بعينها. غير أننا اليوم قد نستبدل هذا المفهوم، عن القدرة اللغوية، بمفهوم آخر يشتمل على القدرة على انتقاء اللغة الملائمة واستخدامها من خلال التقنيات المناسبة لكل موقف على حدة. وهنا، قد يكون البريد الإلكتروني مفيدًا في بعض المواقف، فيما تكون المكالمات الهاتفية أو المحادثات الشخصية أفضل في مواقف أخرى، وغالبًا ما يكون الخيار في يد مستخدم اللغة. وفي بعض الأحيان، قد يكون المدقق الإملائي مصدرًا مفيدًا، لكن هناك ما يدفع لتجاهله في أحيان أخرى. ويكون الخيار في يد

مستخدم اللغة أيضًا. وهذه الخيارات تعتمد، في نهاية المطاف، على مدى كفاءة مستخدم اللغة من حيث التقنية والأسلوب، ذلك أنهما يشكلان مع القدرة اللغوية نوع المكون ذي الصلة باستخدام اللغة بالاستعانة بالتقنية.

وبمعنى آخر، لا بد من النظر إلى القدرة اللغوية التواصلية من منظور الدور المشترك لكل من اللغة والتقنية في عملية التواصل. يتحدث راسول (Rassool, 1999) عن القدرة التواصلية في العصر الحديث قائلاً " تشير القدرة التواصلية إلى العملية التفاعلية التي تنشأ فيها المعاني – على نحو نشط – بين تقنية المعلومات والعالم الذي نعيش فيه" (ص ٢٣٨). وتثير هذه المقولة من منظور التقييم اللغوي القضية الجدلية المتعلقة بمكون اللغة المعتمد على السياق. وقد شرع باحثون في مجال تعلم مبادئ القراءة والكتابة، مثل تاينر (Tyner, 1998) في اكتشاف ما يمكن أن يعنيه هذا: إنهم يرون أن هناك حاجة إلى تجزئة مكونات مبادئ القراءة والكتابة للتعبير عن إيمانهم بأن التقنية تؤثر في طبيعتها المطلوبة للاستخدام اللغوي مع التقنيات المختلفة:

> ترى المناهج الجديدة، في تعليم مبادئ القراءة والكتابة وتعلمها، بأنه بدلاً من التعامل معها على أنها مفهوم ذو وحدة متراصة ... فمن الأفضل أن تقسم هذه المبادئ إلى وحدات كثيرة متعددة، ينفرد كل منها بخصائص فريدة تكشف عن مجموعة متباينة من الأهداف الاجتماعية.... وقد أُطلق على هذه الوحدات المتعددة من هذه المبادئ اسم مبادئ القراءة والكتابة التقنية (محو الأمية التقنية)، ومبادئ القراءة والكتابة المعلوماتية، ومبادئ القراءة والكتابة المرئية، ومبادئ القراءة والكتابة الوسائطية، ... إلخ. وكما تتقارب وسائط الاتصال المعاصرة في مزيج ملموس، كذلك تتقارب الخصائص الفريدة لكل نوع من هذه المبادئ وتتلاقي فيما بينها..." . (Tyner, 1998، ص ٦٠).

قد يختزل هذا الانتشار لوحدات المبادئ المذكورة، مدلول المصطلح ليعني أي قدرة؛ سواء كانت لغوية أم لا. وليس هناك معنى واضح جلي لتقييم المهارات اللغوية. ومع ذلك يبدو أن التقنيات المختلفة قد تؤثر في القدرات محل الاهتمام في التقييم اللغوي. وهذه فكرة تبدو كأنها قضية حرية بأخذها بعين الاعتبار في عملية إعداد الاختبار والتثبت منها حاليًا، سواء أكانت اختبارات اللغة المحوسبة داخلة في ذلك أم لا. ويعبر كل من بروس وهوقان (Bruce and Hogan, 1998) عن هذه الفكرة باعتبار أن التقنية جزء لا يتجزأ من التواصل؛ فهما يشيران إلى أن: أي فرد ليست لديه القدرة على استخدام التقنية، فلن يكون قادراً على التواصل في العديد من المواقف المهمة. وهنا تتجلى قضية مُشابهِة في المحاولات المبذولة لتحديد مكون يمثل محور اختبار ذي أغراض معينة للمهارات اللغوية، إذ إن كلا من المعرفة لأغراض خاصة والقدرة اللغوية والإستراتيجية، يحققان هدف التواصل (Douglas, 2000). وبالمثل، فمن المحتمل أن يكون كلا من مزيج القدرة اللغوية والإستراتيجية مع المعرفة التقنية، هما اللذان يحققان التواصل من خلال التقنية.

بالاستفاضة في مختلف أنماط المبادئ المرتبطة بالتقنية سالفة الذكر، يرى وارشاور (Warschauer) أن احتياجات المتعلمين من مهارات محو الأمية اللازمة لهم في عالم اليوم، تختلف اختلافًا نوعيًا عن احتياجاتهم للمشاركة في الحياة الثقافية، التي لا تتدخل فيها التقنية، كما يصف وارشاور (2000) مهارات مبادئ القراءة والكتابة، والمهارات اللغوية الجديدة المطلوبة للتواصل الفعال من خلال استبدال مكوني القراءة والكتابة، بالقدرات التي يسميها: القراءة/البحث، والكتابة/التأليف على التوالي (ص ٥٢١). هذه المكونات، التي تشمل جوانب القدرة الإستراتيجية المطلوبة لإنجاز العمل بنجاح في بعض البيئات الإلكترونية، تجبر معدي الاختبارات و مستخدميها على مواجهة كيفية تفعيل دور القدرات الإستراتيجية. وهذا لا يشكل أي مشكلة جديدة، بل يمثل

مشكلة تُطرَق وتَكبُر باستخدام اختبارات اللغة المحوسبة. وبذلك توفر هذه الاختبارات الحاجة والفرصة معًا للوصول إلى فهم أفضل للقدرات اللغوية المطلوبة في التواصل عبر الحاسوب. أما فيما يتعلق بالجدل التفسيري الذي ورد ذكره في الفصل الخامس، فقد يحتاج معدو الاختبارات إلى التعبير عن تفسير الدرجات من حيث القدرة على تجميع المعلومات المرئية على شبكة الإنترنت، وليس على أنها "القدرة على القراءة" عمومًا.

وفي محاولتهم لصياغة آراء مبنية على نظرية حول القدرات المطلوبة لاستخدام اللغة بواسطة التقنية، يواجه باحثو اختبارات اللغة تحديًا بشأن دمج التوجهات غير المتوافقة في الظاهر والمتصلة بتقييم اللغات ودراسات مبادئ القراءة والكتابة سالفة الذكر، إذ تتبنى هذه الدراسات نظرة عملية اجتماعية تنطوي على وصف السلوك أكثر منها النظرة الأكثر معرفية المنطوية على مزيد من أعمال التقييم اللغوي. أما من منظور الممارسة الاجتماعية، فيُنظَر إلى محو الأمية الإلكترونية، واكتساب مبادئ القراءة والكتابة متعددة المجالات، على سبيل المثال، على أنها فعل الإنسان باللغة من خلال التقنية، لا ما يحتاجون إلى معرفته عن اللغة والاستراتيجيات اللازمة للتعامل مع اللغة باستخدام التقنية. ويرى بعضهم أن منظور القدرة غير متكافئ مع منظور الممارسة الاجتماعية الخاص بالدراسات الجديدة في مجال المبادئ سالفة الذكر؛ ذلك أن الأول يستلزم تحديد قدرة نظرية مسؤولة عن الأداء في السياقات المختلفة، فيما يقوم الثاني على وصف مختلف صور الأداء في كل سياق على حدة. وقد يشجع اكتشاف الجدل التفسيري لاختبارات اللغة المحوسبة، اللغويين التطبيقيين على فهم أفضل للقدرات الكامنة في محو الأمية الإلكترونية، أو محو الأمية متعددة المجالات، وذلك بطريقة قابلة للقياس تسفر عنها نتائج مفيدة قابلة للتفسير.

اكتساب اللغة الثانية

تؤكد عمليات إعداد اختبارات اللغة المحوسبة وتقييمها على الحاجة إلى تقوية الروابط بين اكتساب اللغة الثانية والتقييم اللغوي؛ فأي امرئ حاول تصميم المواصفات المميزة لطريقة الاختبارات المحوسبة، أو تناول مشكلة تحديد المدلول الجزئي (الدرجات الجزئية) لاستجابات الطلاب، يمكن أن يتفهم هذه القضية محل النظر. ومن قبيل ذلك، إذا قدمنا المدخلات المكتوبة أو الشفهية للطالب، هل يجب تقديم المساعدة أيضًا؟ وإذا حددنا درجة من واحد إلى ستة لكل مقالة مكتوبة، ما الذي يجعل الحصول على خمس درجات أفضل من الحصول على أربع؟ وما الخصائص اللغوية والبلاغية والسياقية التي تدفع باتجاه تحديد درجة معينة على المقياس ذي النقاط الست؟ ومع وضع أساس منطقي لقضايا كهذه، يأمل المرء أن تكون النظرية والبحث المتعلقان باكتساب اللغة الثانية، بمثابة مصادر للمعلومات في ثلاثة مناح على الأقل.

أولاً: قد يبدو أن مراحل تطور البحوث الخاصة باكتساب اللغة يجب أن تؤدي دورًا في إعداد بعض التقييمات وتحديدها بدرجات، على أن يكون فيها للقواعد دور مهم يُراعى عند تقييم الأداء. على سبيل المثال، كتب نوريس (Norris) (قيد الطبع) بنوداً لاختبار تصفية في القواعد بناء على الدراسات المعنية باكتساب اللغة الثانية، وهي الدراسات التي أظهرت فعلاً التطور المبكر نسبيًا لبعض المعرفة في القواعد في جوانب أخرى. وكان متوقعاً من الطلاب ضعاف المستوى، الأداء بصورة أفضل في البنود التي تتطلب ترتيب الكلمات في جمل خبرية يتألف كل منها من عبارة واحدة، مقارنة بالأداء في البنود التي تتطلب الأمر نفسه في جمل ذات عبارات اسمية مضمّنة فيها. وقد صححت هذه البنود إما على أنها صواب أو أنها خطأ؛ ولم يسجَّل أي إنتاج لغوي، بيد أن المرء يأمل في أن يكون قادرًا على الكشف عن تحليل إنتاج المتعلم في ضوء العمل في مجال اكتساب اللغة الثانية.

تناولنا في الفصل الثالث التوتر الذي يعتري معدي الاختبارات الذين يحتاجون إلى أسس منطقية تسوغ الإجراءات اللازمة لتقييم الاستجابات عند تعدد الاستجابات المحتملة للبنود. وحتى الاستجابات غير اللغوية لمهام ترتيب النص التي ناقشها كل من آلدرسون وبيرزيش وزابو (Alderson, Percsich, and Szabo, 2000) بحاجة إلى أساس تنطلق منه لتحديد الدرجات لمختلف الترتيبات التي يقوم بها الطلاب. لكن المشكلة تفاقمت بفعل التباينات الكبيرة التي يمكن أن يدخلها الطلاب الخاضعون لاختبار كونيام للإملاء (Coniam, 1998)، الذي استخدم نظامًا آلياً تلقائيًا للتقييم، أو كما حدث في اختبار القراءة ذي الاستجابات المفتوحة النهايات الذي أعده جاميسون وكامبل ونورفليت وبيربيسادا (Jamieson, Campbell, Norfleet, and Berbisada, 1993)، أو المقالات التي طرحها كل من باورز وبورستين وكودورو وفاولز وكوكيش (Powers, Burstein, Chodorow, Fowles, and Kukich, 2001). في هذه الحالات جميعها، والكثير غيرها مما قد يطرأ على مخيلاتنا، يجب أن يستفيد معدو الاختبارات من قاعدة علمية، يمكنهم – استنادًا إليها – الحكم على بعض أشكال الأداء بمثابة دليل على المعرفة المتقدمة، بينما تكون بعض أشكال الأداء الأخرى ذات مستوى منخفض. علمًا بأن بعض المعارف المتخصصة في مجال اكتساب اللغة الثانية، يمكن الاعتماد عليها فيما يتعلق باختبارات اللغة المحوسبة على وجه مثالي.

يمكن أن يقدم العمل الذي حاول تحديد الخصائص اللغوية ذات الصلة بمستويات كتابة الإنجليزية بوصفها لغة ثانية، بعض الاقتراحات الخاصة بوضع أسباب منطقية لتحليل الاستجابات خلال عملية التقييم. وهنا يناقش كل من وولف – كوينترو، وإيناجاكي، وكيم (Wolfe-Quintero, Inagaki, and Kim, 1998) القضايا والنتائج الخاصة بالبحث المتعلق بالأداء على مستوى التراكيب الدال على المستويات في مؤشرات التطور؛ فهم يأخذون بالتطور إلى جانب أبعاد الطلاقة والدقة والتعقيد. أما هينكل (Hinkel, 2003)

فيحدد الاختيارات المعجمية (المفردات) التي يختارها من يكتبون باللغة الإنجليزية كلغة ثانية، وهى الاختيارات التي تسهم في إدراك عنصري التبسيط المخل وعدم الدقة في الكتابة. بذلك نجد أنفسنا أمام اتجاهين بحثيين مثمرين – على ما يبدو – من حيث الوصول إلى فهم أفضل لمستويات الأداء. في الوقت نفسه، نجد أن مثل هذا البحث محدود بالضرورة إذا فُسّر على أنه طرق لباب المعرفة اللغوية المكتسبة بترتيب غير متباين، أو أن هذه المعرفة لا تتأثر بالظروف التي تظهر فيها. ودون إغفال النتائج التجريبية التي تبرهن على زيادة المعرفة اللغوية أو نقصها، ينبغي أن يراعي الباحثون في اختبارات اللغة العوامل الإدراكية والسياقية التي تتفاعل في أثناء الأداء.

ثمة طريق ثانية تفيد فيها بحوث اكتساب اللغة الثانية في مسألة إعداد الاختبارات، وتقييم الاستجابات، تتمثل في البحث للوقوف على طبيعة الآثار التي تحدثها ظروف العمل على الأداء، ويتمثل الفرض الذي يؤكد على هذا المنحى في هذه البحوث، في أن الأداء يتطلب الشرح ليس من حيث معارف الطالب فحسب، وإنما أيضًا من حيث الظروف التي يجري فيها الأداء. لذلك يفترض باحثون من أمثال سكيهان (Skehan, 1998) وروبنسون (Robinson, 2001) خصائص مهام معينة يُنتَظَر منها المزيد، أو الأقل، من الظروف الصعبة التي يجري فيها الأداء. فعلى سبيل المثال، يتفاعل شرط الطلب من الطالب استجابة مكتوبة بسرعة أو ببطء، مع مستوى المعرفة اللغوية للطالب لإصدار مستوى أعلى أو أقل من الاستجابات؛ أي: أنه يجب مراعاة أبعاد أخرى غير مستوى المعرفة لدى تفسير الأداء.

ثمة بعد ثالث آخر ينبغي ربطه بهذا المنحى، وهو السياق الموضوع لأداء الطالب، إذ يصدر الإنتاج اللغوي المطول المطلوب في الاختبارات المشتملة على التحدث والكتابة استجابةً لطلب يقضي بإعداد تهيئة سياقية خاصة للطالب (Halliday & Hasan, 1989)، وهذا يعطي الطالب مفتاحاً لإنتاج لغة مناسبة لنوع المكون لتحقيق وظيفة تواصلية

مطلوبة (Paltridge, 2001). ومن ثمّ تساعد مثل هذه الاختبارات، إن كانت جيدة التصميم، الطالب في فهم مجال الخطاب وإنشائه (Douglas, 2000)، وهذا يتضمن: موضوع الاستجابة، والجمهور المتلقي، والوظيفة التواصلية ذات الصلة. ونتيجة لذلك، فإن وظيفة تحليل لغة الطالب في هذه المهام، ليست ذات نهاية مفتوحة على الإطلاق أو مشكلة عامة في التحليل اللغوي، بل مشكلة يمكن وصفها بأساليب تجريبية من علماء القواعد الوظيفية (مثل Halliday, 1994) التي يستطيع المتعلمون من خلالها – على اختلاف قدراتهم – استخدام مواردهم اللغوية المحدودة لتحديد المعنى ضمن سياق محدد بدقة. وهذا هو نوع العمل الذي ظل يمارس في اللغويات الحاسوبية والذكاء الاصطناعي لما يزيد على ٤٠ عامًا. إلا أن مثل هذه البحوث اهتمت بأداء المتحدثين الأكفاء في سياقات محددة بدلاً من مستويات أداء المتعلمين.

إن العمل نحو فهم أفضل لهذه العوامل الثلاثة في تشكيل أداء الاختبار يبدو خطوة ضرورية لإجراء التقييمات الآلية في وقتنا هذا، وكذلك أعمال التقييم الذكية الأكثر تقدمًا في المستقبل. ويصف بينيت (1999b) " التقييم الذكي" (ص ٩٩) بأنه تكامل لثلاثة اتجاهات بحثية في خط واحد، وهي: الاختبارات ذات الاستجابات المنشأة (الكتابة)، والذكاء الاصطناعي، والقياس القائم على الأنموذج، وفي ذلك يقول:

> يتمثل التكامل في إنتاج أساليب التقييم التي تتألف من المهام الأقرب إلى المشكلات المعقدة التي تُواجه في بيئات العمل والبيئات الأكاديمية، إذ تُسجل هذه المهام آليًا بأساليب تحاكي سلوك أحد الخبراء، مع توفير تقييم قائم على مقياس ذي مدلول جزئي لأغراض إجمالية، إضافة إلى وصف نوعي مصمم لنقل المعلومات التوجيهية المفيدة. أما الآليات الدافعة التي تقوم عليها هذه المهام، وأسلوب التقييم فيها ... فهي نماذج القياس (الراسخة في علم النفس المعرفي) التي قد تحدد الخصائص المطلوبة للبنود، وتلك البنود التي ينبغي التعامل معها بعد أخذها من

عدد كبير من البنود، إضافة إلى كيفية جمع الاستجابات مع البنود للوصول إلى استنتاجات أكثر شمولاً، وكيفية التعاطي مع جوانب عدم التيقّن (ص ٩٩).

ومن أجل دعم مثل هذا البحث في اختبارات اللغة الثانية، ثمة حاجة إلى تعاون وثيق بين جوانب التقييم اللغوي وبين اكتساب اللغة الثانية – على الأقل – في المجالات الثلاثة سالفة الذكر. وهذه الروابط ليست مستحيلة في مجمل الأمر (انظر مثلاً Brown, Hudson, Norris & Bonk, 2002) على الرغم من الحقيقة التي تفيد اختلاف مجالي اللغويات التطبيقية عن بعضهما، بل إن المكونات العاملة الأكثر أهمية مثل وحدات التحليل، تتميز بالاختلاف، فيما يتحدث الباحثون في التقييم عن مهارات القراءة والكتابة والاستماع والمحادثة، أما الباحثون في اكتساب اللغة الثانية فيتحدثون عن "نظام الزمن والوجهة" (the tense and aspect system) والنفي، أو الطلب المهذب على سبيل المثال. إن المحاولات الرامية إلى توسيع مفاهيم القياس لتشمل البيانات المعقدة محل الاهتمام في اكتساب اللغة الثانية (انظر مثلاً Chapelle, 1996، و Bachman & Cohen, 1998) لا بد من تطويرها لتصبح ضمن برنامج بحثي أكثر انتظامًا من الآن، وهذا ما يراه كل من نوريس وأورتيجا (Norris and Ortega, 2003) عند استعراضهما لممارسات القياس في مجال اكتساب اللغة الثانية: تحتاج أوساط اكتساب اللغة الثانية إلى "الإسهام في اتجاه شامل لجميع المراحل في عملية القياس، (من أجل أن) تصبح أكثر قدرة على تقديم تفسيرات ذات مغزى من الناحية النظرية بشأن المكونات، إلى جانب السعي إلى مراكمة المعرفة الجديرة بالاهتمام من الناحية العلمية" (ص ٧٤٩).

وحتى تفيد هذه المعرفة – في نهاية المطاف – في تصميم اختبارات اللغة المحوسبة، لا بد للباحثين المعنيين باختبارات اللغة أن يميزوا بين معرفة اكتساب اللغة الثانية المرتبطة بالجدل النظري في مجال اكتساب اللغة الثانية من جانب، وتلك المعرفة

التي يمكن أن تفيد عملية التقييم من جانب آخر. وقد قورن بين مجال اكتساب اللغة الثانية عمومًا وبين أكثر المجالات تخصصًا: "اكتساب اللغة الثانية الموجه"؛ إذ يركز الأخير على ظروف الأداء اللغوي واكتساب اللغات، وهي الظروف المتعلقة بالقرارات التعليمية. ومن الموضوعات ذات الصلة في هذا الشأن، البحث الهادف إلى زيادة المعرفة المعتمدة على البحث في المهام التربوية (انظر مثلا Crookes & Gass, 1993 وBygate, Skehan & Swain, 2003)، فهذا هو المجال الذي يولي اهتمامًا مناسبًا يؤهله لأن يكون أساسًا للبحث الافتراضي والتجريبي في تعلم اللغة بمساعدة الحاسوب (CALL). وبالمثل، ثمة حاجة إلى مجال يمكن تسميته باسم "اكتساب اللغة الثانية المقيَّم" ليركز على جوانب من الاكتساب يمكن ملاحظتها تجريبيًا في الأداء ضمن ظروف معرفية خاصة وفي سياقات محددة.

تعليم اللغات

أشرنا في الفصل الأول إلى ضرورة أن يكون لدى معلمي اللغات فهماً عميقاً لعملية التقييم؛ لأنهم يساعدون المتعلمين في وضع استراتيجيات التقييم الذاتي، ويختبرون المتعلمين داخل صفوف الدراسة، ويختارون أو يعدون الاختبارات لبرامج تعليم اللغات، ويتولون أيضًا عملية تهيئة المتعلمين لأداء الاختبارات خارج نطاق الصفوف وخارج برامج تعليم اللغات. إلا أنه ربما تكون أكثر الرؤى تحفيزًا في عملية التقييم اللغوي في الصف، هي الفرصة المتاحة لعمليات التقييم لمساعدة الطلاب والأخذ بيدهم نحو الأفضل وكذا الاعتماد على أنفسهم.

قدمنا في الفصل الثاني وصفًا لبعض الأمثلة لبرامج تعلم اللغة بمساعدة الحاسوب، مثل: برنامج "Longman English Interactive" وبرنامج "Market Leader" اللذين تضمنا اختبارات وتعقيبات مقدمة للمتعلمين ضمن الباقات التوجيهية. وتكمن الفكرة من وراء ذلك، في أنه إذا اطلع المتعلمون بانتظام على مدى جودة معارفهم

وتقدمهم في ظل تعاملهم مع التوجيه المقدم لهم، فسيتمكنون من تحديد الخيار الأفضل بشأن دراستهم، ويصبحون أكثر اعتمادًا على أنفسهم في نهاية المطاف. وبالطبع، إذا كان لابد من وضع هذه القدرات موضع التنفيذ، فلا بد أن يكون لدى معدي برامج الدورات فهماً واضحاً لمبادئ التقييم لدى تقديم خدمات التعلم.

أضف إلى ذلك، لابد أن يعرف المعلمون أن عمليات التقييم لا تُستخدم بالضرورة لأن الناشرين قد أنتجوها. ولذا يحتاج المعلمون إلى الوقوف على إمكانات التقييم المحوسب إذا استقروا على تقديمه للمتعلمين. وبمعنى آخر، نرى أن الفصل الكائن بين معدي اختبارات اللغة ومعلمي اللغة، إنما هو دلالة سلبية إذا ما نظرنا إلى الهدف من توسيع استخدامات التقييم بطرق جديدة مبتكرة. وفي هذا الشأن يرى كل من ستوينوف وتشابيل (Stoynoff and Chapelle, 2005) أنه من الضروري تجاوز هذا الفصل، وأن يلم معلمو اللغات بعمليات التقييم من أجل تحديد الاختبارات وإعدادها للمتعلمين. وينتج عن احتمالات الاستخدامات الجديدة لعمليات التقييم المتكاملة مع مواد التعليم المحوسبة حافز إضافي لتوعية المعلمين بأعمال التقييم. وفي هذا الصدد، يمكن النظر إلى اختبارات اللغة المحوسبة، على أنها فرصة قوية للتأثير الإيجابي ضمن العملية التي تتجاوز أنواع التجميع التي طالما كانت محل اهتمام البحث الحديث (انظر مثلاً: Cheng, Watanabe & Curtis, 2004).

الاتجاهات المستقبلية

تطرح اختبارات اللغة المحوسبة حالياً موضوعات لابد من تقصيها إذا ما أردنا لها التطور بما يكفي لتصبح جزءًا من الطفرة التي تشهدها جوانب التقييم؛ فالتقنيات الحالية تمثل زيادة مفرطة في المعطيات الإيجابية لمعدي الاختبارات، فمن تقديم الاختبارات عن بعد، والتحكم الدقيق في التوقيت، ومدخلات الوسائط المتعددة للطالبين، ناهيك عن

معالجة اللغات الطبيعية ووضع النماذج الطلابية. وقد أصبحت أدوات إعداد الاختبارات معقدة للغاية، فإذا كان من المنتظر أن يستخدم معدو الاختبارات هذه الأدوات، فلا بد من توجيه البحث بأجندة واضحة في مجال اللغويات التطبيقية مدعومًا بمعارف من مجالات أخرى متنوعة.

مشروع متنوع المجالات المعرفية

على الرغم من دخول القضايا الدقيقة التي تثيرها التقنية ضمن نطاق المشكلات التي ينبغي لعلماء اللغويات التطبيقية الإلمام بكيفية التعامل معها، فإنه لا بد من استحداث أدوات للتعامل مع هذه المشكلات واختبارها في مجال يصبح فيه التعاون في مختلف المجالات المعرفية ممكنًا في مثل هذه القضايا. ولقد تناولنا في الفصل الرابع أدوات لتوليف الاختبارات مثل: أنظمة WebCT وRespondus وHot Potatoes وQuiz Center وBlackboard وQuestionmark. وإذا كانت هذه الأنظمة توفر الأدوات المساعدة لإعداد الاختبارات بوجه عام، فقد رأينا أنها لا تتضمن ظواهر محددة تتعلق باللغة، وإنما هي في الظاهر استمداد من استجابات لغوية شفهية ووسيلة لتحليل الاستجابات المدخلة بهدف تحديد درجة جزئية على أساس منطقي. من جانب آخر، لا بد من تطوير أدوات البرامج الخاصة بمتطلبات اختبارات اللغة الثانية تطويرًا تراعى فيه أوجه القصور في الأدوات الحالية المستخدمة في اختبارات اللغة.

وقد تضافرت جهود المعلمين ومعدي البرمجيات لإعداد اختبارات فردية باستخدام أساليب توليفية ذات أهداف عامة أو لغات برمجة بعينها لمدة تزيد على ٣٠ عامًا (انظر مثلاً Boyle, Smith & Eckert, 1976)، لكن إذا أردنا تطوير هذه القاعدة من الخبرات والمعارف على نحو تراكمي، فلا بد من توفير أدوات توليفية ذات جودة عالية في أوساط اللغويات التطبيقية. كما يتطلب استحداث مجموعة كبيرة من الأدوات الملائمة نخبة من المتخصصين في مجالات هندسة البرامج، ومتخصصي التقييم اللغوي،

والمصممين على أقل تقدير. ودون أدوات البرامج الأساسية، التي يستطيع طلاب الدراسات العليا استخدامها لزيادة معارفهم إزاء إجراء الاختبارات، يبدو أن حدود المناقشة – حول تصميم الاختبارات – سيظل مقصورًا على مستوى توقعات غير متخصصة عما قد يكون مجديًا وما قد يحظى بالاهتمام، فعلى سبيل المثال، يعد مشروع Dutch CEF Construct Project (انظر Alderson, Figueres, Kuijper, Nold, Takala & Tardieu, 2004) مثالاً على جزء من برنامج حاسوبي يهدف إلى مساعدة مصممي الاختبارات على إعداد مهام اختبارية وتحليلها وفقًا لإطار قائم على المكونات (مثل الإطار المرجعي الأوروبي المشترك للغات). وقد اتخذت مشاريع مثل DIALANG بعض الخطوات الرامية إلى إعداد مجموعة متنوعة من أنواع البنود، ويرجى لها أن تتمكن من استحداث أدوات توليفية من شأنها مساعدة المولفين الآخرين على تجريب هذه الأدوات (Alderson, 2005).

كما يمكن العثور على لمحات أخرى مما قد يكون متاحاً في البرامج المتقدمة للبحوث التي تناولت التعلم الذكي للغات بمساعدة الحاسوب (انظر مثلاً Chanier, Pengelly, Twidale & Self, 1992)، علمًا بأنها نتاج بحوث تأخذ بضروب متنوعة من علوم ومجالات معرفية شتى. يتضمن هذا النظام العناصر المشابهة لما وصفه بينيت على أنه ضروري للتقييم الذكي – تحليل استجابات المتعلمين الإنشائية (الكتابية)، والأنموذج الطلابي الخاضع للتحديث على أساس تحليل استجابات الطلاب، ونظام متخصص ينتقي التوجيهات للمتعلم ليكتسب من خلالها مزيداً من المعلومات. وليس من الواضح إلى أي مدى تعمل أي من مفاهيم القياسات في هذا النظام الذي لا يهدف على وجه الخصوص إلى التقييم. لكن لب الأمر هو أن هذه النظم المتقدمة تُكتشف في مجالات أخرى، وأن إتاحتها للباحثين في التقييم اللغوي، يتطلب أدوات توليف أكثر تقدماً لإعداد الاختبارات من تلك التي يجدها المرء في إعداد الاختبارات الصفية.

أجندة خاصة باللغويات التطبيقية

على الرغم من الحاجة إلى الاستفادة من الخبرة في مجالات القياس التربوي واللغويات التطبيقية واكتساب اللغة الثانية، والتقنية، فإن ثمة حاجة إلى وضع الأجندة البحثية وتطويرها استنادًا إلى مباعث القلق التي تساور علماء اللغويات التطبيقية والمتعلقة بالتقييم. وحتى فيما يخص مجال اللغويات التطبيقية، فإن ذلك يتطلب تحديد موقف واضح وجلي من التقنية. وبناء على تحليل الاتجاهات المتبع في وضع أجندات البحث والممارسات المعمول بها في التقييم اللغوي، يحدد تشابيل (Chapelle, 2003) ثلاثة اتجاهات يعمل بها، كما هو موضح بالجدول رقم (٦,١).

الاتجاه الأول هو "اتجاه النفق" – كما في معنى "رؤية النفق" – في إشارة إلى الاستعارة التي أوردها براون ودوجويد في دارستهما (Brown and Duguid, 2000) حيث وصفا المنشغلين بالتقنية في مختلف مناحي المجتمع، بأنهم يتحركون بعقلية ذات اتجاه واحد نحو تحقيق أهداف تتسم بالسرعة والكفاءة، دون الالتفات إلى أي شيء آخر. ودائمًا تُفسر التقنية في التقييم اللغوي باعتبارها وسيلة لبناء المزيد من الاختبارات الجيدة. فإذا كانت الجودة هي الهدف، فإن النتائج المنشودة ستكون اختبارات أقصر وأكثر ملاءمة؛ أي أن النظرية التي يمكن لمعدي الاختبارات طرحها ستكون: إن الاختبار المحوسب يمكنه أن يؤدي المهمة نفسها التي تؤديها الاختبارات الأخرى، لكنه أسرع وأقل تكلفة.

في حين ينظر "اتجاه المقارنة" المتبع في اختبارات اللغة المحوسبة إلى *التقنية* بعين الشك، ومن ثم تتمثل مشكلة البحث في أن يميز الاختلافات التي بين الاختبارات المحوسبة وغيرها من الأنواع الأخرى. ويمكن إجراء مثل هذه التحليلات على مستوى الأداء في الاختبار ككل، أو حتى دراسته على مستوى الأداء في البند الواحد. غير أن ما يؤكد هذه النظرية هو الرؤية التي تفيد بأن عدم استخدام التقنية هي الحالة العادية، ومن ثم تظهر المشكلة في تحديد الاختلاف الذي تضيفه التقنية. وبناء على ذلك، يسعى

كلا الاتجاهين (اتجاه النفق واتجاه المقارنة) إلى تحقيق أهداف جديرة بالاهتمام. أما في اللغويات التطبيقية، فمن ذا الذي يقول بأنه لا حاجة لنا في اختبارات أكثر كفاءة وملاءمة؟ ومن ينكر قيمة الفهم الأفضل لكيفية تأثير التقنية في ظروف الأداء ومن ثم في نتائج الاختبارات؟ إلا أنه في حين توافق هاتين الرؤيتين مع مجال اللغويات التطبيقية والتقييم اللغوي، فإن كلا منهما محدود من حيث القدرات التي تؤهله لإحداث طفرة في التقييم اللغوي على النحو الذي وصفه بينيت.

الجدول رقم (٦,١). الافتراضات بشأن التقنية ونتائج اتجاهات كل من: النفق، والمقارنة، والابتكار.
(Chapelle, 2003، ص ١٧٩).

الاتجاه	الافتراضات حول دور التقنية في التقييم	النتائج
النفق	فاعلة	اختبارات قصيرة، وتصحيح آلي، وتقديم النتائج لاستخدامات الاختبارات الحالية.
المقارنة	ينبغي النظر إليها بعين الشك	أنواع مختلفة من الاختبارات للأغراض الاختبارية الحالية؛ والمعرفة المتعلقة بكيفية تأثير التقنية في الاختبارات التقليدية عند تقديمها على شبكة الإنترنت.
الابتكار	ينبغي النظر غليها كمورد	أنواع مختلفة من الاختبارات واستخدامات جديدة للاختبارات؛ والمعرفة المتعلقة بالتفاعل بين التقنية ومجموعة متنوعة من قضايا التقييم.

قد تكمن الطفرة في الاتجاه الابتكاري الذي يعتمد على التقنية بكونها موردًا لاستكشاف مجموعة متنوعة من قضايا التقييم. وهنا يرى تشابيل (Chapelle, 2003) أن مثل هذا الابتكار يستلزم إعداد مجموعة متنوعة من الاختبارات واستخداماتها التي لا يمكن توافرها دون الاستعانة بالتقنية. ولتحقيق ذلك، ثمة حاجة إلى بعض البرامج الخاصة باختبارات اللغة التي سبق ذكرها، إلى جانب استخدام التقنية لتطوير المعارف

حول التفاعل بين التقنية ومختلف قضايا التقييم. أما إيفا بيكر الباحثة في مجال القياس التربوي، فترى أن أجندة الابتكار هذه تقع في قلب الطفرة المستقبلية التي من المنتظر أن تؤدي فيها التقنية دورًا مهماً، وتكمن رؤيتها في أن "التقنية المطبقة لخدمة فهم التعليم الذي نحتاج إليه، تساعدنا على إصلاح ما هو غير قابل للإصلاح حالياً أي: مشكلة الصدق المتأصلة في قلب نظامنا الاختباري" (Baker, 1998، ص ٢٢).

خاتمة

تماثل الاقتراحات والقضايا التي تظهر ضمن نطاق الاتجاه الابتكاري إزاء اختبارات اللغة المحوسبة – في كثير من جوانبها – الاقتراحات والقضايا التي طرحها المبتكرون من الرواد الأوائل في هذا المجال على مدار العقد المنصرم. فقد اقترح كانيل (Canale, 1986) استخدام تقنيات التوجيه الذكية لنمذجة معارف المتعلمين وإفادة العملية التعليمية. أما آلدرسون (Alderson, 1988) فقد أشار إلى أن الحاسوب يمكن أن يستعين بقواعد اللغة لتحليل استجابات المتعلمين الإنشائية (المكتوبة)، في حين تساءل كوربل (Corbel, 1993) عن احتمالات قدرة التقييم الذكي، واختبارات اللغة المحوسبة في المساعدة في التقييم الذاتي، وتقوية الروابط بين التقييم ومجالات اللغويات التطبيقية الأخرى بالاستعانة بالتقنية. وعلى الرغم من التطورات الحديثة في عمليات التقييم الآخذة بجديد التقنية، فإننا لا نستطيع الإعلان عن أية تحولات هائلة في مجال التقييم جرّاء البحث المنتظم في هذه المجالات. وقد تسهم تقنية الحاسوب مستقبلاً في إحداث تحول جوهري في البحث والممارسة في مجال التقييم اللغوي، بيد أن ذلك يتطلب نوعًا من البحث الذي ينفذ إلى صميم التعقيد القائم في هذه القضايا، عابرًا بذلك الحدود الوهمية القائمة بين: التقييم، واللغة، والتقنية بهدف تطوير المسارات التي تعمل على تحقيق أهداف اللغويات التطبيقية.

المراجع

Ackerman, T. (1994). Creating a test information profile for a two-dimensional latent space. Applied Psychological Measurement, 18(3), 257-75.

ACT Inc. (2004). Webpage: http//www.act.org/esl/sample/gammar4.html. Accesses 11 January 2005.

Alderson, J.C. (1980). Scoring procedures for use on cloze tests. In C.A. Yorio, k. perkins & J. Schacter (eds.), On TESOL '79 (pp. 193-205). Washington, DC: TESOL publications.

Alderson, J.C. (1988). Innovations in language testing: Can the microcomputer help? Special Report No I Language Testing Update. Lancaster, UK: University of Lancaster.

Alderson, J.C. (1990). Learner-centered testing through computers: Institutional issues in individual assessment. In J. de Jong & D. K. Stevenson (eds.) Individualizing the assessment of language abilities (pp. 20-7). Clevedon, UK: Multilingual Matters.

Alderson, J.C. (2000). Assessing reading. Cambridge: Cambridge University Press.

Alderson, J.C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment.* London: Continuum.

Alderson, J.C., Figueras, N., Kuijper, H., Nold, G., Takala, S. & Tardieu, C. (2004). *The development of specifications for item development and classification within the Common European Framework of Reference for Language: Learning, Teaching, Assessment, Reading and Listening.* Final Report of the Dutch CEF construct Project. Mimeo.

Alderson, J.C. & Hamp-Lyons, L. (1996). TOEFL preparation courses: a study of washback. *Language Testing*, 13 (3), 280-297.

Alderson, J.C., percsich, R. & Szabo, G. (2000). Sequencing as an item type. *Language testing*, 17 (4), 423-47.

Alderson, J.C. & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115-29.

Alessi, S. M. & Trollop, S. R. (1985). *Computer-based instruction: Methods and development*, Englewood Cliffs, NJ: prentice-Hall.

Almond, R. G., Steinberg, L. S. & Mislevy, R. J. (2002). Enhancing the design and delivery of assessment systems: A four-process architecture. *Journal of*

Tecnonolgy, Learning and Assessment, 1(5), Available from http://www.jtla.org.

American Educational Research Association, American Psychological association, National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: American University Press.

Bachman, L.F. (1991). What does language testing have to offer? *TESOL Quarterly*, 25 (4), 671-704.

Bachman, L.F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford University Press.

Bachman, L.F. (2000). Modern language testing at the turn of the century: assuring that what we count counts. *Language Testing*, 17(1), 1-42.

Bachman, L.F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.

Bachman, L.F. (in press). Building and supporting a case for test use. *Language Assessment Quarterly*.

Bachman, L.F. & Cohen, A. D. (eds.) (1998). *Interfaces between second language acquisition and language testing research*. Cambridge: Cambridge University Press.

Bachman, L.F., Lynch, B. & Mason, M. (1995.). investigating variability in tasks and rater judgments in a performance test of foreign language speaking. *Language Testing*, 12(2),238-58.

Bailey, K. (1996). Working for washback: A review of the washback concept in language testing, *Language Testing*. 13(3), 257-79.

Baker, E. L. (1998). *Understanding educational quality: Where validity meets technology. The fifth annual William Angoff Memorial Lecture*. Princeton, NJ: Educational Testing Service.

Baker, E. L. (1989). Computer technology in test construction and processing. In R. L. Linn (Ed.) *Educational measurement*, 3rdj edn (pp. 409-28). NY: Macmillan publishing Co.

Bajar, I. I. (1985). Speculations on the future of test design. In S. Embretson (ed.), *Test design-Developments in psychology and psychometrics* (pp. 279-94). Orlando: Academic Press.

Bajar, I. & Braun, H. (1994). On the synergy between assessment and instruction: Early lessons from computer-based simulations. *Machine-Mediated Learning*, 4(1)< 5-25.

Bennett, R. E. (1993). On the meanings of constructed. In R. E. Bennett & W. C. Ward (eds.), *Construction versus responses choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (pp. 1-27). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bennett, R. E. (1999a). How the Internet will help large-scale assessment reinvent itself. *Education Policy Analysis Archives*, 9(5), 1-25.

Bennett, R. E. (1999b). using new technology to improve assessment. *Educational Measurement: Issues and Practice*, 18(3), 5-12.

Bennett, R. E. (2001). How the Internet will help large-scale assessment reinvent itself. *Education Policy Analysis Archives*, 9(5), 1-27. Online: http//epaa.asu.edu/epaa/v9n5.html (Accessed 11 January 2005).

Birenbaum, M. & Tatsuoka, K. K. (1987). Open-ended versus multiple-choice response formats ñ It does made a difference for diagnostic purposes. *Applied Psychological Measurement*, 11(4). 385-95.

Blackboard. (2004). Blackboard Learning System. Blackboard, inc. http://blackdoard.com (Accessed11 January 2005).

Blais, J-G & Laurier, M. D. (1993). The dimensionality of a placement test from several analytical perspectives. *Language Testing*, 10(2), 72-98.

Boyle, T. A., Smith, W. F. & Eckert, R. G. (1976). Computer-mediated testing: A branched program achievement test. *Modern Language Journal*, 60, 428-40.

Bradlow, E. T., Wainer, H. & Wang, X. (1999). A Bayesian random effects model for testlets. *Psychometrika*, 64, 153-68.

Brown, A. & Iwashita, N. (1996). The role of language background in the validation of a computer-adaptive test. *System*, 24(2), 199-206.

Brown, J. D. (1997). Computer in language testing: Present research and some future directions. *Language Learning & Technology*, 1(1), 44-59.

Brown, J. D. Hudson, T., Norris, J. & Bonk, W. J. (2002). *An investigation of second language task-based performance assessments, Technical Report* #24, Honolulu: Second Language Teaching & Curriculum Center, University of Hawaii at Manoa.

Brown, J. S. & Duguid, P. (2000). *The social life of information*. Boston, MA: Harvard Business School Press.

Bruce, B. C. & Hogan, M. P. (1998). The disappearance of technology: Toward an ecological model of literacy. *In Handbook of literacy and technology: Transformations in a post-typographic world* (pp. 296-81). Mahwah, NJ: Lawrence Erlbaum Associates.

Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.

Burstein, J., Frase, L., Ginther, A. & Grant, L. (1996). Technologies for language assessment. *Annual Review of Applied Linguistics*, 16, 240-60.

Burston, J. & Monville-Burston, M. (1995). Practical design and implementation considerations of a computer-adaptive foreign language test: The Monash/Melbourne French CAT. CALICO *Jounal*, 13(1), 26-46.

Bygate, M., Skehan, P. & Swain, M. (eds.) (2003). *Researching pedagogic tasks: second language learning teaching, and testing*. Harrow, England: Longman.

Canale, M. (1986). The promise and threat of computerized adaptive assessment of reading comprehension. In C. stansfield (ed.), *Technology and language testing* (pp. 30-45). Washington, DC: TESOL publications.

Canale, M. (1987). Language assessment: the method is the message. In D. Tannen & J. E. Alatis (eds.), *The interdependence of theory, data, and application* (pp. 246-62). Washington, DC: Georgetown University Press.

Carr, N., Pan, M. & Xi, X. (20020. Construct refinement and automated scoring in web-based testing. Paper presented at the Language Testing Research Colloquium, Hong Kong, December 2002.

Carroll, J. B. (1968). The psychology of language testing. In A. Davies (ed.), *Language testing symposium: A Psycholinguistic Perspective* (pp. 46-69). Oxford: Oxford University Press.

Chalhoub-Deville, M. (ed.) (1999). *Development and research in computer adaptive language testing.* Cambridge: University of Cambridge Examinations Syndicate/Cambridge university Press.

Chalhoub-Deville, M. & Deville, C. (1999). Computer adaptive testing in second language contexts. *Annual review of Applied Linguistics*, 19, 273-99.

Chanier, T., Pengelly, M. Twidale, M. & Self, J. (1992). Conceptual modelling in error analysis in computer-assisted language learning systems. In M. L. Swartz & M. Yazdani (eds.), *Intelligent tutoring system for foreign language learning* (pp. 125-50). Berlin: Springer-Verlag.

Chapelle, C. A. (1993). Issues in computer-assisted analysis for one-word test responses. *Assessment-Transactions of the 1993 CALICO Symposium* (pp. 28-30). Durham, NC: CALICO.

Chapelle, C. A. (1996). Validity issues in computer-assisted strategy assessment. *Applied Language Learning*, 7(1), 47-60.

Chapelle, C. A. (1999). Validity in language assessment. *Annual Review of Applied Linguistics*, 19, 254-72.

Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching testing, and research.* Cambridge: Cambridge University Press.

Chapelle, C. A. (2003). *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology.* Amsterdam: John Benjamins Publishing.

Chapelle, C. A. & Abraham, R. (1990). Cloze method: What difference does it make? *Language Testing*, 7(2), 121-45.

Chapelle, C. A., Enright, M. K. & Jamieson, J. (forthcoming). Challenges in developing a test of academic English. In C.A. Chapelle, M. K. Enright & J. Jamieson (eds.), *Building a validity argument for TOEFL.* Mahwah, NJ: Lawrence Erlbaum Associates.

Chapelle, C. A. & Jamieson, J. (2001). *Longman English Assessment.* White Plaines, NY: Pearson Education.

Chapelle, C. A. & Jamieson, J. &Hegelheimer, V. (2003). Validation of a web-based ESL test. *Language Testing*, 20(4), 409-93.

Cheng, L., Watanabe, Y. & Curtis, A. (eds.) (2004). Washback in language testing Research contexts and methods. Mahwah, NJ: Lawrence Erlbaum Associates.

Choi, I.-C., Kim, K. S. & Boo, J. (2003). Comparability of a paper-based language test and a computer-based language test. *Language Testing*, 20(3), 295-320.

Clark, J. L. D. (1989). Multipurpose language tests: Is a conceptual and operational synthesis possible? In J.E. Alatis (ed.), *Georgetown University Round Table on Language and Linguistics. Language teaching, testing, and technology: Lessons from the past with a view toward the future* (pp. 206-15). Washington, DC: Georgetown University Press.

Cohen, A. (1998). Strategies and processes in test-taking and SLA. In L.F. Bachman & A. D. Cohen (eds.), *Interfaces between second language acquisition and language testing research* (pp. 90-111). Cambridge University Press.

Cole, N. (1993). Comments on Chapters 1-3. In N. Fredreiksen, R.J. Mislevy & I. I. Bejar (eds.). *Test theory for a new generation of tests* (pp. 72-7). Hillsdale, NJ: Lawrence Erlbaum Associates.

Coniam, D. (1996). Computerized dictation for assessing listening proficiency CALICO *Journal*, 13(2-3), 73-85.

Coniam, D. (1998). Interactive evaluation of listening comprehension: How the context may help. *Computer-assisted Language Learning*, 11(1) 35-53.

Coniam, D. (2001). The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study. *System*, 29, 1-14.

Corbel. C. (1993). *Computer-enhanced language assessment*. In f. Brindley (Series ed.) Research Report Series 2. National Center for English Language Teaching and Research, Macquarie University, Sydney, Australia.

Council of Europe (2001). *Common European Framework of Reference: Learning teaching, assessment*. Cambridge University Press.

Crookes, G & Gass, S. M. (eds.) (1993). *Tasks and Language Learning: Integrating theory and practice*. Philadelphia: Multilingual Matters.

Crystal, D. (2001). *Language and the Internet*. Cambridge: Cambridge University Press.

Davidson, F. (1996). *Principles of statistical data handling*. Thousand Oaks, CA: Sage Publications.

DIALANG. (2001). Online: http://dialang.org/info/english/index.htm (Accessed 11 January 2005).

DIALANG Beta Test. (2002). DIALANG Beta Version. Jyvaskyla, FI: DIALANG Project, Jyvaskyla University.

Discovery School. (2004). *Quiz Center*. Discoveryschool.com: http://school.discovery.com/customclassroom/about.html (Accessed 11 January 2005).

Douglas, d. (1998). Testing methods in context-based second language research. In L. G. Bachman & A. D. Cohen (eds.), *Language testing-SLA interfaces* (pp. 141-55). Cambridge: Cambridge University Press.

Douglas, d. (2000). *Assessing language for specific purposes*. Cambridge: Cambridge University Press.

Dunkel, p. (eds.) (1991). Computer-assisted language learning testing: Research issues and practice. New York: Newbury House.

Dunkel, p. (1999). Research and development of a computer-adaptive test of listening comprehension in the less commonly-taught language Hausa. In M. Chalhoub-Deville (ed.), *Development and research in computer adaptive language testing* (pp. 91-121). Cambridge: Cambridge University Press.

Educational Testing Service (1986). *Test of Spoken English*. Princeton, NJ: Educational Testing Service.

Educational Testing Service (1999). TOEFL *Sampler: An Introduction to Taking the TOEFL Test on Computer*. CR ROM. Princeton, NJ: Educational Testing Service.

Educational Testing Service (2005a). *Criterion*. Online: http://www.ets.org/criterion/index.html (Accessed 21 September 2005).

Educational Testing Service (2005b). *criterion tour*. Online: http://www.ets.org/MMedia/Products/Criterion.tour2/open.swf (Accessed 21 September 2005).

Embretson, S. (1983). Construct validity: Construct representation versus nomothetic span. Psychological Bulletin, 93(1), 179-97.

Embretson, S. (ed.). (1985). Test Design: Developments in psychology and psychometrics. Orlando, FL: Academic Press.

Enlight, A. B. (2004). *Test Station Manager*. Stockholm: http://www.enlight.net/site/Viewpage.action?siteNodeId=226&languageId=1&contentId=-1 (Accessed 11 January 2005).

Fulcher, G. (2000). Computers in language testing. In P/ Brett and G. Motteram (eds.), *A special interest in computers* (pp. 93-107). Manchester: IATEFL Publications.

Fulcher, G. (2003). Interface design in computer-based language testing *Language Testing*, 20(4), 384-408.

GeoHive (2005). GeoHive: global Statistics. Online: http://www.geohive.com/index.php (Accessed 19 March 2005).

Green, B. F. (1988). Construct validity of computer-based language tests. In H. Wainer & H. I. Braun (eds). *Test validity* (pp. 77-1030. Hillsdale, NJ: Lawrence Erlbaum Associates.

Green, B. F. Book, R. D., Humphreys, L. B., Linn, R. L. & Reckase, M. D. (1984). Technical guidelines for assessing computer adaptive tests. *Journal of Educational Measurement*, 21, 374-60.

Gruba, P. (1997). Exploring digital video material. In r. Debski, J. Gassin & M. Smith (eds.), *Language learning through social computing* (pp. 109-40). Parkville, Vic: Applied Linguistics Association of Australia.

Haas, C. (1996). *Writing technology: Studies on the materiality of literacy*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hagen, L. K. (1994). Constructs and measurement in parameter models of second language acquisition. In E. E. tarone, S. M. Gass & A. D. Cohen. *Research methodology in second-language acquisition* (pp. 61-87). NJ: Lawrence Erlbaum Associates.

Half-baked. (2004). Hot Potatoes. Version 6.0.3 Half-baked Software, inc. http://web.uvic.ca/hrd.halfbaked/ (Accessed 11 January 2005).

Halliday, M. A. K. (1994). *An introduction to functional grammar*, 2nd, ed. London: Edward Arnold.

Halliday, M. A. K. & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective.* Oxford: Oxford University Press.

Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of item response theory*. London: Sage publications.

Hegelheimer, V. & Chapelle, C. A. (2000). Methodological issues in research on lerner-computer interactions in CALL. *Language Learning and Technology*, 4(1), 41-59. Online: http://llt.msu.deu/vol4num1/hegchap/default.html.

Henning, G. (1987). *A guide to language testing: Development, evaluation research.* Cambridge, MA: Newbury House.

Henning, G., Anbar, M., Helm, c. & D'Arcy, S. (1993). Computer-assisted testing of reading comprehension: comparisons among multiple-choice and open-ended scoring methods. In D. Douglas & C. A. Chapelle (eds.), *A new decade of language testing research* (pp. 123-31). Alexandria, VR: TESL Publications.

Higgins, J.J., Lawrie, A. M. & White, A. G. (1999). Recognising coherence; The use of a text game to measure and reinforce awareness of coherence in text. *System*, 27, 339-49.

Hinkel, E. (2003). Simplicity without elegancy: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, 37, 275-302.

Holland, V. M. (1994). Intelligent tutors for foreign language: how parsers and lexical semantics can help learners and assess learning. In R. m. Kaplan & J. C. Burstein (eds.). *Proceedings of the Educational Testing Service conference on natural language processing techniques and technology in assessment and education* (pp. 95-108). Princeton, NJ: Educational Testing Service.

Holland, V. M., Kaplan, J. & Sams, M. (eds.). (1994). *Intelligent language tutors: Theory shaping technology* (pp. 153-74). Hikllsdale, NJ: Lawrence Erlbaum Associates.

Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.

IBM Corporation. (2002). *They Also Served: An Album of IBM Special Products*, VOL. 1. Online: http://www.1.ibm.com/ibm/history/exhibits/specialprod1/specialprod1_9.html. (Accessed 11 January 2005).

Jemieson, J. & Chapelle, C. A. (1987). Working styles on computers as evidence of second language learning strategies. *Language Learning*, 37, 523-44.

Jemieson, J., Campbell, J., Norfleet, L. & Berbisada, N. (1993). Reliability of a computerized scoring routine for an open-ended task. *System*, 21(3), 3.5-22.

Jemieson, J., Kersch, I., Taylor, C. & Eignor, D. (1999). *Designing and evaluating a computer-based TOEFL tutorial.* TOEFL Research Report 62.

Jemieson, J., Taylor, C., Kersch, I. & Eginor, D. (1998). Designing and evaluating a computer-based TOEFL tutorial. *System*, 26, 485-513.

Kane, M. t. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112, 527-35.

Kane, M. t. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38, 319-42.

Kane, M. Crooks, T. & Choen, A. (1999). Validating measures of performance. *Educational Measurement: issues and Practice, 18, (2), 5-17.*

Kaya-Carton, E., Carton, A. S. & Dandonoli, P. (1991). Developing a computer-adaptive test of French reading proficiency. In P. Dunkel (ed.), *Computer-assisted language learning and testing: Research issues and practice* (pp. 259-84) New York: Newbury House.

Kendrick, M. (1997). Cultural rage and computer literacy: A response to Theodore Roszak. *ALKI:The Washington Library Association Journal* 13.2. Online: http://www.wla.org/alki/jul97/roszak2.html (Accessed 11 January 2005).

Larson, J.W. & Madsen, H.S. (1985). Computer-adaptive language testing: Moving beyond computer-assisted testing. *CALICO Journal*, 2(3), 32-6.

Laurier, M. (1999). The development of an adaptive test for placement in French. In M. Chalhoud-Deville (ed.)., *Development and research in computer adaptive language testing* (pp.122-35). Cambridge: Cambridge University Press. Examinations Syndicate/ Cambridge University Press.

Laurier, M. (2000). Can computerized testing be authentic? ReCALL, 12(1) 93-104.

Kiskin-Gasparro, J.E. (1984). The ACTFL proficiency guidelines: A historical perspective. In. T.V. Higgs (ed.), *Teaching for proficiency: The organizing principle* (pp. 26-28). ACTFL Foreign Language Education Series, 15. Lincolnwood, IL: National Textbook Company.

Longman. (2002). *Business English: Market Leader Interactive.* Online: http://www.longman.com/ae/multimedia/programs/mil.htm# (Accessed 11 January 2005).

Lord, F.M. (1980). *Application of item response theory to practical testing problems.* Hillsdale. NJ: Lawrence Erlbaum Associates.

Louma, S. & Tarnanen, M. (2003) creating a self-rating instrument for second language writing: From idea to implementation. *Language Testing*, 20(4), 440-65.

Madsen, H. S. (1986). Evaluating a computer-adaptive ESL placement test, *CALICO Journal*, 4(2), 41-50.

Madsen, H. S. (1991). Computer-adaptive testing of listening and reading comprehension: The Brigham Young approach. In P. Kunkel (ed.), *Computer-assisted language learning and testing: Research issues and practice* (pp. 237-57). New York: Newbury House.

Malabonga, V., Kenyon, D. & Carpenter, H. (2002). Computer assisted rating: Reliability, efficiency and perceptions on the COPI. Paper presented at the Language Testing Research Colloquium, Hong Kong, 12-15 December.

Marty, f. (1901). Reflections on the use of computers in second language acquisition. *Studies in Language Learning*, 3(1), 25-53.

McNamara, T. (1996). *Measuring second language performance*. London: Longman.

Messick, S. (1989). Validity. In R. L. Linn (ed.), *Educational Measurement*, 3rd edn (pp. 13-103). NY: Macmillan.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessment. *Educational Researcher*, 23(2), 13-23.

Meunier, L. E. (1994). Computer adaptive language tests (CALT) offer a great potential for functional testing. Yet, why don't they? CALICO *Journal*, 11(4) pp. 23-39.

Mislevy, R. J. (1993). A framework for studying differences between multiple-choice and free-response test items. In R. E. Bennett & W. C. Ward (eds.), *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (pp. 75-106). Hillsdale, NJ: Lawrence Erlbaum Associates.

Molholt, G. & Presler, A. M. (1986). Correlation between human and machine ratings of Test of Spoken English passages. In C. W. Stansfield (ed.) *Technology and language testing* (pp. 111-28). Washington DC: Teachers of English to Speakers of Other Language.

Monaghan, W. & Bridgeman, B. (2005). *E-rater as a quality control on human raters. R & D connections*, April. Princeton, NJ: Educational Testing Service. Online: http://ftp.ets.org/pub/res/researcher/RD-Connections2.pdf.

Murray, D. (2000). Protean communication: The language of computer-mediated communications. *TESOL Quarterly*, 34(4), 397-421.

Noijons, J. (1994). Testing computer assisted language tests: Towards a checklist for CALT. *CALICO Journal*, 12(1), 37-58.

Norris, J. (in press). Using developmental sequences to estimate English grammar ability: Preliminary design and investigation of a wed-based test. *University of Hawaii Working Papers in Applied Linguistics.*

Norris, J. & Ortega, L. (2003). Defining and measuring SLA. In C. J. Doughty & M. H. Long (eds.), *The handbook of second language acquisition* (pp. 71-61). Malden, MA: Blackwell Publishing.

Oller, J. (1979). *Language tests at school.* London: Longman.

Ordinate Corporation. (2002a). validation summary for PhonePass SET-10 Menlo Park, CA: Author. Online: http://www.ordinate.com (Accessed 11 January 2005).

Ordinate Corporation. (2002b). PhonePass. Online: http://www.ordinate.com (Accessed 11 January 2005).

Otto, S. (1989). Assessment, articulation, accountability: New roles for the language lab. In J.E. Elatis (ed.), *Georgetown University sound table on languages and linguistics, 1989. Language teaching, testing, and technology: Lessons from the past with a view toward the future* (pp. 276-87). Washington, DC: Georgetown University Press.

Paltridge, B. (2001). *Genre and the language learning classroom.* Ann Arbor: University of Michigan Press.

Powers, D. E., Burstein, J. C., Choldorow, M., Fowles, M. E. & Kukich, K. (2001). *Stumping E-Rater: Challenging the validity of automated essay scoring. ETS RR 01-03.* Princeton, NJ: Educational Testing Service.

Powers, D. E., Fowles, M.E., Farnum, M. & Ramsey, P. (1994). Will they think less of handwritten essay if others wordprocess theirs? Effects on essay scores of intermingling handwritten and word-processed essays. *Journal of Educational Measurement,* 31(5), 220-30.

Promissor. (2004). *Quality Test Development Service.* Evanston, IL: Promissor, inc. http://www.promissor.com/ (Accessed 11 January 2005).

Purdue University. (2001). *Oral English Proficiency Test.* Version 1.9. West Lefayette, IN: Purdue Universtiy.

Questionmark. (2004). *Questionmark Perception.* Questionmark Corporation. htpp://www.questionmark.com/us/perception/index.htm (Accessed 11 January 2005).

Rassooln N. (1999). *Literacy for sustainable development in the age of information.* Clevedon, UK: Multilingual Matters.

Read, J. (2000). *Assessing vocabulary.* Cambridge: Cambridge University Press.

Reid, J. (1986). Using the Writer's Workbench in composition teaching and testing in C. stansfield (ed.), *Technology and language testing* (pp. 167-88). Washington, DC: TESOL publications.

Respondus. (2004). Version 2.0. Respondus, inc. http://www.respondus.com (Accessed 11 January 2005).

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge: Cambridge University Press.

Roever, C. (2001). Web-based language testing. *Language Learning & Technology* 5(2), 84-94.

Rost, M. (2003). *Longman English Interactive.* New York: Pearson Education.

Sawaki, Y. (2001). Comparability of conventional and computerized tests of reading in a second language. *Language Learning & Technology,* 5(2), 38-59.

Schaeffe, G. A. Reese, C. M. Steffen, M., McKinley, R. L. &Mills, C. N. (1993). *Field test of a computer-based GRE general test, ETS RR 93-07.* Princeton, NJ: Educational Testing Service.

Singleton, D. &Little, D. (1991). The second language lexicon: some evidence from university-level learners of French and German. *Second Language Research,* 7, 62-81.

Skehan, P. (1998). *A cognitive approach language learning.* Oxford: Oxford University Press.

Smith, f. (1984). The promise and threat of microcomputers for language learners in J. Hanscombe, R. Orem &B. Taylor (eds.), *On TESOL '83: The question of control* (pp. 1-18). Washington DC: Teachers of English to Speakers of Other Languages.

Stansfield, C. (ed.) (1986). *Technology and language testing.* Washington, DC: TESOL publications.

Steinberg, L., Thissen, D. &Wainer, H. (2000). Validity. In H. Wainer, N.J. Dorans, D. Eignor, R. Flaugher, B. F. Green, R. J. Mislevy, L. Steinberg & D. Thissen (2000). *Computer adaptive testing: A primer,* 2nd edn (pp. 185-229). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stevenson, J. & gross, S. (1991). Use of a computerized adaptive testing model for EDOL/bilingual entry/exit decision making. In. P. Kunkel (ed.), *Computer-assisted language learning and testing: Research issues and practice* (pp. 223-36). New York: Newbury House.

Stoynoff, S. & Chapelle, C. A. (2005). *ESOL tests and testing: A resource for teachers and administrators*. Alexandria, VA: TESOL Publications.

Susser, B. (2001). A defense of checklists for courseware evolution. *ReCALL,* 13(2), 261-76.

Taylor, C., Jamieson, J. & Eignor, D. (2000). Trends in computer use among international students. *TESOL Quarterly,* 34(3), 575-85.

Taylor, C., Kivsch, I., Eignor D. & Jamieson, J. (1999). Examining the relationship between computer familiarity and performance on computer-based language tasks. *Language Learning,* 49(2), 219-74.

Tung, P. (1986). Computer adaptive testing: Implications for language test developers. In C. Stansfield (ed.), *Technology and language testing* (pp. 13-28). Washington, DC: TESOL publications.

Tyner, K. (1998). Literacy in a digital world: *Teaching and learning in the age of information.* Mahwah, NJ: Lawrence Erlbaum Associates.

University of California, Los Angeles. (2001). WebLAS Web Page: www.humnet.ucla.edu/lasslab/wedlas.htm (Accessed 11 January 2005).

University of California, Los Angeles. (2002). WebLAS Demonstration Test. Online: http://www.humnet.ucla.edu/web/deparments/alt/weblas_esl_demo/demo_listen_psych1_vid.htm (Accessed 11 January 2005).

UCLA Applied Linguistics and Center for Digital humanities. 2001-2003. WebLAS documents. Online: http://www.weblas.ucla.edu/ (Accessed 11 January 2005).

University of Cambridge Local Examination Syndicate. Online: http://www.bulats.org (Accessed 11 January 2005).

University of Surrey. (2001). Online Self-Access Quizzes: http://www.surrey.ac.uk/ELI/sa/thesisl.html (Accessed 11 January 2005).

Vispoel, W. P. (1998). Psychometric characteristics of computer-adaptive and self-adaptive vocabulary tests: The role of answer feedback and test anxiety. *Journal of Educational Measurement,* 35(2), 155-67.

Vispoel, W. P. Hendrikcson, A. B. & Bleiler, T. (2000). Limiting answer and change computerized adaptive vocabulary tests: psychometric and attitudinal results. *Journal of Educational Measurement* 37(1), 21-38.

Wainer, H. & Braun, H. I. (eds.). (1988). *Test validity.* NJ: Lawrence Erlbaum Associates.

Wainer, H. & Eignor, D. (2000). Caveats, pitfalls, and unexpected consequences of implementing large-scale computerized testing. In H. Wainer, N.J. Dorans, D. Eignor, R. flaugher, B. f. green, R.J. Mislevy, L. Steinberg & D. Thissen (2000). *Computer adaptive testing: A primer,* 2nd edn (pp. 271-99). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wainer, H. Dorans, N. J., Eignor, D., Flaugher, R, Green, B. F., Mislevy, R. J., Steinberg, L. & Thissen, D. (2000). *Computer adaptive testing: A primer,* 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wainer, H. Dorans, N. J., Flaugher, R, Green, B. F., Mislevy, R. J., Steinberg, L. & Thissen, D. (1990). *Computer adaptive testing: A primer,* Hillsdale, NJ: Lawrence Erlbaum Associates.

Wall, D. (1997). Impact and washback in language testing. In C. Clapham and D. Corson (eds.), *Encyclopedia of language and education, vol. VII: Language testing and assessment* (pp. 291-302). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Warchauer, M. (2000). The changing global economy and the future of English teaching: *TESOL Quarterly,* 34(3), 511-35.

Warchauer, M. (1989). Researching technology in TESOL: Determinist, instrumental, and critical approaches. *TESOL Quarterly,* 32(4), 757-61.

Web-based Education Commission. (2002). *The power of the Internet for learning.* Report of the Web-based Education Commission to the President and Congress of the United States. Online: http://www.hpcnet.org/webcommission (Accessed 11 January 2005).

WebCT. (2004). Version 3.0. WebCT, inc. http://www.webct.com (Accessed 11 January 2005).

Weigle, S. C. (2002). *Assessing writing.* Cambridge: Cambridge University Press.

Williamson, D. M., Befar, I.I. & Hone, A. S. (1999). "Mental model" comparison of automated and human scoring. *Journal of Educational Measurement,* 36, 185-84.

Wolfe-Wuintero, K. Inagaki, S. & Kim, H-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity.* Honolulu: University of Hawaii Press.

Wresch, W. (1993). The imminence of grading essays by computer-25years later. *Computers and composition,* 10(2), 45-58.

Young, R., Shermis, M. d., Brutten, S. & Perkins, K. (1996). From conventional to computer adaptive testing of ESL reading comprehension. *System,* 24(1), 32-40.

**Ability see communicative competenc; language ability.*

ثبت المصطلحات

أولاً: عربي – إنجليزي

أ

الاتجاه الابتكاري	Innovative approach
اتجاه المقارنة	Comparison approach
اتجاه النفق	Tunnel approach
الاتساق في التوجيهات	Consistency in instructions
الاتصالات التقنية	Technology-mediated communication
أتمتة تصحيح المقال	Automated essay scoring
أثر الطريقة	Method effect
إجراءات أداء الاختبار	Procedures for test taking
إجراءات الاستجابة	Procedures of responding
اختبار التتمة المنتظمة	Cloze test
الاختبار التشخيصي المحوسب	Diagnostic computer-delivered test
اختبار التوفل المحوسب	CBT

CALT reading test	اختبار القراءة المحوسب
Proficiency test	اختبار الكفاية اللغوية
TOEFL	اختبار اللغة الإنجليزية بوصفها لغة أجنبية
Second language testing	اختبار اللغة الثانية
Computer-Based Eurocentres Test	اختبار المراكز الأوروبية المعتمد على الحاسوب
Vocabulary tests	اختبار المفردات
Placement test	اختبار تحديد المستوى
Achievement test	اختبار تحصيلي
Diagnostic test	اختبار تشخيصي
Screening test	اختبار تصفية
Computer-adaptive test	اختبار حاسوبي مكيف
WBET	الاختبار عبر شبكة الإنترنت
Usability testing	اختبار قابلية الاستخدام
Paper-and-pencil test	اختبار ورقي
Computer-based reading tests	اختبارات القراءة المحوسبة
Language tests	اختبارات اللغة
CALT	اختبارات اللغة والمحوسبة
High-stakes tests	الاختبارات المصيرية
Objective tests	الاختبارات الموضوعية
Practice tests	اختبارات تدريبية/تجريبية
Communicative test	اختبارات تواصلية

Communicative task-based tests	اختبارات تواصلية قائمة على المهام
Individualized class room testing	اختبارات صفية ذات طابع فردي
Low-stakes tests	الاختبارات غير المصيرية
Criterion-Referenced tests	اختبارات محكية المرجع
Large scale tests	اختبارات واسعة النطاق
Adaptive test selection	اختيار تكيفي للبنود
Multiple choice	الاختيار من متعدد
Differential performance	الأداء التفاضلي
Performance on computer-based tests	الأداء في الاختبارات المحوسبة
Language performance	أداء لغوي
CALT tools	أدوات اختبارات اللغة المحوسبة
Assessment tools	أدوات التقييم
Authoring tools	أدوات توليف
Correlation	الارتباط
Bivairate correlation	ارتباط ذو متغيرين
Technical guidelines	إرشادات التقنية
Navigation buttons	أزرار التنقل
linguistic responses	الاستجابات اللغوية
Expected responses	الاستجابات المتوقعة
Extended responses	الاستجابات الموسعة
Non-linguistic responses	الاستجابات غير اللغوية

Multiple-response	الاستجابة المتعددة
Correct response	استجابة صحيحة
Creating test tasks	استحداث المهام الاختبارية
Language use	الاستخدام اللغوي
Inferences	استدلالات
Self-assessment strategies	استراتيجيات التقييم الذاتي
Communicative strategies	استراتيجيات التواصل
Securely retrieving of data	الاسترجاع الآمن للبيانات
Viewing records	استعراض النتائج
Matching questions	أسئلة المزاوجة
Open-ended questions	أسئلة مفتوحة
Toolbars	أشرطة الأدوات
Interactional authenticity	الأصالة التفاعلية
Situational authenticity	الأصالة الموقفية
Test usefulness framework	إطار الاستفادة من الاختبار
Common European Framework	الإطار الأوروبي المشترك
Settings	إعداد (شاشة)
Developing CALT	إعداد اختبارات اللغة المحوسبة
Test developing	إعداد الاختبار
Breakdown	الأعطال
Language acquisition	اكتساب اللغة

اكتساب اللغة الثانية	SLA
اكتساب اللغة الثانية المقيَّم	Assessed SLA
اكتساب اللغة الثانية الموجه	Instructed SLA
آلة تصحيح الاختبارات	Test scoring machine
الألفة باستخدام الحاسوب	Familiarity with computer use
الألفة بالحاسوب	Computer familiarity
إنتاج البنود آليًّا	Automatic item generation
الإنتاج اللغوي المطول	Extended language production
انتخاب البنود	Item selection
أنشطة متزامنة	Concurrent activities
أنظمة توليف	Authoring tools
أيقونات	Icons

ب

بدء (شاشة)	Start
برامج التوليف	Authoring software
البرمجيات	Software
برمجيات التوليف	Authoring software
برنامج معالجة النصوص	Word-processing program
بروتوكول التذكر	Recall protocols
بطارية ميتشجان لتقييم اللغة الإنجليزية	MELAB
بناء (زر)	Create

بناء الاختبار	Creating the test
بند	Item
بنك البنود	Item bank
بنك المهام	Task bank
بنود الاختبار	Test items
بيانات	Data

ت

التأثير التحويلي	Transformative impact
التأثير السلبي	Negative impact
التأليف عبر الإنترنت	Composing online
التجربة العملية	Field trial
تحرير (شاشة)	Edit
تحقيق الصدق	Validation
التحليل الإحصائي	Statistical analysis
تحليل الأداء العام	General performance analysis
التحليل الأساسي للتكرار	Fundamental frequency analysis
تحليل الاستجابات	Response analysis
تحليل الاستفادة	Usefulness analysis
التحليل الآلي للاستجابات	Automatic response analysis
تحليل البروتوكول الشفهي	Verbal protocol analysis

Analysis of variance	تحليل التباين
Computer-assisted response analysis	التحليل الحاسوبي للاستجابات
Discourse analysis	تحليل الخطاب
Authenticity analysis	تحليل الموثوقية
Pronunciation analysis	تحليل النطق
Analyzing reliability	تحليل ثبات الاختبار
Test bias	التحيز في الاختبار
Measurement bias	التحيز في القياس
Storing results	تخزين النتائج
Rater training	تدريب المقيِّمين
Scoring complex language performance	تصحيح الأداء اللغوي المعقد
Automatic response scoring	تصحيح الاستجابات آليًّا
Automated scoring	التصحيح المؤتمت
Human scoring	تصحيح بشري
Computer-mediated scoring	تصحيح بواسطة الحاسوب
Detailed scoring	تصحيح تفصيلي
Dichotomus scoring	تصحيح ثنائي الخيارات
Partial-credit scoring	تصحيح جزئي
Polytomus scoring	تصحيح متعدد الخيارات
Automated scoring	تصحيح مؤتْمت
Page layout	تصميم الصفحة

Interface design	تصميم واجهة الاستخدام
Designing interface	تصميم واجهة المستخدم
Intelligent applications	التطبيقات الذكية
Application of word processing and the internet	تطبيقات معالجة النصوص والإنترنت
Modifying the input	تعديل المدخلات
Software breakdown	تعطل البرمجيات
Equipment breakdown	تعطل المعدات
Summary feedback	تعقيب موجز
Feedback	تعقيب/ تغذية راجعة
Intelligent learning	التعلم الذكي
Instant feedback	تغذية راجعة فورية
Interaction between input and responses	التفاعل بين المدخلات والاستجابات
Interactiveness	التفاعلية
Interpretation of performance	تفسير الأداء
Score interpretation	تفسير الدرجات
Interpretation of test scores	تفسير نتائج الاختبارات
Rating	التقدير
Human and machine rating	التقدير الآلي والبشري
Holistic rating	التقدير الشامل
Test delivering	تقديم الاختبار
Instructional technology	التقنيات التعليمية

Interactive technologies	التقنيات التفاعلية
Computer technology	تقنية الحاسوب
Evaluating CALT	تقويم اختبارات اللغة المحوسبة
Tests evaluation	تقويم الاختبارات
Evaluation of linguistic responses	تقويم الاستجابات اللغوية
Evaluating scoring methods	تقويم طرائق التصحيح
Assessing computer-adaptive tests	تقييم الاختبارات الحاسوبية المتكيفة
Constructed responses assessment	تقييم الاستجابات الإنشائية
Human rating	التقييم البشري
Diagnostic assessment	التقييم التشخيصي
Integrated assessment	التقييم التكاملي
Self assessment	التقييم الذاتي
Intelligent assessment	التقييم الذكي
Classroom assessment	التقييم الصفي
Assessing languages for specific purposes	تقييم اللغات لأغراض خاصة
Evaluation of CALT	تقييم اللغة المحوسب
Language assessment	التقييم اللغوي
Computer-based language assessment	التقييم اللغوي المعتمد على الحاسوب
Web LAS	التقييم اللغوي عبر الإنترنت
Computer-assisted assessment	التقييم المحوسب
High-stakes assessment	التقييم المصيري

التقييم بمساعدة التقنية	Technology-assisted assessment
تقييم ذاتي	Self-rating
التقييم عبر الإنترنت	Online assessment
تقييم مهارة التحدث التلقائي	Phone pass
التواصل	Communication
التواصل البشري	Personal communication
التواصل الفعال	Effective communication
توجيهات الاختبار ومدخلاته	Test rubrics and input
التوجيهات المحوسبة	Computer-assisted instructions
توليف	Authoring
توليف اختبارات اللغة	Authoring language tests
توليف التقييم اللغوي	Authoring language assessments

ث

الثبات	Reliability
ثروة لفظية	Vocabulary knowledge

ج

الجدل التفسيري	Interpretive argument
جودة الاختبارات	Quality of tests

ح

Evaluative impetus	حافز تقييمي
Test wiseness	الحنكة الاختبارية

خ

Server	خادم
Characteristics of assessment	خصائص التقييم
Computer algorithm	الخوارزميات الحاسوبية
Options for feedback	خيارات التعقيب

د

Score	درجة
Screen resolution	دقة الشاشة
Significance	دلالة

ذ

Artificial intelligence	الذكاء الاصطناعي

ز

Download time	زمن التنزيل

س

Test security	سرية الاختبار وأمنه
Context	السياق
Communicative context	السياق التواصلي
Testing contexts	السياقات الاختبارية
Educational contexts	السياقات التربوية

ش

Screen	شاشة

ص

Test validity	صدق الاختبار
Validity of inferences	صدق الاستدلالات
Construct validity	صدق البناء
Concurrent validity	الصدق التزامني
Predictive validity	الصدق التنبئي
Face validity	الصدق الظاهري
Content validity	صدق المحتوى
Item difficulty	صعوبة البند
Home page	الصفحة الرئيسة

ض

الضبط	Fine tuning

ط

طباعة البيانات	Data printing
طرائق اختبارية	Testing methods
طرائق تجريبية	Empirical methods
طفرة التقييم	Assessment evaluation

ظ

الظروف الزمانية والمكانية (للاختبار)	Physical and temporal circumstances

ع

عامل الأداة	Utility factor
عامل الصلة	Relevance factor
عامل الكفاية	Sufficiency factor
العروض التقديمية	Power point
علامات البناء الخطابي	Markers of discourse structure
علم النفس المعرفي	Cognitive psychology
عملية العرض	Presentation process
عناصر التحكم	Controls

ق

Database	قاعدة البيانات
Browser compatibility	قدرات برنامج المتصفح
Strategic competence	القدرة الاستراتيجية
Communicative competence	القدرة التواصلية
Communicative language competence	القدرة اللغوية التواصلية
Writing ability	قدرة كتابية
Test anxiety	قلق اختباري
Standardized templates	قوالب معيارية
Check lists	قوائم التحقق
Transformative powered technology	القوة التحويلية للتقنية
Educational measurement	القياس التربوي
Model-based measurement	القياس القائم على الأنموذج
CALT limitations	قيود اختبار اللغة المحوسب
Constraints of technology	قيود التقنية

ك

Strategic competence	الكفاية الاستراتيجية
Content words	كلمات المحتوى

ل

Applied linguistics	اللغويات التطبيقية
Computational linguistics	اللغويات الحاسوبية
Control panel	لوحة تحكم
Adaptive algorithm	لوغاريثم تكيفي

م

Metalinguistic	ما وراء اللغة
Literacy	مبادئ القراءة والكتابة
Computer programmer	مبرمج الحاسوب
Needs of assessment	متطلبات التقييم
Contrasting groups	مجموعات التقابل
Testlet	النص والأسئلة المرتبطة به
Computer-based simulation	المحاكاة المحوسبة
Input	المدخلات
Test input	مدخلات اختبارية
Rich contextualized input	المدخلات السياقية الثرية
Linguistic input	المدخلات اللغوية
Audio input	مدخلات صوتية
Language teachers	مدرسو اللغة

Spell-checker	المدقق الإملائي
Procedures CALT advantages	مزايا اختبار اللغة المحوسب
Item-level adaptivity	مستوى التكيف للبند
Differential problems	مشكلات التفاضل
Response processing	معالجة الاستجابات
Data processing	معالجة البيانات
Natural language processing (NLP)	معالجة اللغات الطبيعية
Coefficient	معامل
Spearman rho correlation coefficient	معامل ارتباط سبيرمان رو
Discrimination index	معامل التمييز
Reliability coefficient	معامل الثبات
Coefficient-alpha reliabilities	معامل ألفا
Preview and publish	معاينة ونشر (شاشة)
Standards	معايير
Rubrics	معايير الأداء
Standard evaluation criteria	معايير التقويم القياسية
Test developers	معدو الاختبارات
Full-motion video	مقاطع الفيديو الحركية
E-rater	المقدر الإلكتروني
Human rater	المقدر البشري
Likert scale	مقياس لكيرت

Test constructs	مكونات الاختبار
Language constructs	المكونات اللغوية
Language test tasks	مهام اختبارات اللغة
Computer-assisted test tasks	مهام الاختبارات المحوسبة
Test tasks	المهام الاختبارية
Academic tasks	المهام الأكاديمية
Completion tasks	مهام الإكمال
Language assessment tasks	مهام التقييم اللغوي
Computer-assisted writing tasks	مهام الكتابة المحوسبة
Complex tasks	المهام المعقدة
Threats to CALT	مخاوف اختبارات اللغة المحوسبة
Threats to validity	مخاوف الصدق
CALL materials	مواد تعليم اللغة بالحاسوب
Computer specifications	مواصفات الحاسوب

ن

Publishing the test on the web	نشر الاختبار على شبكة الإنترنت
Data management system	نظام إدارة البيانات
Task development system	نظام إعداد المهام
Tense and aspect system	نظام الزمن والوجهة
IELTS	النظام العالمي لاختبار اللغة الإنجليزية

Criterion	نظام الكريتيرين
Task delivery system	نظام تقديم المهام
Item response theory (IRT)	نظرية الاستجابة للبند
Utilization argument	نظرية الاستخدام
Interpretive argument	النظرية التفسيرية
Evidence model	أنموذج الدليل
Rasch model	أنموذج راش

و

Interface	واجهة الاستخدام
Interface	واجهة المستخدم
Communication media	وسائط الاتصال
Technology means	وسائط التقنية
Audio/video methods	وسائط الصوت والصورة
Multimedia	الوسائط المتعددة
Multimedia texts	الوسائط النصية
Mode of test delivery	وسيلة تقديم الاختبار
Quick access	الوصول السريع
Self-access	وصول ذاتي
Communicative function	الوظيفة التواصلية

ثانياً: إنجليزي – عربي

A

Academic tasks	المهام الأكاديمية
Achievement test	اختبار تحصيلي
Adaptive algorithm	لوغاريثم تكيفي
Adaptive test selection	اختيار تكيفي للبنود
Analysis of variance	تحليل التباين
Analyzing reliability	تحليل ثبات الاختبار
Application of word processing and the internet	تطبيقات معالجة النصوص والإنترنت
Applied linguistics	اللغويات التطبيقية
Artificial intelligence	الذكاء الاصطناعي
Assessed SLA	اكتساب اللغة الثانية المقيَّم
Assessing computer-adaptive tests	تقييم الاختبارات الحاسوبية المتكيفة
Assessing languages for specific purposes	تقييم اللغات لأغراض خاصة
Assessment evaluation	طفرة التقييم
Assessment tools	أدوات التقييم
Audio input	مدخلات صوتية
Audio/video methods	وسائط الصوت والصورة
Authenticity analysis	تحليل الموثوقية
Authoring	توليف

Authoring language assessment	توليف التقييم اللغوي
Authoring language tests	توليف اختبارات اللغة
Authoring software	برامج التوليف
Authoring software	برمجيات التوليف
Authoring tools	أدوات توليف
Authoring tools	أنظمة توليف
Automated scoring	التصحيح المؤتمت
Automated essay scoring	أتمتة تصحيح المقال
Automatic item generation	إنتاج البنود آليًّا
Automatic response analysis	التحليل الآلي للاستجابات
Automatic response scoring	تصحيح الاستجابات آليًّا

B

Bivairate correlation	ارتباط ذو متغيرين
Breakdown	الأعطال
Browser compatibility	قدرات برنامج المتصفح

C

CALL materials	مواد تعليم اللغة بالحاسوب
CALT	اختبارات اللغة المحوسبة
CALT limitations	قيود اختبار اللغة المحوسب

CALT reading test	اختبار القراءة المحوسب
CALT tools	أدوات اختبارات اللغة المحوسبة
CBT	اختبار التوفل المحوسب
Characteristics of assessment	خصائص التقييم
Check lists	قوائم التحقق
Class room assessment	التقييم الصفي
Cloze test	اختبار التتمة المنتظمة
Coefficient	معامل
Coefficient-alpha reliabilities	معامل ألفا
Cognitive psychology	علم النفس المعرفي
Common European Framework	الإطار الأوروبي المشترك
Communication	التواصل
Communication media	وسائط الاتصال
Communicative competence	القدرة التواصلية
Communicative context	السياق التواصلي
Communicative function	الوظيفة التواصلية
Communicative language competence	القدرة اللغوية التواصلية
Communicative strategies	استراتيجيات التواصل
Communicative task-based tests	اختبارات تواصلية قائمة على المهام
Communicative test	اختبارات تواصلية
Comparison approach	اتجاه المقارنة

Completion tasks	مهام الإكمال
Complex tasks	المهام المعقدة
Composing online	التأليف عبر الإنترنت
Computational linguistics	اللغويات الحاسوبية
Computer algorithm	الخوارزميات الحاسوبية
Computer familiarity	الألفة بالحاسوب
Computer programmer	مبرمج الحاسوب
Computer specification	مواصفات الحاسوب
Computer technology	تقنية الحاسوب
Computer-adaptive test	اختبار حاسوبي مكيف
Computer-assisted assessment	التقييم المحوسب
Computer-assisted instructions	التوجيهات المحوسبة
Computer-assisted response analysis	التحليل الحاسوبي للاستجابات
Computer-assisted test tasks	مهام الاختبارات المحوسبة
Computer-assisted writing tasks	مهام الكتابة المحوسبة
Computer-Based Eurocentres Test	اختبار المراكز الأوروبية المعتمد على الحاسوب
Computer-based language assessment	التقييم اللغوي المعتمد على الحاسوب
Computer-based reading tests	اختبارات القراءة المحوسبة
Computer-based simulation	المحاكاة المحوسبة
Computer-mediated scoring	تصحيح بواسطة الحاسوب
Concurrent activities	أنشطة متزامنة

Concurrent validity	الصدق التزامني
Consistency in instructions	الاتساق في التوجيهات
Constraints of technology	قيود التقنية
Construct validity	صدق البناء
Constructed responses assessment	تقييم الاستجابات الإنشائية
Content validity	صدق المحتوى
Content words	كلمات المحتوى
Context	السياق
Contrasting groups	مجموعات التقابل
Control panel	لوحة تحكم
Controls	عناصر التحكم
Correct response	استجابة صحيحة
Correlation	الارتباط
Create	بناء (زر)
Creating test tasks	استحداث المهام الاختبارية
Creating the test	بناء الاختبار
Criterion	نظام الكريتيرين
Criterion-Referenced tests	اختبارات محكية المرجع

D

Data	بيانات

Data management system	نظام إدارة البيانات
Data printing	طباعة البيانات
Data processing	معالجة البيانات
Database	قاعدة البيانات
Designing interface	تصميم واجهة المستخدم
Detailed scoring	تصحيح تفصيلي
Developing CALT	إعداد اختبارات اللغة المحوسبة
Diagnostic assessment	التقييم التشخيصي
Diagnostic computer-delivered test	الاختبار التشخيصي المحوسب
Diagnostic test	اختبار تشخيصي
Dichotomous scoring	تصحيح ثنائي الخيارات
Differential performance	الأداء التفاضلي
Differential problems	مشكلات التفاضل
Item difficulty	صعوبة البند
Discourse analysis	تحليل الخطاب
Discrimination index	معامل التمييز
Download time	زمن التنزيل

E

Edit	تحرير (شاشة)
Educational contexts	السياقات التربوية
Educational measurement	القياس التربوي

Effective communication	التواصل الفعال
Empirical methods	طرائق تجريبية
Equipment breakdown	تعطل المعدات
E-rater	المقدر الإلكتروني
Evaluating CALT	تقويم اختبارات اللغة المحوسبة
Evaluating scoring methods	تقويم طرائق التصحيح
Evaluation of CALT	تقييم اللغة المحوسب
Evaluation of linguistic responses	تقويم الاستجابات اللغوية
Evaluative impetus	حافز تقييمي
Evidence model	أنموذج الدليل
Expected responses	الاستجابات المتوقعة
Extended language production	الإنتاج اللغوي المطول
Extended responses	الاستجابات الموسعة

F

Face validity	الصدق الظاهري
Familiarity with computer use	الألفة باستخدام الحاسوب
Feedback	تعقيب/ تغذية راجعة
Field trial	التجربة العملية
Fine tuning	الضبط
Full-motion video	مقاطع الفيديو الحركية
Fundamental frequency analysis	التحليل الأساسي للتكرار

G

General performance analysis	تحليل الأداء العام

H

High-stakes assessment	التقييم المصيري
High-stakes tests	الاختبارات المصيرية
Holistic rating	التقدير الشامل
Home page	الصفحة الرئيسة
Human and machine rating	التقدير الآلي والبشري
Human rater	المقدر البشري
Human rating	التقييم البشري
Human scoring	تصحيح بشري

I

Icons	أيقونات
IELTS	النظام العالمي لاختبار اللغة الإنجليزية
Individualized class room testing	اختبارات صفية ذات طابع فردي
Inferences	استدلالات
Innovative approach	الاتجاه الابتكاري
Input	المدخلات
Instant feedback	تغذية راجعة فورية

Instructed SLA	اكتساب اللغة الثانية الموجه
Instructional technology	التقنيات التعليمية
Integrated assessment	التقييم التكاملي
Intelligent applications	التطبيقات الذكية
Intelligent assessment	التقييم الذكي
Intelligent learning	التعلم الذكي
Interaction between input and responses	التفاعل بين المدخلات والاستجابات
Interactional authenticity	الأصالة التفاعلية
Interactive technologies	التقنيات التفاعلية
Interactiveness	التفاعلية
Interface	واجهة الاستخدام
Interface	واجهة المستخدم
Interface design	تصميم واجهة الاستخدام
Interpretation of performance	تفسير الأداء
Interpretation of test scores	تفسير نتائج الاختبارات
Interpretive argument	الجدل التفسيري
Interpretive argument	النظرية التفسيرية
Item	بند
Item bank	بنك البنود
Item response theory (IRT)	نظرية الاستجابة للبند
Item selection	انتخاب البنود
Item-level adaptivity	مستوى التكيف للبند

L

Language acquisition	اكتساب اللغة
Language assessment	التقييم اللغوي
Language assessment tasks	مهام التقييم اللغوي
Language constructs	المكونات اللغوية
Language performance	أداء لغوي
Language teachers	مدرسو اللغة
Language test tasks	مهام اختبارات اللغة
Language tests	اختبارات اللغة
Language use	الاستخدام اللغوي
Large scale tests	اختبارات واسعة النطاق
Likert scale	مقياس لكيرت
Linguistic input	المدخلات اللغوية
linguistic responses	الاستجابات اللغوية
Literacy	مبادئ القراءة والكتابة
Low-stakes tests	الاختبارات غير المصيرية

M

Markers of discourse structure	علامات البناء الخطابي
Matching questions	أسئلة المزاوجة
Measurement bias	التحيز في القياس

MELAB	بطارية ميتشجان لتقييم اللغة الإنجليزية
Metalinguistic	ما وراء اللغة
Method effect	أثر الطريقة
Mode of test delivery	وسيلة تقديم الاختبار
Model-based measurement	القياس القائم على الأنموذج
Modifying the input	تعديل المدخلات
Multimedia	الوسائط المتعددة
Multimedia texts	الوسائط النصية
Multiple choice	الاختيار من متعدد
Multiple-response	الاستجابة المتعددة

N

Natural language processing (NLP)	معالجة اللغات الطبيعية
Navigation buttons	أزار التنقل
Needs of assessment	متطلبات التقييم
Negative impact	التأثير السلبي
Non-linguistic responses	الاستجابات غير اللغوية

O

Objective tests	الاختبارات الموضوعية
Online assessment	التقييم عبر الإنترنت
Open-ended questions	أسئلة مفتوحة

Options for feedback	خيارات التعقيب

P

Page layout	تصميم الصفحة
Paper-and-pencil test	اختبار ورقي
Partial-credit scoring	تصحيح جزئي
Performance on computer-based tests	الأداء في الاختبارات المحوسبة
Personal communication	التواصل البشري
Phone pass	تقييم مهارة التحدث التلقائي
Physical and temporal circumstances	الظروف الزمانية والمكانية (للاختبار)
Placement test	اختبار تحديد المستوى
Polytomus scoring	تصحيح متعدد الخيارات
Power point	العروض التقديمية
Practice tests	اختبارات تدريبية/تجريبية
Predictive validity	الصدق التنبئي
Presentation process	عملية العرض
Preview and publish	معاينة ونشر (شاشة)
Procedures CALT advantages	مزايا اختبار اللغة المحوسب
Procedures for test tasking	إجراءات أداء الاختبار
Procedures of responding	إجراءات الاستجابة
Proficiency test	اختبار الكفاية اللغوية
Pronunciation analysis	تحليل النطق

Publishing the test on the web	نشر الاختبار على شبكة الإنترنت

Q

Quality of tests	جودة الاختبارات
Quick access	الوصول السريع

R

Rasch model	أنموذج راش
Rater training	تدريب المقيِّمين
Rating	التقدير
Recall protocols	بروتوكول الاستذكار
Relevance factor	عامل الصلة
Reliability	الثبات
Reliability coefficient	معامل الثبات
Response analysis	تحليل الاستجابات
Response processing	معالجة الاستجابات
Rich contextualized input	المدخلات السياقية الثرية
Rubrics	معايير الأداء

S

Score	درجة
Score interpretation	تفسير الدرجات

Scoring complex language performance	تصحيح الأداء اللغوي المعقد
Screen	شاشة
Screen resolution	دقة الشاشة
Screening test	اختبار تصفية
Second language testing	اختبار اللغة الثانية
Securely retrieving of data	الاسترجاع الآمن للبيانات
Self assessment	التقييم الذاتي
Self-access	وصول ذاتي
Self-assessment	التقييم الذاتي
Self-assessment strategies	استراتيجيات التقييم الذاتي
Self-rating	تقييم ذاتي
Server	خادم
Settings	إعداد (شاشة)
Significance	دلالة
Situational authenticity	الأصالة الموقفية
SLA	اكتساب اللغة الثانية
Software	البرمجيات
Software breakdown	تعطل البرمجيات
Spearman rho correlation coefficient	معامل ارتباط سبيرمان رو
Spell-checker	المدقق الإملائي
Standard evaluation criteria	معايير التقويم القياسية

Standardized templates	قوالب معيارية
Standards	معايير
Start	بدء (شاشة)
Statistical analysis	التحليل الإحصائي
Storing results	تخزين النتائج
Strategic competence	القدرة الاستراتيجية
Strategic competence	الكفاية الاستراتيجية
Sufficiency factor	عامل الكفاية
Summary feedback	تعقيب موجز

T

Task bank	بنك المهام
Task delivery system	نظام تقديم المهام
Task development system	نظام إعداد المهام
Technical guidelines	إرشادات التقنية
Technology means	وسائط التقنية
Technology-assisted assessment	التقييم بمساعدة التقنية
Technology-mediated communication	الاتصالات التقنية
Tense and aspect system	نظام الزمن والوجهة
Test anxiety	قلق اختباري
Test bias	التحيز في الاختبار

Test constructs	مكونات الاختبار
Test delivering	تقديم الاختبار
Test developers	معدو الاختبارات
Test developing	إعداد الاختبار
Test input	مدخلات اختبارية
Test items	بنود الاختبار
Test rubrics and input	توجيهات الاختبار ومدخلاته
Test scoring machine	آلة تصحيح الاختبارات
Test security	سرية الاختبار وأمنه
Test tasks	المهام الاختبارية
Test usefulness framework	إطار الاستفادة من الاختبار
Test validity	صدق الاختبار
Test wiseness	الحنكة الاختبارية
Testing contexts	السياقات الاختبارية
Testing methods	طرائق اختبارية
Testlet	النص والأسئلة المرتبطة به
Tests evaluation	تقويم الاختبارات
Threats to CALT	مخاوف اختبارات اللغة المحوسبة
Threats to validity	مخاوف الصدق
TOEFL	اختبار اللغة الإنجليزية بوصفها لغة أجنبية
Toolbars	أشرطة الأدوات
Transformative impact	التأثير التحويلي

Transformative powered technology	القوة التحويلية للتقنية
Tunnel approach	اتجاه النفق

U

Usability testing	اختبار قابلية الاستخدام
Usefulness analysis	تحليل الاستفادة
Utility factor	عامل الأداة
Utilization argument	نظرية الاستخدام

V

Validation	تحقيق الصدق
Validity of inferences	صدق الاستدلالات
Verbal protocol analysis	تحليل البروتوكول الشفهي
Viewing records	استعراض النتائج
Vocabulary knowledge	ثروة لفظية
Vocabulary tests	اختبار المفردات

W

WBET	الاختبار عبر شبكة الإنترنت
Web LAS	التقييم اللغوي عبر الإنترنت
Word-processing program	برنامج معالجة النصوص
Writing ability	قدرة كتابية

كشاف الموضوعات

أ

اتجاه المقارنة ١٧٠

إجراءات أداء الاختبار ١٢٢

اختبار

التتمة المنتظمة ٢٣

تحديد مستوى ٥٠

تشخيصي ١٥٣

الاختبارات

المصيرية ١، ٣٦، ٤٠، ٨٨، ٨٩، ١٣٦

غير المصيرية ٤٠، ٨٩

أدوات توليف ١٠٩، ١١٠، ١١٧، ١٥٥، ١٦٨، ١٦٩

الاستجابات

اللغوية ٨٢، ٨٤، ٨٦، ١٢٩

غير اللغوية ٨٠، ١٦٢

استراتيجيات التقييم الذاتي ١٦٦

أشرطة الأدوات ١٢٥

الأصالة التفاعلية ١٤

الأصالة الموقفية ١٣

إعداد اختبارات اللغة المحوسبة ٢٩، ٩٦، ١١٩، ١٢٣، ١٤٢، ١٥٦، ١٦١

اكتساب اللغة ٣، ١٣، ٢٥، ٢٧، ١٥٠، ١٥٣، ١٥٧، ١٦١، ١٦٢، ١٦٣، ١٦٥

اكتساب اللغة الثانية ٣، ٢٥، ١٥٠، ١٥٣، ١٥٧، ١٦١، ١٦٢، ١٦٣، ١٦٥

الإنتاج اللغوي المطول ١٦٣

آلة تصحيح الاختبارات ٥٢، ٥٣
أنظمة توليف ١١٦

ب

برمجيات التوليف ١٤٢
بناء الاختبار ٥١، ٩٨، ١٠١، ١٢٠
بنك البنود ٨٩
بنك المهام ١١٥
بنود الاختبار ٩، ١٢، ١٩، ٤٩، ٦٣، ٦٦، ٧٧، ١١٤، ١٢٢

ت

تحليل الاستجابات ٨٤، ١١٤
تغذية راجعة ٩
التقييم
الذاتي ٩، ١٧، ٤٠، ١٧٢
الصفي ٥، ٩
المحوسب ٣، ١٦٧

ث

الثبات ١١، ٤٩، ٦٧، ٧٦، ١٣٣، ١٣٤، ١٣٥، ١٣٦، ١٣٧

ج

الجدل التفسيري ١٦٠

خ

خصائص التقييم ٣٦، ٣٧، ٥٢
خيارات التعقيب ١٢١

س

السياق التواصلي ٥٢
السياقات التربوية ١٤٧

ص

صدق الاختبار ٢٣، ٢٧
صدق البناء ١٢٠، ١٢٧، ١٣٥، ١٣٧، ١٣٨، ١٣٩
صدق المحتوى ١٢٧

ط

طرائق اختبارية ١٥

ع

عامل الكفاية ١٩

عملية العرض ١١٣
عناصر التحكم ١٢٥

ق

قاعدة البيانات ١٠٥، ١١١، ١١٢، ١١٤، ١١٥، ١٢٦
قوالب معيارية ٩٨
قيود اختبار اللغة المحوسب ٣٦

ل

اللغويات التطبيقية ٥٧، ٩٨، ١١٦، ١٤٢، ١٥٦، ١٥٧، ١٦٥، ١٦٨، ١٧٠، ١٧١، ١٧٢
اللغويات الحاسوبية ١٦٤

م

مبرمج الحاسوب ١٥٢
مدخلات اختبارية ٦٥
مزايا اختبار اللغة المحوسب ٣٦
مهام الاختبارات المحوسبة ١٣٥
مواصفات الحاسوب ١٢٤

ن

النص والأسئلة المرتبطة به ٥١
نظام إدارة البيانات ١١٦
نظام إعداد المهام ١١٥
نظرية الاستجابة للبند ٤٩، ١٢٨، ١٣٨

و

واجهة الاستخدام ١٢١
وسائط الاتصال ١٥٨
الوسائط المتعددة ١٠، ١٣، ١٤، ٣٦، ١١٦، ١٢٥، ١٣٦، ١٦٧