

صدق البناء الداخلي لاختبار القدرة اللفظية ومقارنة مجموعة نماذج بنائية بديلة: التكامل بين التحليل العاملي الاستكشافي والتوكيدي

راشد سيف المحرزي *

ملخص

هدفت الدراسة إلى فحص صدق البناء الداخلي لاختبار القدرة اللفظية في اختبار القدرات العامة الذي يعده المركز الوطني للقياس والتقويم في التعليم العالي بالمملكة العربية السعودية باستخدام التحليل العاملي الاستكشافي والتحليل العاملي التوكيدي باستخدام نسختين من الاختبار. دعمت النتائج صدق البناء الداخلي لاختبار القدرة اللفظية وفقاً لمواصفات مكونات الاختبار. بغرض تطوير قياس اختبار القدرة اللفظية، قامت الدراسة أيضاً بفحص جودة ملاءمة مجموعة من النماذج البنائية البديلة للبناء الداخلي الحالي. وأشارت النتائج إلى توافر درجة ملاءمة لأربعة نماذج مختزلة ويقوم قربية من جودة ملاءمة النموذج الحالي. وخلصت الدراسة إلى تقديم عدد من التوصيات حول البناء الداخلي لاختبار القدرة اللفظية وطرق تطويره.

الكلمات الدالة: صدق البناء الداخلي، التحليل العاملي الاستكشافي، التحليل العاملي التوكيدي، القدرة اللفظية.

المقدمة

استحضار هذه المفردات في وقت قصير، واستخدامها بشكل مناسب للسياق ولتنوع الأفراد الموجه إليهم الخطاب. ويذكر خير الله (1974) أن القدرة اللفظية قدرة مركبة من عدة قدرات ثانوية بسيطة تجمع الأساليب والأنشطة المختلفة التي تتعلق بمعالجة الأفكار والمعاني عن طريق استخدام الألفاظ. وتشكل القدرة على الاستيعاب القرائي أعقد عمليات القدرة اللفظية، لأنها تتطلب نشاطاً ذهنياً فاعلاً لإدراك الألفاظ في النص والعلاقات بين الجمل والفقرات استناداً إلى المخزون المعرفي السابق لفهم النص وتحليله ونقده (عافشي، 2012). وتتطلب عملية الاستيعاب القرائي التفاعل النشط بين القارئ والنص من خلال الاستراتيجيات القرائية مثل فك الرموز ومعرفة المعاني الدقيقة للكلمات واستيعابها في حيز السياق، واستراتيجيات فهم الجمل والفقرات (عاشور، 2006، التل ومقدادي، 1989).

ويذكر سويت (1993) Sweet وتشابمان (1993) Chapman أن القراءة هي عملية ديناميكية تتطلب عملية تمثيل عقلي للنص، واستدعاء المعلومات ذات الصلة من الذاكرة، وتقييم الاختلافات بين النص وما يعرفه ويعتقد به القارئ سابقاً، وعمل استدلالات معينة لملء الفراغات في الفهم السابق ولتوضيح المعنى الجديد، وتكامل المعلومات الجديدة في النص مع القاعدة المعرفية السابقة للقارئ، وتمييز النقاط المهمة وغير المهمة في النص، وكيفية استخدام المعلومات الموجودة. ويختلف الأفراد في قدرتهم اللفظية حسب مستوى تمكنهم من القدرات المكونة لعملية القراءة.

تعد القدرة اللفظية إحدى القدرات الأساسية التي تتطلبها البرامج الدراسية بمؤسسات التعليم العالي (Du Boulay, 1999)، ذلك أن دراسة جميع العلوم تعتمد على مقدار اللغة التي يمتلكها الطالب، ومقدار تمكنه من فهم محتوى ومعنى ما يقرؤه، والقدرة على استنباط الأفكار من النصوص والربط بينها، والوصول إلى تعميمات صحيحة تستند إلى أدلة وشواهد. وذكر نست وسمبسون (2000) Nist & Simpson أن الدراسات في هذا المجال أظهرت أن 85% تقريباً من جميع مجالات التعلم بالجامعات تتطلب القراءة، وأن الكتب الدراسية جزء أساسي في التعلم على جميع المستويات والمراحل التعليمية.

وتختلف النظريات المعرفية في علم النفس المعرفي في تعريف القدرة اللفظية؛ لذا يصعب الوصول إلى تعريف علمي دقيق للقدرة اللفظية، نتيجة تأثرها بمتغيرات كثيرة كالمستوى الدراسي للطالب المراد قياس قدرته اللفظية، والهدف من عملية القياس، والغرض الذي ستستخدم فيها الدرجات المحصلة. ويُعرف أندريو، وكوب وجيميترو (Andrew, Cobb & Giamietro, 2005) القدرة اللفظية بأنها مقدار سهولة تقديم الفرد لأفكاره على شكل كلمات سواء أكانت كتابة أم تحدثاً، والتي تتطلب امتلاك رصيد كبير ومتنوع من المفردات اللغوية، والقدرة على

* جامعة السلطان قابوس، عُمان. تاريخ استلام البحث 2012/10/23، وتاريخ قبوله 2013/5/13.

بدراسة تهدف إلى التعرف على عمليات القراءة البسيطة المتمثلة في معاني الكلمات، وفهم المعنى في السياق، وفهم المسموع، وتأثير كل من عمليات القراءة البسيطة (Simple Reading processes)، والمهارة اللفظية (Verbal Proficiency)، والطلاقة القرائية (Reading Fluency) في الاستيعاب القرائي (Reading Comprehension) لدى 271 من الطلبة القارئيين في الصفوف الرابع والسابع والتاسع. وأظهرت الدراسة تأثير عمليات القراءة البسيطة في الاستيعاب القرائي في جميع الصفوف الدراسية الثلاثة مع نقصان في مقدار تفسيرها للبتابن في الاستيعاب القرائي في الصفوف العليا. كما أضافت كل من المهارة اللفظية والطلاقة القرائية تأثيراً دالاً احصائياً في الاستيعاب القرائي على التأثير الناتج من عمليات القراءة البسيطة في جميع الصفوف الدراسية الثلاثة.

وأظهرت دراسة التل ومقادي (1989) أن القدرة اللفظية بصفة عامة، والاستيعاب القرائي بصفة خاصة هي نتاج لعملية التعلم وطرق التدريس وطرق التقويم التكويني المستخدمة من قبل المعلمين بالمراحل التعليمية المختلفة. ففي التجربة التي أجراها الباحثان حول فاعلية استخدام الأسئلة الصفية من المستويات العقلية العليا على زيادة الاستيعاب القرائي باستخدام مجموعتين تكونت كل منهما من 30 طالباً من طلاب الصف الخامس بمدينة إربد بالملكة الأردنية الهاشمية، أظهرت النتائج وجود فرق دال إحصائياً في مستوى الاستيعاب القرائي بين المجموعتين التجريبية والضابطة لصالح المجموعة التي أستخدمت فيها الأسئلة التي تتطلب قدرات عقلية عليا.

وقام الصارمي وإبراهيم (2004) بمقارنة مستوى الاستيعاب القرائي لدى 122 طالباً وطالبة من طلاب الصف الثاني الثانوي بسلطنة عمان حسب متغيرات الجنس والتخصص (علمي وأدبي) ونوع النص المستخدم. وتم تحليل فقرات الاستيعاب القرائي (14 فقرة) في اختبار القدرة اللفظية ضمن اختبار القدرات الدراسية المُعد كاختبار للقبول بالتعليم الجامعي. ويتكون اختبار القدرة اللفظية من ثلاثة مكونات: مكون المترادفات والمتشابهات اللفظية، ومكون إكمال الجمل، ومكون الاستيعاب القرائي. واحتوى مكون الاستيعاب القرائي في هذا الاختبار نصين: أحدهما في العلوم الطبيعية والآخر في العلوم الإنسانية. وأظهرت الدراسة عدم وجود فروق دالة إحصائياً بين الجنسين في الاستيعاب القرائي، بينما ظهرت الفروق بين الجنسين عند أخذ تخصص الطلاب بعين الاعتبار. فأظهرت الدراسة أن الاستيعاب القرائي كان أفضل لدى الإناث منها لدى الذكور في التخصص العلمي، بينما كان نمط الفروق متعكسا في التخصص الأدبي. وتتطابق هذه النتائج في نوعي

ويذكر بورتون وويلش وكوستن وفان ايسن (Burton, Welsh, Kostin, & Van Essen 2009) أن هناك ثمان عمليات يمكن من خلالها قياس القدرة اللفظية في اختبارات القبول بالتعليم العالي، وهي: (1) الاستيعاب القرائي، ويتضمن معاني الكلمات والجمل والنصوص، (2) تفسير المقروء، ويتضمن التحليل والوصول إلى استنتاجات حول النص، (3) تقييم المقروء، ويتضمن التعرف على جوانب القوة والضعف في النص، (4) الربط بين المقروء والقاعدة المعرفية والمعتقدات السابقة، ويتضمن تقييم الاختلاف بين محتوى المقروء والقاعدة المعرفية والمعتقدات السابقة للقارئ، (5) تكوين فهم جديد، ويتضمن الانتقال إلى أبعد من استقبال المعلومة إلى استخدامها وتطبيقها، (6) السعي لحل المشكلة، ويتضمن تحديد الأجزاء التي تتطلب البحث عن أفكار جديدة، (7) التواصل، ويتضمن القدرة على الكتابة والعرض والشرح والإقناع والتفاعل مع العاملين في التخصص ذاته، (8) تحكم الفرد في طريقة فهمه واستدلالاته وعملياته العقلية، ويتضمن استخدام محكات مختلفة للتحكم في الفهم أثناء القراءة، وتعديل الاستراتيجيات المساعدة في فهم النص. وبرغم أهمية هذه العمليات اللازمة للقدرة اللفظية وترابطها، لا يمكن لاختبار ما أن يقيس جميع هذه العمليات. بل تختلف القدرة على قياس هذه العمليات نظراً لصعوبة قياس بعضها، واختلاف ظروف الاختبار، والبيئة التعليمية في المجتمع، ومتطلبات التعليم العالي. فبرغم أن بعض الأفراد يستخدمون القدرة الاستدلالية اللفظية في النصوص المسموعة، فإن أغلب اختبارات القبول تركز على قياس القدرة الاستدلالية باستخدام الألفاظ المكتوبة أو المقروءة. كما أن أغلب اختبارات القدرات اللفظية لا تقيس المهارات اللغوية الأخرى مثل الكتابة، والإنصات، والتحدث، والأدب اللغوي.

ومن بين الدراسات التي تناولت مفهوم القدرة اللفظية دراسة ستيرنبرج (Sternberg 1985) والتي هدفت إلى فهم الذكاء اللفظي Verbal Intelligence من خلال التعرف على مكونات الاستيعاب اللفظي عن طريق إجراء سلسلة تجارب باستخدام سياقات مختلفة داخلية وخارجية على عينات مختلفة من الطلبة. وأظهرت الدراسة ارتباط القدرة على استيعاب الألفاظ في السياق بالتفاعل بين عمليات معالجة المعلومات والإشارات المتوافرة في السياق الذي استخدمت فيه الألفاظ. وأظهرت الدراسة اختلاف عمليات استيعاب المواد اللفظية عن عمليات استيعاب المواد المجردة كالرموز والأشكال والرسومات.

كما قام تيلستا وماك ماستر وفان دن برويك وكنديوي وراب (Tilstra, McMaster, Van den Broek, Kendeou, & Rapp 2009)

النصوص المستخدمة.

وتشكل القدرة اللفظية إحدى القدرات الأساسية في اختبارات الذكاء والاستعداد الدراسي بالإضافة إلى الاستدلال الكمي والتفكير المنطقي والمكاني كاختبار التقييم الدراسي Scholastic Assessment Test (SAT)، واختبار الكليات الأمريكية American College Testing (ACT) واختبارات سجل المتخرج Graduate Record Examinations (GRE) في الولايات المتحدة الأمريكية، وغيرها (أبو دنيا ومحروس، 2002). وحيث إن اختبارات القدرات والاستعداد الدراسي التي تستخدم لأغراض القبول بمؤسسات التعليم العالي تركز على المهارات الأكاديمية الأساسية التي يمكن تطبيقها في قطاع عريض من التخصصات والمجالات المختلفة، فإن قياس القدرة اللفظية فيها يركز على المهارات العليا في القدرة اللغوية، مثل القراءة الناقدة والقدرة على الاستدلال، واللتين ينبغي أن يمتلكهما الطالب المعتمد على ذاته في اكتساب المعرفة في التعليم العالي (الصارمي وإبراهيم، 2004؛ Burton, Welsh, Kostin, & VanEssen, 2009).

ونظراً لاختلاف الطلاب في مستوى قدرتهم اللفظية، فإنه يتطلب من اختبارات القبول قياس جوانب متعددة من القدرة اللفظية بمستويات متنوعة من الصعوبة لتغطية جميع مستويات الطلاب المختلفة. ويمكن إدراك صعوبة قياس القدرة اللفظية وتحديد مكوناتها وعملياتها الضرورية للتعليم العالي من خلال دراسة التعديلات المتكررة والمتعاقبة عبر السنوات لتطوير اختبار التقييم الدراسي (SAT) - Scholastic Assessment Test - على سبيل المثال - لتحسين مستوى صدق نتائجه في قياس القدرة اللفظية (Burton, et. al., 2009). حيث عدل مجلس الجامعات الأمريكي (College Board) من أقسام القدرة اللفظية في اختبار SAT عدة مرات. فبدأ في عام 1901م بقياس قدرات مختلفة من ضمنها بعض عمليات القدرة اللفظية، ثم عدل في عام 1936م ليقاس القدرة اللفظية بشكل مستقل، وتضمنت أربعة مكونات وهي: معاني الكلمات، والتناظر اللفظي، والتعريفات المزدوجة، وقراءة النصوص، ثم في عام 1946م، أُستبدل مكون التعريفات المزدوجة بمكون إكمال الجمل، واستمر كذلك حتى عام 1994م، عندما تم حذف مكون معاني الكلمات، ثم حذف مكون التناظر اللفظي بعد ذلك في عام 2005م. وفي العام نفسه، عدل مسمى القدرة اللفظية إلى القراءة الناقدة (Critical Reading)، لكي تتضمن مكونين هما: إكمال الجمل وقراءة النصوص القصيرة والطويلة. وأصبح يركز الاختبار على قياس قدرة الطالب على التأمل والقيام باستدلالات واستنتاجات حول النص المقروء (Burton, et. al.,

2009). وقد أُجريت هذه التعديلات استجابة لتطور مفهوم القدرة اللفظية في الأدب النظري في علم النفس المعرفي، وانتقادات مؤسسات التعليم العالي المستخدمة لنتائج الاختبار، والدراسات الميدانية على الاختبار، وغيرها من المتغيرات.

ومن بين الأساليب المهمة لدراسة صدق نتائج الاختبارات ما يعرف بصدق المكونات البنائية (البناء الداخلي) Internal Structure Validity. فيقوم صدق البناء الداخلي للاختبار بدور كبير في تحديد جوانب القدرة العقلية التي يتطلب قياسها في اختبارات القدرات (Nitko & Brookhart, 2007). ويهتم صدق البناء الداخلي للاختبار بتوفير الأدلة على اتساق مكونات الاختبار ونظام تصحيحه مع الإطار المفاهيمي للمكونات البنائية للقدرة المقيسة والعلاقات فيما بينها (Messick, 1993). فيجب أن تتسق نوعية وعدد العوامل التي تجمع الارتباطات بين فقرات الاختبار لنوعية وعدد عوامل مجال السمة المقاسة بالاختبار، ويتطلب من واضعي الاختبار بذل المستطاع لاستهداف هذا البناء من خلال تمثيل السمة من حيث عدد الفقرات، وما تقيسه هذه الفقرات، ونظام التصحيح المتبع وتفسير درجات الاختبار. كما يسهم صدق البناء الداخلي في تقديم سبل تطوير قياس القدرة من خلال فحص ومقارنة نماذج مختلفة من البناء الداخلي لفقرات الاختبار،

تتم مقارنة النماذج المختلفة للبناء الداخلي لفقرات الاختبار بطريقة وصفية أو بطريقة كمية. ويفضل ميسك (Messick 1993) أن تتم مقارنة النماذج المختلفة للبناء الداخلي كميًا من خلال حساب مؤشرات لجودة ملاءمة هذه النماذج لبيانات الاختبار. ويعد التحليل العاملي الاستكشافي Exploratory Factor Analysis والتحليل العاملي التوكيدي Confirmatory Factor Analysis من أكثر الطرق المستخدمة لدراسة المكونات البنائية لفقرات الاختبار. ويهدف التحليل العاملي الاستكشافي إلى اشتقاق عدد محدد من المحاور الأساسية التي تجمع الارتباطات البيئية لفقرات الاختبار، وتكون فيما بينها -بأوزان معينة- تفسيراً للتباينات المشاهدة في درجات الطلاب (Cattell, 1978). وتُفسر العوامل الناتجة من التحليل العاملي الاستكشافي العمليات المشتركة بين الفقرات التي تنتسب في كل عامل. وكلما كانت هذه العوامل الناتجة متسقة مع طبيعة السمة المقاسة والعمليات التي تتطلبها، كانت النتائج داعمة لطريقة وضع الاختبار وشاهدة على صدق البناء الداخلي له. ولكن قد يؤدي التحليل العاملي الاستكشافي في حالات عديدة إلى عوامل قد لا تتسق مع السمة المقيسة ومكوناتها نتيجة لعينات الطلاب، وطريقة استجاباتهم لفقرات الاختبار، ونوعية الفقرات، والطريقة المتبعة في تطبيق التحليل العاملي

أساسيتين للنجاح في التعليم العالي وهما القدرة اللفظية والقدرة الكمية.

بالاستفادة من الإطار النظري في علم النفس المعرفي لقياس القدرة اللفظية، والخبرات السابقة في دول العالم في بناء اختبارات القبول بمؤسسات التعليم العالي، بُني اختبار القدرة اللفظية بالمركز الوطني للقياس والتقويم في التعليم العالي منذ البدء في تطبيق اختبار القدرات العامة في عام 1423/1422هـ ليقاس أربعة مكونات للقدرة اللفظية، متمثلة في: (1) معاني الكلمات، (2) إكمال الجمل، (3) التناظر اللفظي، (4) الاستيعاب القرائي. ويهدف قياس القدرة اللفظية في الاختبار إلى تزويد الجامعات بمعلومات عن مستوى الطالب ومقدار تمكنه من فهم محتوى ومعنى ما يقرؤه وقدرته على استنباط الأفكار من النصوص والربط بينها والوصول إلى تعميمات صحيحة تستند إلى أدلة وشواهد.

ونظرا لأهمية القرارات التي تبنى على نتائج الطلبة في اختبار القدرة اللفظية سواء على مستوى الطالب بشكل فردي، من حيث قبوله من عدمه، أم على مستوى مؤسسات التعليم العالي، من حيث فعالية انتقاء أفضل المدخلات، تظهر الحاجة الماسة لدراسة العلاقات البيئية بين فقرات الاختبار للتعرف على البناء الداخلي لمكونات القدرة اللفظية المقيسة. وتوفر مثل هذه الدراسة معلومات عن صدق بناء اختبار القدرة اللفظية واقتراح جوانب تطويرية لتحسين قياس القدرة اللفظية. فمن المعايير الأساسية التي يجب على أي مؤسسة معنية ببناء وتطوير الاختبارات على تحقيقها ضرورة إجراء دراسات حول صدق البناء الداخلي للاختبارات، وتكرارها بشكل دوري لتقديم أدلة وحجج على مصداقية درجات الطلاب الحالية في القدرة اللفظية والاستفادة منها في تطوير طرق قياس القدرة اللفظية.

من كل ذلك، تبرز الحاجة الماسة لإجراء دراسات على صدق البناء الداخلي لاختبار القدرة اللفظية في اختبار القدرات العامة. وتهدف هذه الدراسة إلى فحص البناء الداخلي لمكونات اختبار القدرة اللفظية من خلال اختبار جودة ملاءمة النموذج الأصلي الذي بُني عليه الاختبار، ومقارنة عدد من النماذج المختزلة البديلة التي تفترض إمكانية دمج بعض مكونات اختبار القدرة اللفظية بسبب ترابطها الإحصائي والمنطقي. وسيتم توظيف التحليل العملي الاستكشافي والتحليل العملي التوكيدي لتحقيق أهداف الدراسة. ويمكن ترجمة أهداف الدراسة في الأسئلة الآتية:

1. ما مدى تطابق المكونات العملية لاختبار القدرة اللفظية الناتجة من التحليل العملي الاستكشافي مع مواصفات البناء العملي للاختبار؟

الاستكشافي. لذا يمكن اعتبار نتائج التحليل العملي الاستكشافي افتراضات، يجب فحصها بشكل تجريبي وجمع بيانات أخرى تدعم هذه الافتراضات أو تدحضها.

وفي المقابل، يهدف التحليل العملي التوكيدي إلى تقدير قيم تشبعات (Loading) الفقرات مع المكونات النظرية للسمة المقيسة التي وُضعت مسبقاً من خلال الإطار النظري أو نتائج التحليل العملي الاستكشافي أو تجريب فرضيات بديلة. ويقدم أيضاً أساليب إحصائية لفحص مستوى جودة ملاءمة بيانات الاختبار لتشبع الفقرات مع المكونات، ومقارنتها مع نماذج بديلة أخرى، عن طريق ما يعرف بمؤشرات ومحكات جودة ملاءمة النموذج Model Fit Indices. ويعد التحليل العملي التوكيدي من أكثر الطرق المستخدمة في دراسة صدق البناء الداخلي للاختبارات وسبل تطويرها (Holdnack, Zhou, Larrabee, 1984; Tanka & Huba, 1984; Millis & Salthouse, 2011).

وتؤكد معايير الاختبارات التربوية والنفسية - التي أعدتها الجمعية الأمريكية للبحوث التربوية، والجمعية النفسية الأمريكية، والمجلس الوطني للقياس في التربية - على ذلك لزيادة موثوقية النتائج والقرارات التي يتخذها مستخدمو نتائج الاختبارات (AERA, APA, & NCME, 1999). وتكتسب دراسة البناء الداخلي لاختبار القدرة اللفظية أهميتها من احتمالية حدوث تغيرات في البناء الداخلي للاختبار مع الوقت نظرا للتغيرات المحتمل حدوثها في قدرات المتقدمين للاختبار عبر السنوات نتيجة للتغيرات الحادثة في أي مجتمع في الجوانب التعليمية والثقافية والاجتماعية والاقتصادية. كما قد تطرأ على القدرات اللغوية للمتقدمين تغيرات نتيجة للتدريب على مكونات الاختبار الذي قد يخضع له الطلاب. وقد يؤدي هذا التغير في القدرات اللغوية للمتقدمين إلى تغير البناء الداخلي لاختبار القدرة اللفظية. كما يتطلب من مطوري اختبارات القدرة اللفظية مواكبة التطورات الحادثة في علم النفس المعرفي حول مفهوم القدرة اللفظية وطرق قياسها.

مشكلة الدراسة

يعد اختبار القدرات العامة الذي يقدمه المركز الوطني للقياس والتقويم بالمملكة العربية السعودية من اختبارات قياس الاستعداد الدراسي للقبول بمؤسسات التعليم العالي، ويسعى إلى تقديم معلومات صادقة عن القدرة المستقبلية للطالب على التعامل مع المهارات والعمليات التي تتطلبها البرامج الدراسية بمؤسسات التعليم العالي. وتعتمد أغلب الجامعات السعودية على نتائج اختبار القدرات العامة في فرز الطلاب حسب نتائجهم كأحد معايير قبول الطلاب بمختلف التخصصات والبرامج الدراسية. ويقاس اختبار القدرات العامة قدرتين

المستخدمة بالدراسة، حيث تم استخدام عينتين من طلبة المرحلة الثانوية الذكور (الصفين الثاني والثالث الثانوي) من مختلف مناطق المملكة العربية السعودية في العام الأكاديمي 1431/1430هـ. كما أن نتائج الدراسة محدودة باستخدام نسختين فقط من اختبار القدرة اللفظية في اختبار القدرات العامة الذي يقدمه المركز الوطني للقياس والتقييم بالمملكة العربية السعودية.

الطريقة والإجراءات

مجتمع الدراسة وعينتها

يتمثل مجتمع الدراسة في جميع طلبة المرحلة الثانوية الذكور (الصفين الثاني والثالث الثانوي) من مختلف مناطق المملكة العربية السعودية في العام الأكاديمي 1431/1430 الذين يتقدمون لاختبار القدرات العامة. وتكونت عينة الدراسة من عينتين من طلبة المرحلة الثانوية الذكور (الصفين الثاني والثالث الثانوي) من مختلف مناطق المملكة العربية السعودية في العام الأكاديمي 1431/1430هـ، تم اختيارهم بطريقة عشوائية. بلغ حجم العينة الأولى 43452 طالباً، بينما تكونت العينة الثانية من 43687 طالباً. قدمت العينة الأولى النسخة 1030 من اختبار القدرة اللفظية، وقدمت العينة الثانية النسخة 1035. ويقوم نظام المركز الوطني للقياس والتقييم على تقديم نسخة واحدة للاختبار لجميع الطلاب المتقدمين في كل يوم من أيام تطبيق الاختبار. ويقوم الطالب بنفسه باختيار اليوم الذي يقدم فيه الاختبار، والذي يحدد نسخة الاختبار التي يقدمها. ويساعد هذا النظام على تحقيق عشوائية توزيع عينات الطلاب على نسخ الاختبار المختلفة.

أداة الدراسة وخصائصها السيكومترية:

استخدمت هذه الدراسة نسختين من اختبار القدرة اللفظية في اختبار القدرات العامة، وهي: 1030، 1035. وتتكون جميع نسخ اختبار القدرة اللفظية في الأصل من 68 فقرة من نوع الاختيار من متعدد (لكل منها أربعة بدائل) موزعة على أربعة مكونات بتوزيع معين ومتشابه عبر النسخ المختلفة. وتتمثل مكونات القدرة اللفظية في الاختبار في: (1) معاني الكلمات، ويشكل ما وزنه 20% تقريباً، و(2) إكمال الجمل، ويشكل ما وزنه 26% تقريباً، و(3) التناظر اللفظي، ويشكل ما وزنه 31% تقريباً. كما يتبع المركز آلية واضحة في عملية بناء النسخ المختلفة للاختبار من خلال وضع جدول مواصفات محدد للاختبار واختيار فقرات الاختبار وتوزيعها على النسخ

2. ما مدى جودة ملائمة بيانات اختبار القدرة اللفظية مع نموذج مواصفات البناء الداخلي للاختبار باستخدام التحليل العاملي التوكيدي؟

3. ما مؤشرات تعديل النموذج (Model Modification) في النموذج الأصلي للبناء الداخلي لاختبار القدرة اللفظية؟

4. ما مدى جودة ملائمة النماذج المختزلة (Reduced Models) البديلة للبناء الداخلي الحالي لاختبار القدرة اللفظية؟ وما أفضلها جودة؟

5. هل يتحقق الصدق البيئي (Cross Validation) للنماذج المختزلة البديلة للبناء الداخلي لاختبار القدرة اللفظية في نسخة أخرى من الاختبار؟

أهمية الدراسة

1. توفر الدراسة أدلة على صدق البناء الداخلي لاختبار القدرة اللفظية في اختبار القدرات العامة، مما يزيد من ثقة واضعي الاختبار ومستخدميه في مصداقية درجات الطلاب وقرارات القبول بمؤسسات التعليم العالي.

2. تساعد دراسة نماذج مختلفة للبناء الداخلي في تطوير اختبار القدرة اللفظية في اختبار القدرات العامة، والتي تؤدي إلى تحسين عمليات بناء فقرات ومكونات الاختبار.

3. تعضد هذه الدراسة التكامل بين التحليل العاملي الاستكشافي والتحليل العاملي التوكيدي في دراسة البناء العاملي في تطوير أدوات قياس القدرات العقلية واختبارات الاستعداد الدراسي.

مصطلحات الدراسة

القدرة اللفظية: تعرف القدرة اللفظية إجرائياً في هذه الدراسة بمجموع الدرجات التي يحصل عليها الطالب في كل من مكون معاني الكلمات، وإكمال الجمل، والتناظر اللفظي، والاستيعاب القرآني في اختبار القدرة اللفظية.

صدق البناء الداخلي: يعرفها ميسيك (Messick 1993) بالأساليب المختلفة التي توفر الأدلة اللازمة على اتساق مكونات الاختبار ونظام تصحيحه مع الإطار المفاهيمي للمكونات البنائية للقدرة المقيسة من الاختبار، والعلاقات فيما بينها. وتتمثل أساليب صدق البناء الداخلي في هذه الدراسة في التحليل العاملي الاستكشافي والتوكيدي لفحص ومقارنة نماذج مختلفة من العوامل التي تربط فقرات اختبار القدرة اللفظية.

النماذج المختزلة Reduced Models: تلك النماذج التي تحتوي على عدد أقل من العوامل التي يتضمنها النموذج الأصلي أو الكامل.

محددات الدراسة

تحدد القدرة على تعميم نتائج هذه الدراسة بالعينات

النسختين، حيث بلغ معامل ثبات ألفا كرونباخ 0.876 للنسخة 1030، وبلغ 0.891 للنسخة 1035، مما يشير إلى ثبات عالٍ. كما حسبت الاحصاءات الوصفية (المتوسطات والانحرافات المعيارية) ومعاملات الارتباط بين درجات الطلاب في المكونات الأربعة لاختبار القدرة اللفظية مع بعضها ومع القدرة اللفظية الكلية في كلتا النسختين كمؤشرات للصدق التكويني للاختبار كما يظهر في الجدول (2). وتظهر النتائج تقارب الإحصاءات الوصفية ومعاملات ارتباط المكونات الأربعة لاختبار القدرة اللفظية مع بعضها، حيث تراوحت معاملات الارتباط الداخلية بين 0.592-0.678. كما أن درجة ارتباط كل من هذه الأبعاد مع القدرة اللفظية الكلية كانت متقاربة وعالية، حيث تراوحت بين 0.814-0.890.

المختلفة بعد تجريبيها وتوظيف خصائصها السيكمترية لتحقيق أعلى درجات التكافؤ فيما بينها.

تمت دراسة الخصائص السيكمترية لفقرات الاختبار من خلال حساب معاملات الصعوبة والتمييز، وأظهرت التحليلات أن جميع الفقرات تتمتع بمعاملات صعوبة وتمييز مقبولة ومناسبة ما عدا مفردتين هما الفقرة 25 والفقرة 52 (معاملات تمييز أقل من 0.1) في النسخة 1030، والفقرة 15 والفقرة 49 (معاملات تمييز أقل من 0.1) في النسخة 1035، وتم حذفها، وبالتالي تكونت النسختان من 66 فقرة. ويظهر الجدول (1) توزيع فقرات الاختبار في النسختين على المكونات الفرعية للقدرة اللفظية. وتم التأكد من ثبات نتائج اختبار القدرة اللفظية في

الجدول (1). توزيع فقرات الاختبار في نسختي اختبار القدرة اللفظية على المكونات الفرعية للقدرة اللفظية

النسخ		المكونات
1035	1030	
14	13	معاني الكلمات
14	16	إكمال الجمل
17	17	التناظر اللفظي
21	20	الاستيعاب القرائي
66	66	الإجمالي

الجدول (2). الاحصاءات الوصفية ومصفوفة الارتباط للمكونات الأربعة في نسختي اختبار القدرة اللفظية

النسخة 1030							
المكون	مدى الدرجات	المتوسط الحسابي	الانحراف المعياري	إكمال الجمل	التناظر اللفظي	الاستيعاب القرائي	القدرة اللفظية
معاني الكلمات	13-0	7.084	2.675	0.631	0.592	0.621	0.821
إكمال الجمل	16-0	9.763	3.006		0.626	0.660	0.831
التناظر اللفظي	17-0	9.757	3.730			0.636	0.890
الاستيعاب القرائي	20-0	9.628	3.798				0.839
القدرة اللفظية	66-0	29.147	9.194				
النسخة 1035							
المكون	مدى الدرجات	المتوسط الحسابي	الانحراف المعياري	إكمال الجمل	التناظر اللفظي	الاستيعاب القرائي	القدرة اللفظية
معاني الكلمات	14-0	7.584	2.523	0.569	0.649	0.643	0.821
إكمال الجمل	14-0	8.674	3.102		0.619	0.578	0.831
التناظر اللفظي	17-0	9.854	3.681			0.678	0.890
الاستيعاب القرائي	21-0	10.025	3.812				0.839
القدرة اللفظية	66-0	30.012	10.002				

الداخلي للقدرة اللفظية.

4. للإجابة عن السؤال الرابع، تم تطبيق سبعة نماذج بديلة مختزلة (Reduced Model) للبناء الداخلي لاختبار القدرة اللفظية، والحصول على مؤشرات جودة ملاءمتها، ومقارنتها مع قيم محكات جودة الملاءمة للنموذج الكامل Full Model (نموذج المكونات الأربعة الأصلية) للاختبار باستخدام بيانات النسخة 1030. وحيث إن قيم محكات جودة ملاءمة النموذج الأصلي (النموذج الكامل) دائماً ما تكون أفضل من أي نموذج مختزل، فيتم التمييز بين النماذج البديلة المختزلة من حيث قرب قيم محكات جودة ملاءمتها من قيم محكات جودة ملاءمة النموذج الكامل. أي يهدف هذا التحليل إلى التوصل إلى نموذج بديل مختزل له قيم جودة مشابهة لقيم جودة النموذج الكامل، وبالتالي يمكن لهذا النموذج المختزل أن يحل محل النموذج الكامل.

5. للإجابة عن السؤال الخامس، أُجري تحليل عملي توكيدي على بيانات النسخة الأخرى للاختبار (1035) باستخدام خمسة نماذج للبناء الداخلي للاختبار: النموذج الأصلي وأربعة نماذج بديلة مختزلة تمثل أفضل النماذج البديلة التي حصلت على أفضل مؤشرات الجودة في السؤال الرابع. ويهدف هذا التحليل إلى التأكد من صدق النتائج التي تم والتوصل إليها، وهو ما يعرف بمسمى الصدق البيئي (Cross Validation).

النتائج

نتائج السؤال الأول: التحليل العملي الاستكشافي

أجري تحليل عملي استكشافي على نسختي اختبار القدرة اللفظية (1030، 1035) باستخدام تحليل المكونات الأساسية. وأظهرت نتائج التحليل العملي الاستكشافي على كلتا النسختين (كما في الجدول (3)) تشابهاً بين البناء الداخلي لفقرات اختبار القدرة اللفظية، من حيث عدد العوامل التي تربط الفقرات ونسبة التباين المفسر من هذه العوامل. أظهر التحليل تسعة عوامل حصلت على جذور كامنة أكبر من واحد في كلتا النسختين، وفسرت في مجملها ما قيمته 29.29% من التباين في النسخة 1030، و29.40% من التباين في النسخة 1035. وأشارت نتائج كلتا النسختين إلى وجود عامل واحد بجذر كامن كبير مقارنة ببقية العوامل الأخرى (9.59، 9.70 في النسختين 1030، 1035، على التوالي). كما كانت قيم الجذور الكامنة للعاملين الثاني والثالث متقاربة وصغيرة نسبياً (1.82، 1.52 في النسخة 1030 مثلاً)؛ بينما تبتعد نسبياً الجذور الكامنة

وعند دراسة الإحصاءات الوصفية (المتوسطات الحسابية والانحرافات المعيارية والالتواء والتقاطع) والتوزيعات التكرارية لدرجات الطلاب في القدرة اللفظية في نسختي الاختبار، يظهر تقارب هذه الإحصاءات الوصفية والتوزيعات التكرارية للنسختين، مما يشير إلى تكافؤ درجات الطلاب في القدرة اللفظية في نسختي الاختبار. ويجدر بالذكر أنه لم يتم الاعتماد على الاختبارات الاستدلالية للفروق بين المتوسطين الحسابيين نظراً للحجم الكبير للعينتين الذي يظهر دلالة إحصائية لأي فرق ولو كان صغيراً جداً وليس له مدلول عملي كما هو الحال في هذه البيانات (الفروق بين المتوسطين (0.865) أقل من درجة واحدة مقارنة بمدى الدرجات وهو 66).

الإجراءات والتحليلات الإحصائية

استخدمت الدراسة العديد من الأساليب الإحصائية للإجابة عن أسئلة الدراسة وتحقيق أهدافها، وفيما يلي تفصيل بالأساليب الإحصائية المستخدمة، والمعالجات، والخطوات المتبعة في التحليلات الإحصائية:

1. للإجابة عن السؤال الأول، أُجري التحليل العملي الاستكشافي باستخدام برنامج SPSS على بيانات النسختين بغرض تحديد عدد العوامل الكامنة لفقرات الاختبار من خلال مقارنة قيم الجذور الكامنة (Eigenvalues) ونسبة التباين المفسر (Explained Variance) للعوامل. ثم أُجري التحليل العملي الاستكشافي مرة أخرى مع تحديد عدد العوامل بأربعة عوامل، بغرض تحديد الفقرات المتشعبة بكل عامل. ولُخصت نتائج هذا التحليل بتحديد العامل الذي تشعب معه كل فقرة بأكثر مقدار مقارنة بتشعبها مع العوامل المتبقية (Christofferson 1975). تم بعد ذلك مقارنة نتائج التحليل العملي مع البناء الداخلي الأصلي للاختبار بمقارنة تشابه العامل الناتج من التحليل العملي لكل فقرة مع العامل الأصلي حسب مواصفات بناء الاختبار. وتم تكرار هذه العمليات في نسختي الاختبار.
2. للإجابة عن السؤال الثاني، أُجري التحليل العملي التوكيدي للبناء الداخلي لاختبار القدرة اللفظية وفقاً لمواصفات الاختبار باستخدام برنامج LISREL 8.52 (Jöreskog & Sörbom, 2001)، ودراسة مؤشرات ملاءمة النموذج باستخدام بيانات النسخة 1030.
3. للإجابة عن السؤال الثالث، حُسبت مؤشرات تعديل النموذج (Model Modification) لنموذج المكونات الأربعة الأصلية مع المكون العام في النسخة 1030، من خلال دراسة ترابط الفقرات مع المكونات الأربعة للقدرة اللفظية، واستخدامها في الوصول إلى نماذج بديلة مختزلة للبناء

على المكونات الأربعة، حيث تشبعت 5 فقرات من معاني الكلمات في مكون إكمال الجمل و5 فقرات في مكون الاستيعاب القرائي، بينما لم تنتسب أي فقرة في مكون التناظر اللفظي. أما فقرات إكمال الجمل فتشبت مفردتان في مكون معاني الكلمات و5 فقرات في مكون الاستيعاب القرائي، بينما لم تنتسب أية فقرة في مكون التناظر اللفظي. وتشبت أغلب فقرات الاستيعاب القرائي في مكونها الأصلي بينما تشبتت 3 فقرات فقط في مكون إكمال الجمل ولم تنتسب أية فقرة في مكون معاني الكلمات أو مكون التناظر اللفظي. وأخيرا تشبتت أغلب فقرات التناظر اللفظي في مكونها الأصلي وتشبتت الفقرات المتبقية (4 فقرات) في مكون الاستيعاب القرائي فقط.

يظهر الجدول (5) توزيع تشبعت فقرات النسخة 1035 على المكونات الأربعة، حيث أظهرت اختلافا في توزيع تشبعت فقرات النسخة 1035 عن فقرات النسخة 1030. فيظهر من الجدول تشبعت عدد من فقرات معاني الكلمات في مكون إكمال الجمل، وتشبعت عدد من فقرات إكمال الجمل في مكون معاني الكلمات (فقرة واحدة فقط)، وتشبعت عدد من فقرات التناظر اللفظي في كل من مكون معاني الكلمات ومكون إكمال الجمل. بينما لم تنتسب فقرات مكون الاستيعاب القرائي في أي مكون من المكونات الأخرى.

العوامل المتبقية وتقرب من الواحد الصحيح. وأظهرت نتائج التدوير المتعامد للعوامل باستخدام طريقة التباين الأقصى (Varimax) تقارب قيم الجذور الكامنة للعوامل الثلاثة الأولى في النسخة 1030 (تراوحت بين 4.19-3.24)، واختلافها عن قيم الجذور الكامنة للعوامل الستة الأخرى، بينما كانت قيم الجذور الكامنة في النسخة 1035 متقاربة للعوامل الأربعة الأولى (تراوحت بين 3.18-2.87)، ومختلفة عن قيم الجذور الكامنة للعوامل الخمسة الأخرى. ويمكن تلخيص هذه النتائج بأن هناك عاملا واحدا سائدا يربط فقرات الاختبار ويتمثل في القدرة اللفظية مع وجود ما بين ثلاثة إلى أربعة عوامل فرعية تربط مجموعات مختلفة من الفقرات.

بعد ذلك، تم تطبيق تحليل عاملي استكشافي مع تحديد عدد العوامل بأربعة (العدد الأكبر للعوامل) مع استخدام التدوير المتعامد للتأكد من كيفية توزيع الفقرات على العوامل الأربعة الناتجة، ومقارنتها بالمكونات الأربعة الأصلية للقدرة اللفظية في الاختبار وتوزيع الفقرات عليها. يظهر الجدولان (4) و(5) توزيع فقرات اختبار القدرة اللفظية على المكونات الناتجة من التحليل في نسختي الاختبار حسب القيمة الكبرى لتشبتت الفقرة مع ذلك البعد.

يظهر الجدول (4) توزيع تشبعت فقرات النسخة 1030

الجدول (3). العوامل ذات قيم جذور كامنة أكبر من واحد ونسبة تفسيرها للتباين في النسختين 1030 و1035 في اختبار القدرة اللفظية

العامل	النسخة 1030				النسخة 1035			
	قبل التدوير		بعد التدوير		قبل التدوير		بعد التدوير	
	التباين	الجذر	التباين	الجذر	التباين	الجذر	التباين	الجذر
1	14.53	9.59	14.70	9.70	6.35	4.19	14.53	9.59
2	2.76	1.82	2.81	1.85	5.74	3.79	2.76	1.82
3	2.31	1.52	2.13	1.41	4.92	3.24	2.31	1.52
4	1.76	1.16	1.85	1.22	2.80	1.85	1.76	1.16
5	1.67	1.10	1.69	1.12	2.35	1.55	1.67	1.10
6	1.60	1.06	1.62	1.07	1.97	1.30	1.60	1.06
7	1.58	1.04	1.55	1.02	1.87	1.23	1.58	1.04
8	1.55	1.02	1.53	1.01	1.80	1.19	1.55	1.02
9	1.53	1.00	1.52	1.00	1.60	1.06	1.53	1.00

الجدول (4). توزيع فقرات اختبار القدرة اللفظية (النسخة 1030) على المكونات الأربعة حسب قيم تشبعاتها

المكونات الأصلية	عدد الفقرات في المكونات الناتجة				عدد الفقرات في المكون الناتج
	معاني الكلمات	إكمال الجمل	التناظر اللفظي	الاستيعاب القرائي	
معاني الكلمات	3	5		5	13
إكمال الجمل	2	9		5	16
التناظر اللفظي			13	4	17
الاستيعاب القرائي		3		17	20
	5	17	13	31	66
	7.6	25.8	19.6	47.0	
	19.7	24.2	25.8	30.3	

الجدول (5). توزيع فقرات اختبار القدرة اللفظية (النسخة 1035) على المكونات الأربعة حسب قيم تشبعاتها

المكونات الأصلية	المكونات الناتجة				عدد الفقرات في المكون الناتج
	معاني الكلمات	إكمال الجمل	التناظر اللفظي	الاستيعاب القرائي	
معاني الكلمات	10	4			14
إكمال الجمل	6	7	1		14
التناظر اللفظي		7	10		17
الاستيعاب القرائي		4		17	21
	16	22	11	17	66
	24.2	33.3	16.7	25.8	
	21.2	21.2	25.8	31.8	

نتائج السؤال الثاني: التحليل العاملي التوكيدي

فحصت الدراسة البناء الداخلي لاختبار القدرة اللفظية باستخدام أسلوب التحليل العاملي التوكيدي من المستوى الثاني (Second-order Confirmatory Factor Analysis) الذي يقوم على وضع نماذج افتراضية للعوامل، واختبرت جودة ملامتها في البيانات، بحيث يُعرف النموذج الأفضل بالنموذج الذي يحقق أفضل مؤشرات جودة الملاءمة. وتم استخدام برنامج LISREL 8.52 لإجراء التحليل العاملي التوكيدي باستخدام معاملات الارتباط بين فقرات النسخة 1030. واستخدمت الدراسة سبعة محكات لمقارنة جودة ملاءمة النماذج، ويعرض الجدول (6) هذه المحكات والقيمة التي تعبر عن النموذج الأفضل لملاءمة (Browne & Cudeck, 1993). تم البدء بفحص جودة ملاءمة ثلاثة نماذج للمكونات

الأربعة للقدرة اللفظية حسب مواصفات بناء الاختبار، فيما تختلف فيما بينها في طبيعة الارتباط بين هذه المكونات الأربعة، حيث كانت كالتالي: (1) نموذج المكونات الأربعة المستقلة، (2) نموذج المكونات الأربعة المترابطة، (3) نموذج المكونات الأربعة مع مكون عام. يظهر الجدول (7) قيم محكات جودة ملاءمة هذه النماذج للمكونات الأربعة لاختبار القدرة اللفظية.

وتظهر النتائج أن قيم محكات جودة ملاءمة نموذج المكونات الأربعة المستقلة أقل من قيم نموذج المكونات الأربعة المترابطة وقيم نموذج المكونات الأربعة والمكون العام، مما يشير إلى وجود ترابط بين المكونات الأربعة. وتشير النتائج إلى أن النموذج الثالث (نموذج المكونات الأربعة مع المكون العام) أعطى قيما متطابقة لمحكات جودة ملاءمة نموذج المكونات

الأربعة المترابطة، مما يشير إلى أن المكون العام والمتمثل في الفرعية. ويدعم جودة ملائمة هذا النموذج مصفوفة الارتباط القدرة اللفظية العامة يفسر الترابطات بين المكونات الأربعة بين المكونات الأربعة النظرية الناتجة من التحليل التوكيدي لاختبار القدرة اللفظية. ويشير النموذج الثالث إلى وجود العامل العام الذي يربط المكونات الأربعة، مع وجود المكونات الأربعة

الجدول (6). معايير الحكم على محكات جودة ملائمة النماذج

المحكات	المعيار	Criteria
مربع كاي	الأصغر قيمة	CHI-SQUARE
مربع كاي / df	الأصغر قيمة	Chi-Square/df
الجذر التربيعي لمتوسط مربعات خطأ التقريب (RMSEA)	الأصغر قيمة	RMSEA
مؤشر جودة الملائمة (GFI)	الأكبر قيمة والأقرب للواحد	GOODNESS OF FIT INDEX (GFI)
مؤشر جودة الملائمة المعدل (AGFI)	الأكبر قيمة والأقرب للواحد	ADJUSTED GOODNESS OF FIT INDEX (AGFI)
مؤشر الملائمة المعياري (NFI)	الأكبر قيمة والأقرب للواحد	NORMED FIT INDEX (NFI)
مؤشر الملائمة المقارن (CFI)	الأكبر قيمة والأقرب للواحد	COMPARATIVE FIT INDEX (CFI)
مؤشر الملائمة الإضافي (IFI)	الأكبر قيمة والأقرب للواحد	INCREMENTAL FIT INDEX (IFI)

الجدول (7). ملائمة نماذج المكونات الأربعة لاختبار القدرة اللفظية (النسخة 1030)

المحكات	نموذج المكونات الأربعة المستقلة	نموذج المكونات الأربعة المترابطة	نموذج المكونات الأربعة + مكون عام
Chi-Square	217203	85509	85577
Chi-Square/df	104.78	41.25	41.24
RMSEA	0.053	0.035	0.035
GFI	0.85	0.93	0.93
AGFI	0.84	0.92	0.92
NFI	0.96	0.98	0.98
CFI	0.96	0.98	0.98
IFI	0.96	0.98	0.98

الجدول (8). مصفوفة الارتباط بين المكونات الأربعة حسب نموذج المكونات الأربعة المترابطة والمكون العام في اختبار القدرة اللفظية (النسخة 1030).

المكونات	معاني الكلمات	إكمال الجمل	التناظر اللفظي	الاستيعاب القرآني	المكون العام
معاني الكلمات	0.93	0.81	0.92	0.95	
إكمال الجمل		0.83	0.94	0.97	
التناظر اللفظي			0.84	0.86	
الاستيعاب القرآني				0.97	

نتائج السؤال الثالث: مؤشرات تعديل النموذج (Model Modification)

عند فحص معاملات الارتباط بين المكونات الأربعة في الجدول (8)، يظهر وجود ارتباط قوي بين كل من معاني الكلمات وإكمال الجمل والاستيعاب القرآني، بينما يقل ترابط مكون التناظر اللفظي مع كل من المكونات الثلاثة الأخرى. وتشير هذه النتائج إلى تقارب المكونات الثلاثة (معاني الكلمات، وإكمال الجمل، والاستيعاب القرآني) فيما تقيسه، وتختلف عن مكون التناظر اللفظي. وللفحص الدقيق لافتراض أن هذه المكونات الثلاثة متداخلة، لجأت الدراسة إلى أحد مؤشرات تعديل النموذج (Model Modification) التي يوفرها برنامج LISREL، وهو معامل الارتباط المتوقع لكل فقرة من فقرات الاختبار مع المكونات الأربعة كما لو تم افتراض ذلك (Steiger, 1990). ويظهر الجدول (9) معاملات ارتباط فقرات الاختبار مع كل مكون من المكونات الأربعة (الأرقام الغامقة تمثل ارتباط الفقرة مع المكون الذي تنتمي إليه). وتظهر النتائج أن ارتباط الفقرة بالمكون الذي حددت فيه حسب النموذج أكبر من ارتباطها بالمكونات الأخرى حسب ما هو متوقع.

بينما تشير معاملات ارتباط الفقرات مع المكونات الأخرى غير كونها الأصلي إلى بعض الدلالات؛ فجدد أولاً أن فقرات مكون التناظر اللفظي ترتبط بالمكونات الأخرى بقيم أقل نسبياً من ارتباطها بمكونها الأصلي، فمثلاً ترتبط الفقرة 2 من مكون التناظر اللفظي بقيم 0.70 مع كونها الأصلي، بينما ترتبط بكل من مكون معاني الكلمات، ومكون إكمال الجمل، ومكون الاستيعاب القرآني بقيم 0.56، 0.58، 0.58، على التوالي. ثانياً ترتبط فقرات المكونات الثلاثة الأخرى بدرجة أقل بمكون التناظر اللفظي مقارنة بارتباطها بمكوناتها الأصلية والمكونين الآخرين. فمثلاً ترتبط الفقرة 3 في مكون إكمال الجمل بقيمة 0.55 بمكون التناظر اللفظي، بينما ترتبط بقيمة 0.67 بمكونها الأصلي، وقيمة 0.62 بمكون معاني الكلمات، وقيمة 0.63 بمكون الاستيعاب القرآني. وهذا يشير إلى انتماء فقرات مكون التناظر اللفظي إلى كونها الأصلي، واختلاف هذا المكون عن بقية المكونات.

ثالثاً، يظهر الجدول (9) أن معظم فقرات مكون معاني الكلمات ترتبط مع كل من إكمال الجمل والاستيعاب القرآني بقيم متقاربة، وتقل بمقدار قليل جداً عن قيم ارتباطها ببعدها الأصلي، فمثلاً، ترتبط الفقرة 4 من مكون معاني الكلمات بمكونها الأصلي بقيمة 0.62، وترتبط بقيمة 0.58 مع مكون إكمال الجمل، وقيمة 0.57 مع مكون الاستيعاب القرآني، بينما ترتبط بقيمة 0.50 مع مكون التناظر اللفظي. رابعاً يتكرر النمط السابق في ارتباط فقرات كل من مكون إكمال الجمل

ومكون الاستيعاب القرآني. وتدفع هذه النتائج إلى احتمالية تداخل فقرات المكونات الثلاثة (معاني الكلمات، وإكمال الجمل، والاستيعاب القرآني) بشكل كبير جداً، وفي المقابل تختلف هذه المكونات عن مكون التناظر اللفظي.

نتائج السؤال الرابع: جودة ملائمة النماذج المختزلة (Reduced Models)

تشير النتائج السابقة إلى تداخل فقرات المكونات الثلاثة لاختبار القدرة اللفظية (معاني الكلمات، وإكمال الجمل، والاستيعاب القرآني)، وبالتالي يأتي التساؤل حول مستوى هذا التداخل بين المكونات الثلاثة، وما نوعه؟ فهل تتداخل المكونات الثلاثة معاً بحيث أنها تشكل مكوناً واحداً أم أن اثنين منها يتداخلان بدرجة كبيرة مع اختلافهما عن المكون المتبقي؟ وهل تحقق نماذج البناء الداخلي للقدرة اللفظية المكونة من عدد أقل من المكونات، والناجمة من اختزال المكونات إلى ثلاثة مكونات أو مكونين قيم محكات جودة ملائمة عالية ومقاربة مع النموذج الأصلي الكامل؟ ويمكن الإجابة عن هذه التساؤلات من خلال وضع نماذج مختزلة للمكونات تعكس جميع الاحتمالات الممكنة لنوعية التداخل، ومقارنتها مع بعضها، ومقارنتها مع نموذج المكونات الأربعة الأصلية مع وجود المكون العام من خلال التحليل العملي التوكيدي باستخدام بيانات النسخة 1030. وتم اختبار سبعة نماذج مختزلة (Reduced Models) كالتالي:

1. ستة نماذج تختزل عدد مكونات القدرة اللفظية إلى ثلاثة مكونات فقط بدلاً من أربعة مع وجود المكون العام، والتي تنتج من دمج أحد المكونات مع مكون آخر في كل نموذج مع إبقاء المكونين الآخرين كما هما.
 2. نموذج المكونين للقدرة اللفظية مع وجود المكون العام، والمتمثل في دمج المكونات الثلاثة (معاني الكلمات، والتناظر اللفظي، والاستيعاب القرآني) في مكون واحد مع بقاء مكون التناظر اللفظي.
- يظهر الجدول (10) قيم محكات جودة ملائمة النماذج الستة للمكونات الثلاثة لاختبار القدرة اللفظية في النسخة 1030. وتظهر النتائج أن قيم محكات جودة ملائمة النماذج الثلاثة الناتجة عن دمج مكونين من بين مكون معاني الكلمات، ومكون إكمال الجمل، ومكون الاستيعاب القرآني متقاربة جداً مع بعضها بعضاً، ومتقاربة جداً مع قيم محكات جودة ملائمة نماذج المكونات الأربعة الأصلية (الجدول 7). بينما أظهرت قيم محكات النماذج الثلاثة الأخرى الناتجة عن دمج مكون التناظر اللفظي مع أي من المكونات الثلاثة الأخرى قيمة أقل ملائمة عن نموذج المكونات الأربعة الأصلية (الجدول 7).

الجدول (9). معاملات الارتباط المتوقعة لفقرات اختبار القدرة اللفظية (النسخة 1030) مع المكونات الأربعة حسب التحليل العاملي التوكيدي

الفقرات	معاني الكلمات	إكمال الجمل	التناظر اللفظي	الاستيعاب القراني	الفقرات	معاني الكلمات	إكمال الجمل	التناظر اللفظي	الاستيعاب القراني
الفقرة 1	0.63	0.59	0.52	0.58	الفقرة 1	0.40	0.44	0.36	0.41
الفقرة 2	0.37	0.34	0.30	0.34	الفقرة 2	0.45	0.48	0.40	0.45
الفقرة 3	0.21	0.20	0.18	0.20	الفقرة 3	0.62	0.67	0.55	0.63
الفقرة 4	0.62	0.58	0.51	0.57	الفقرة 4	0.58	0.63	0.52	0.59
الفقرة 5	0.33	0.30	0.26	0.30	الفقرة 5	0.16	0.17	0.14	0.16
الفقرة 6	0.58	0.54	0.47	0.53	الفقرة 6	0.50	0.54	0.45	0.50
الفقرة 7	0.34	0.32	0.28	0.32	الفقرة 7	0.41	0.44	0.37	0.41
الفقرة 8	0.57	0.53	0.46	0.52	الفقرة 8	0.47	0.51	0.42	0.48
الفقرة 9	0.16	0.15	0.13	0.15	الفقرة 9	0.63	0.68	0.56	0.64
الفقرة	0.67	0.62	0.54	0.62	الفقرة	0.32	0.34	0.28	0.32
الفقرة	0.60	0.55	0.48	0.55	الفقرة	0.54	0.58	0.48	0.54
الفقرة	0.37	0.35	0.30	0.34	الفقرة	0.44	0.48	0.40	0.45
الفقرة	0.46	0.43	0.37	0.42	الفقرة	0.44	0.47	0.39	0.44
الفقرة 1	0.38	0.39	0.35	0.42	الفقرة	0.50	0.54	0.45	0.51
الفقرة 2	0.43	0.43	0.39	0.46	الفقرة	0.60	0.65	0.54	0.61
الفقرة 3	0.38	0.38	0.34	0.41	الفقرة	0.19	0.21	0.17	0.20
الفقرة 4	0.49	0.50	0.45	0.54	الفقرة 1	0.50	0.62	0.52	0.52
الفقرة 5	0.45	46.	0.41	0.49	الفقرة 2	0.56	0.70	0.58	0.58
الفقرة 6	0.59	0.60	0.53	0.64	الفقرة 3	0.44	0.55	0.45	0.46
الفقرة 7	0.43	0.44	0.39	0.47	الفقرة 4	0.44	0.13	0.11	0.44
الفقرة 8	0.13	0.13	0.12	0.14	الفقرة 5	0.41	0.51	0.42	0.43
الفقرة 9	0.28	0.29	0.26	0.31	الفقرة 6	0.46	0.57	0.47	0.48
الفقرة	0.64	0.65	0.58	0.70	الفقرة 7	0.42	0.52	0.44	0.44
الفقرة	0.38	0.39	0.34	0.41	الفقرة 8	0.49	0.61	0.51	0.51
الفقرة	0.42	0.43	0.38	0.46	الفقرة 9	0.58	0.72	0.60	0.60
الفقرة	0.48	0.49	0.43	0.52	الفقرة	0.49	0.60	0.50	0.50
الفقرة	0.45	0.46	0.41	0.49	الفقرة	0.24	0.30	0.25	0.25
الفقرة	0.46	0.47	0.42	0.50	الفقرة	0.53	0.66	0.55	0.55
الفقرة	0.40	0.41	0.36	0.44	الفقرة	0.56	0.70	0.58	0.58
الفقرة	0.15	0.16	0.14	0.17	الفقرة	0.45	0.56	0.46	0.46
الفقرة	0.30	0.31	0.28	0.33	الفقرة	0.57	0.71	0.59	0.59
الفقرة	0.30	0.31	0.28	0.33	الفقرة	0.43	0.53	0.44	0.44
الفقرة	0.28	0.29	0.26	0.31	الفقرة	0.41	0.51	0.42	0.42

الجدول (10). ملاءمة نماذج المكونات الثلاثة المختلفة مع المكون العام لاختبار القدرة اللفظية (النسخة 1030)

المحكات	معاني الكلمات مع إكمال الجمل	معاني الكلمات مع الاستيعاب القرائي مع إكمال الجمل	معاني الكلمات مع التناظر اللفظي	معاني الكلمات مع التناظر اللفظي مع الاستيعاب القرائي مع التناظر اللفظي	إكمال الجمل مع التناظر اللفظي
Chi-Square	87671	88217	87859	117131	104849
Chi-Square/df	42.23	42.49	42.28	56.39	50.50
RMSEA	0.036	0.036	0.036	0.043	0.042
GFI	0.92	0.92	0.92	0.89	0.90
AGFI	0.92	0.92	0.92	0.89	0.89
NFI	0.98	0.98	0.98	0.98	0.98
CFI	0.98	0.98	0.98	0.98	0.98
IFI	0.98	0.98	0.98	0.98	0.98

يظهر الجدول (11) قيم محكات جودة ملاءمة نموذج المكونين، والنتائج عن دمج كل من مكون معاني الكلمات، ومكون إكمال الجمل، ومكون الاستيعاب القرائي في مكون واحد مع إبقاء مكون التناظر اللفظي. وتشير النتائج إلى ملاءمة هذا النموذج نسبياً ولكن بدرجة أقل من ملاءمة نموذج المكونات الأربعة الأصلية (الجدول 7). وكذلك حصل هذا النموذج على قيمة لاختبار مربع كاي (90202) أكبر من قيم اختبار مربع كاي لأي من نماذج المكونات الثلاثة الناتجة عن دمج أي من مكون معاني الكلمات، أو إكمال الجمل، أو الاستيعاب القرائي (الجدول 10).

نتائج السؤال الخامس: الصدق البيئي (Cross Validation)

وللتأكد من صدق نتائج النماذج المختزلة للبناء الداخلي لاختبار القدرة اللفظية التي تم الحصول عليها في بيانات النسخة 1030، فحصت الدراسة جودة ملاءمة هذه النماذج في بيانات النسخة 1035، وهذه النماذج هي: نموذج المكونات الأربعة الأصلية مع المكون العام، ونموذج المكونات الثلاثة (دمج معاني الكلمات مع الاستيعاب القرائي) مع المكون العام، ونموذج المكونات الثلاثة (دمج معاني الكلمات مع إكمال الجمل) مع المكون العام، ونموذج المكونات الثلاثة (دمج معاني الكلمات مع إكمال الجمل) مع المكون العام، ونموذج المكونات الثلاثة (دمج معاني الكلمات مع إكمال الجمل والاستيعاب القرائي) مع المكون العام، ونموذج المكونات الثلاثة (دمج معاني الكلمات مع إكمال الجمل والاستيعاب القرائي) مع المكون العام. يظهر الجدول (12) قيم محكات جودة الملاءمة للنماذج الخمسة. وتشير هذه النتائج إلى تشابه النتائج التي تم الحصول عليها سابقاً باستخدام النسخة 1030، حيث أظهرت

وتدل هذه النتائج على ملاءمة نماذج المكونات الثلاثة المختزلة والنتيجة عن دمج أي مكونين من المكونات الثلاثة (مكون معاني الكلمات، ومكون إكمال الجمل، ومكون الاستيعاب القرائي)، وبمستويات متقاربة جداً من نموذج المكونات الأربعة الأصلي. بينما تدل النتائج على عدم ملاءمة جميع نماذج المكونات الثلاثة المختزلة والنتيجة عن دمج مكون التناظر اللفظي مع أي من المكونات الثلاثة الأخرى. ويمكن تلخيص هذه النتائج في أن النماذج الناتجة عن دمج أي مكونين من المكونات الثلاثة التالية (معاني الكلمات، إكمال الجمل، الاستيعاب القرائي) حصلت على محكات جودة ملاءمة متقاربة وجيدة ولا تختلف كثيراً عن محكات جودة ملاءمة نموذج المكونات الأربعة الأصلية مع المكون العام.

الجدول (11). ملاءمة نموذج المكونين مع المكون العام لاختبار القدرة اللفظية (النسخة 1030)

المحكات	مكونان (دمج معاني الكلمات وإكمال الجمل)
Chi-Square	90202
Chi-Square/df	43.41
RMSEA	0.036
GFI	0.92
AGFI	0.92
NFI	0.98
CFI	0.98
IFI	0.98

الاستيعاب القرائي) مع المكون العام على قيم جودة ملائمة أكثر قرباً من نموذج المكونات الأصلية. بينما حقق نموذج المكونين قيمة أقل جودة مقارنة ببقية النماذج.

أن نماذج المكونات الثلاثة حققت درجة ملائمة متقاربة جداً مع نموذج المكونات الأصلية. ومن بين هذه النماذج الثلاثة، حصل نموذج المكونات الثلاثة (دمج معاني الكلمات مع

الجدول (12). الصدق البيئي لمحكات جودة ملائمة النماذج الأربعة في اختبار القدرة اللفظية (النسخة 1035)

المحكات	3 مكونات (دمج معاني الكلمات، وإكمال الجمل) + المكون العام	3 مكونات (دمج معاني الكلمات، وإكمال الجمل) + المكون العام	3 مكونات (دمج معاني الكلمات، وإكمال الجمل) + المكون العام	4 مكونات + المكون العام	مكونان + المكون العام
Chi-Square	98345	97154	95622	90651	100275
Chi-Square/df	47.37	46.79	46.06	43.66	47.48
RMSEA	0.038	0.038	0.037	0.037	0.038
GFI	0.91	0.91	0.91	0.92	0.91
AGFI	0.91	0.91	0.91	0.92	0.91
NFI	0.98	0.98	0.98	0.98	0.98
CFI	0.98	0.98	0.98	0.98	0.98
IFI	0.98	0.98	0.98	0.98	0.98

الاختبار في المكونات الأصلية كما هو مفترض خاصة مكون معاني الكلمات، ومكون إكمال الجمل. فقد أشارت النتائج إلى قلة عدد الفقرات التي تشبعت في مكون معاني الكلمات، بينما توزعت بقية فقرات هذا المكون على مكون إكمال الجمل، ومكون الاستيعاب القرائي، كما أشارت النتائج إلى وجود فقرات مشتركة ومتداخلة بين ثلاثة مكونات وهي معاني الكلمات، والاستيعاب القرائي، وإكمال الجمل.

كما دعمت نتائج التحليل العاملي التوكيدي هذه النتائج حول البناء الداخلي لاختبار القدرة اللفظية من خلال قيم محكات جودة ملائمة عدد من النماذج المختزلة للبناء الداخلي لاختبار القدرة اللفظية. فقد أظهرت نتائج التحليل العاملي تقارب قيم محكات جودة ملائمة نماذج المكونات الثلاثة الناتجة من دمج كل من مكون معاني الكلمات، وإكمال الجمل، والاستيعاب القرائي (كل اثنين مع بعضهما) مع قيم محكات جودة نموذج المكونات الأربعة الأصلية. كما أظهر نموذج المكونين الناتج من دمج المكونات الثلاثة السابقة قيمة جيدة لمحكات جودة الملاءمة.

ومن خلال النظر إلى طبيعة المكونات الأربعة لاختبار القدرة اللفظية من الجانب النظري، وتعريف كل مكون من المكونات الأربعة، يمكن دعم نتائج النماذج المختزلة للبناء

مناقشة النتائج والاستنتاجات

سعت الدراسة إلى فحص صدق البناء الداخلي لاختبار القدرة اللفظية كأحد اختبارات القدرات العامة بالمركز الوطني للقياس والتقويم في التعليم العالي بالمملكة العربية السعودية من خلال التحليل العاملي الاستكشافي، والتحليل العاملي التوكيدي بطريقة تكاملية. أظهرت الدراسة تكاملاً بين نتائج التحليل العاملي الاستكشافي والتحليل العاملي التوكيدي حول البناء الداخلي لمكونات اختبار القدرة اللفظية. فأظهرت نتائج الدراسة أولاً مستوى ملائمة عالي للبناء الداخلي لاختبار القدرة اللفظية والمتمثلة في أربعة مكونات (معاني الكلمات، إكمال الجمل، الاستيعاب القرائي، التناظر اللفظي) مع المكون العام للقدرة اللفظية حسب مواصفات بناء الاختبار، حيث أظهرت النتائج قيم محكات جودة ملائمة عالية في التحليل العاملي التوكيدي. كما أظهرت نتائج الدراسة وجود مؤشرات لإمكانية تطوير البناء الداخلي لاختبار القدرة اللفظية. فأظهرت النتائج ارتباطات عالية بين ثلاثة مكونات لاختبار القدرة اللفظية وهي مكون معاني الكلمات، ومكون إكمال الجمل، ومكون الاستيعاب القرائي. فأظهرت نتائج التحليل العاملي الاستكشافي أن ثلاثة إلى أربعة عوامل كافية لتفسير درجات الطلاب في القدرة اللفظية. كما أظهرت النتائج عدم تشعب فقرات مكونات

إكمال الجمل متطلبا لقدرة الطالب على الاستيعاب القرائي، والإجابة عن الأسئلة حول النص المقروء؛ فلا يمكن للطالب أن يفهم ما يقرؤه على شكل فقرات إلا إذا كان متقنا لمعاني الكلمات المتضمنة، ومستوعبا للجمل المقروء، ومدركا للعلاقات بين الجمل (Andrew, Cobb & Giampietro, 2005). لذا تقترح الدراسة أيضا إمكانية قصر اختبار القدرة اللفظية على مكونين فقط: مكون التناظر اللفظي ومكون الاستيعاب القرائي مع إضافة عدد من الفقرات التي تقيس معاني الكلمات.

ويقترح أن يتضمن مكون الاستيعاب القرائي نصوصا قصيرة ونصوصا طويلة في مواضيع مختلفة لزيادة عدد النصوص المستخدمة وتنوعها لمراعاة التباين في الخلفيات والخبرات المختلفة للطلاب المتقدمين لاختبار القدرة اللفظية وزيادة مصداقية وثبات درجات القدرة اللفظية. وحيث يؤثر مستوى اللغة المستخدمة في النصوص في صدق قياس القدرة اللفظية، فإنه من المفضل استخدام فقرات وكلمات متنوعة ومتناسبة مع المستويات القرائية للطلاب المتقدمين لاختبار القدرة اللفظية.

ويساعد اختزال مكونات اختبار القدرة اللفظية إلى مكونين أو ثلاثة مكونات في تقليل عدد فقرات الاختبار وقصر زمن الاختبار، الأمر الذي يساعد في عدم إشعار الطالب بالملل من طول الاختبار وخاصة لدى الطلاب ذوي القدرات اللفظية العالية. كما يسمح بقياس قدرات أخرى إضافية في اختبار القدرة اللفظية مثل القدرة على الكتابة التي تعد قدرة أساسية للطلاب في مؤسسات التعليم العالي، وقد دخلت في العديد من اختبارات الاستعداد الدراسي مثل اختبار SAT واختبار ACT.

وبرغم ما قدمه كل من التحليل العاملي الاستكشافي، والتحليل العاملي التوكيدي في هذه الدراسة من معلومات وشواهد على صدق البناء الداخلي لاختبار القدرة اللفظية، تقترح الدراسة إجراء المزيد من الدراسات على البناء الداخلي لاختبار القدرة اللفظية باستخدام بيانات أخرى في الأعمار الدراسية المختلفة؛ لتأكيد النتائج التي تم الحصول عليها في هذه الدراسة. كما يمكن دراسة اختلاف البناء الداخلي لاختبار القدرة اللفظية حسب المناطق التعليمية للطلاب، ونوعهم الاجتماعي، وغيرها من البيانات الديموغرافية للطلبة. كما أنه من الضروري أيضا دعم نتائج التحليلات العاملية للبناء الداخلي لاختبار القدرة اللفظية بدراسات صدق حول العمليات العقلية التي يستخدمها الطلاب في الاستجابة عن فقرات الاختبار، سواء من خلال مراجعة المحكمين والمتخصصين في القدرة اللفظية أو تحليل العمليات العقلية التي يمر بها الطالب أثناء إجابته عن الأسئلة (ما يسمى بالتفكير بصوت عال). كما

الداخلي لاختبار القدرة اللفظية. فحيث إن فقرات مكون معاني الكلمات تتطلب من الطالب انتقاء مرادف ومعنى الكلمة المعطاة من بين عدد من الخيارات، فيمكن اعتبار مكون معاني الكلمات مهارة جزئية ومتطلبا لإكمال الجمل (عافشي، 2012)، فلا يمكن للطلاب أن يكمل جملة (كما في مكون إكمال الجمل) إلا إذا عرف معاني الكلمات المتضمنة في الجمل المعطاة. وكذلك الحال في العلاقة بين معاني الكلمات والاستيعاب القرائي، فلا يمكن للطلاب أن يستوعب ما يقرؤه (كما في مكون الاستيعاب القرائي) ويجيب عن الأسئلة المعطاة بعد قراءة النص المعطى إلا إذا كان ملما بمعاني الكلمات المستخدمة في النص (التل، ومقادي، 1989؛ Sternberg, 1985).

كما يتطلب استيعاب النص المقروء فهما من الطالب للجمل المستخدمة في النص المعطى، وقدرة على إدراك العلاقات بين الكلمات في الجملة الواحدة، والعلاقات بين الجمل المترابطة بأي طريقة ربط مناسبة، كالعلاقة السببية، والموضحة والمفسرة، والداعمة والمتناقضة والمستدركة وغيرها (عاشور، 2006). لذا لا يستطيع الطالب الإجابة عن أسئلة مكون الاستيعاب القرائي إلا إذا كان لدى الطالب قدرة على فهم تركيب الجمل، وإكمال الكلمات الناقصة لإتمام المعنى المقصود من الجملة المعطاة.

ومن بين الطرق الثلاث لإمكانية دمج مكونين من كل من معاني الكلمات، والاستيعاب القرائي، وإكمال الجمل في نموذج المكونات الثلاثة، أظهرت الدراسة أن نموذج المكونات الثلاثة الناتجة عن دمج مكون معاني الكلمات في الاستيعاب القرائي كان أعلى درجة ملاءمة نسبيا مقارنة بالنموذجين الآخرين. ويعضد هذه النتيجة العلاقة النظرية بين المكونات الثلاثة المذكورة سابقا. وتتفق هذه النتائج كذلك مع التعديلات التي أدخلت على اختبار SAT، المتمثلة في حذف مكون معاني الكلمات وتضمينها في مكون الاستيعاب القرائي (Burton, et. al., 2009). لذا تقترح الدراسة إمكانية تبني هذا النموذج في تطوير اختبار القدرة اللفظية ليتكون من ثلاثة مكونات وهي مكون التناظر اللفظي، ومكون إكمال الجمل، ومكون الاستيعاب القرائي، وإلغاء مكون معاني الكلمات وتضمينه في مكون الاستيعاب القرائي، من خلال إضافة فقرات تقيس معاني الكلمات كما وردت في سياق النص.

وحول نموذج المكونين لاختبار القدرة اللفظية، والمتمثل في دمج معاني الكلمات، والاستيعاب القرائي، وإكمال الجمل، فيمكن دعم هذا النموذج أيضا من خلال العلاقة النظرية بين كل من معاني الكلمات، وإكمال الجمل، والاستيعاب القرائي. إذ يمكن أن تعد القدرة على معرفة معاني الكلمات، والقدرة على

الكمية، والسمات التي تؤثر فيها كالتحصيل في التعليم العام، والسمات التي تتأثر بها كالتحصيل والنجاح في مؤسسات التعليم العالي وقطاعات سوق العمل.

توصي الدراسة بإجراء دراسات حول صدق البناء الخارجي External Structure Validity للاختبار من خلال دراسة علاقة القدرة اللفظية بغيرها من السمات المتشابهة معها كالكتابة والقدرات العقلية الأخرى، والسمات المختلفة عنها كالقدرة

assessing model fit. In K. A. Ballen & J. S. Long (Editors): *Testing Structural Equation Models*. Sage Publication.

Burton, N. W., Welsh, C., Kostin, I., & Van Essen, T. 2009. *Toward a Definition of Verbal Reasoning in Higher Education*. ETS, Princeton, New Jersey.

Chapman, A. 1993. *Making sense: Teaching critical reading across the curriculum*. New York: The college Board.

Christofferson, A. 1975. Factor analysis of dichotomized variables. *Psychometrika*, 40, 5-32.

Du Boulay, D. 1999. Argument in reading: What does it involve and how can students become better critical readers? *Teaching in Higher Education*, 4, 147-162.

Holdnack, J. A., Zhou, X., Larrabee, G. J., Millis, S. R. & Salthouse, T. A. 2011. Confirmatory factor analysis of the WAIS-IV/W). *MS-IV. Assessment*, 18, 178-191.

Jöreskog, K. & Sörbom, D. 2001. *Lisrel 8: User's reference guide*. 2nd edition, Lincolnwood: Scientific Software International.

Messick, S. 1993. Validity. In R. L. Linn (Ed.), *Educational measurement*, (3rd ed.). Phoenix, AZ: Oryx Press.

Nie, N. H., Golde, S. & Butler, D. M. 2009. *Education and verbal ability over time: Evidence from three multi-time sources*. The Stanford Institute for the Quantitative Study of Society: Stanford, CA.

Nist, S. T., & Simpson, M.L. 2000. College studying. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III). Mahwah, NJ: Lawrence Erlbaum Associates.

Nitko, A. J. & Brookhart, S. M. 2007. *Educational assessment of students*. (5th ed.) Upper Saddle River, NJ: Pearson Education.

Steiger, J. H. 1990. Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral research*, 25, 173-180.

Sternberg, R. J. 1985. *Components of verbal intelligence: Final report*. Ed.267090

Sweet, A. P. 1993. *State of the art: Transforming ideas for*

المراجع

أبو دنيا، نادية عبده، ونادية رشدي، 2002، أثر تثبيت شيوع العامل العام على العلاقة بين اختبارات القدرة اللفظية والقدرة الرياضية . دراسات تربوية وإجتماعية، مصر، 8.

الثل، شادية أحمد؛ ومحمد فخري مقدادي، 1989، دراسة تجريبية في تأثير استخدام الاسئلة التي تتطلب قدرات عقلية عليا في الاستيعاب القرائي. المجلة التربوية، الكويت، 6.

الحموري، هند عبدالمجيد وأحمد إسماعيل الكلوت، 2006، البناء الكامنة لاستبانة هني وممفورد لأنماط التعلم: تحليل عاملي توكيدي . مجلة العلوم التربوية والنفسية، البحرين، 7.

خير الله، سيد، 1978، سلوك الإنسان: أسسه النظرية والتجريبية. القاهرة: الأنجلو المصرية.

الصارمي، عبدالله محمد وعلي محمد إبراهيم، 2004، القدرة على الاستيعاب القرائي وعلاقتها بمتغيرات الجنس والتخصص ونوع النص لدى طلاب الصف الثاني الثانوي بسلطنة عمان. مجلة العلوم التربوية والنفسية، البحرين، 5.

عاشور، راتب قاسم، 2006، درجة تطبيق طلبة التربية العملية في الجامعة الأردنية - تخصص معلم صف - مهارات الاستيعاب القرائي في التدريس . مجلة كلية التربية - عين شمس، مصر، 30.

عاشفي، ابتسام عباس، 2012، فاعلية استراتيجية الأسئلة المستندة إلى التفكير في تنمية مهارات الاستيعاب القرائي والتفكير الناقد للنصوص الأدبية لدى طالبات قسم اللغة العربية بجامعة الأميرة نورة بنت عبدالرحمن. مجلة القراءة والمعرفة، مصر، 126.

عيد، غادة خالد؛ وآخرون، 2009، الخصائص السيكمترية والتحليل العاملي التوكيدي لمقياس أعراض اضطراب الوسواس القهري لدى عينة من طلاب جامعة الكويت. مجلة العلوم التربوية والنفسية - البحرين، 10.

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

Andrew, M. D., Cobb, C. D. & Giamietro, P. J. 2005. Verbal ability and teacher effectiveness. *Journal of Teacher Education*, 56, 343-354.

Browne, M. W. & Cudeck, R. 1993. Alternative ways of

635. teaching and learning to read. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement Education Information.
- Tilstra, J, McMaster, K., Van den Broek, P., Kendeou, P. & Rapp, D. 2009. Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading*, 32, 383-401.
- Tanka, J. S., & Huba, G. J. 1984. Confirmatory hierarchical factor analysis of psychological distress measures. *Journal of Personality and Social Psychology*, 46, 621-

Internal Structure Validity of Verbal Ability Test with Comparison among Various Alternative Structural Models: Integration between Exploratory and Confirmatory Factor Analysis

*Rashid Saif Almehrzi **

ABSTRACT

The study aimed at investigating the internal structure validity of the Verbal Ability Test in the General Abilities Test (GAT) developed by the National Center of Measurement in Higher Education in the Kingdom of Saudi Arabia using exploratory and confirmatory factor analysis. Two test forms were used. The results supported the current internal structure validity of the Verbal ability test according to the specified test components. For improving further the assessment of Verbal ability, the study aimed also at investigating the fit of various alternative reduced models for the current internal structure of the Verbal ability test. Four reduced models showed good model fit with very close values to the full model. Recommendations for internal structure of the verbal ability test and ways for improvement were provided.

Keywords: Internal Structure Validity, Exploratory and Confirmatory Factor Analysis, Verbal Ability.

* Sultan Qaboos University, Oman. Received on 23/10/2012 and Accepted for Publication on 13/5/2013.