بعض الخصائص السيكومترية لاختبار واطسون – جليسر للتفكير الناقد، الصورة "S" لدى عينة من طلاب الجامعة

دكتور/ عماد الدين محمد السكري

مدرس علم النفس التربوي

كلية التربية – جامعة المنوفية

ملخص البحث:

يهدف البحث الحالي إلى التعرف على بعض الخصائص السيكومترية لاختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" في البيئة المصرية. و تكونت العينة النهائية للبحث من (٥٠٣) طالبا وطالبة من طلاب وطالبات الفرقة الرابعة شعبة اللغة الإنجليزية بكلية التربية جامعة المنوفية، و تراوحت أعمار العينة ما بين (١٩) عاما، و(٢٣) عاما بمتوسط (٢١,٧٣)، وانحراف معياري (١,٥٣).

وأكدت النتائج صدق وثبات الاختبار، حيث تم حساب الصدق بثلاث طرق هي: صدق المحكمين، و الصدق العاملي، وصدق الاتساق الداخلي (معاملات الارتباط)، كما تم حساب الثبات بثلاث طرق أيضا هي: الاتساق الداخلي (معامل ألفا كرونباك)، والتجزئة النصفية، وإعادة الاختبار. أما التحليل العاملي فقد أكد صدق البناء العاملي للاختبارحيث أثبت وجود خمسة عوامل تتشبع عليها جميع مفردات الاختبار. وفسرت النتائج في ضوء الدرسات السابقة والإطار النظري.

وبشكل إجمالي أثبتت نتائج البحث تمتع الترجمة العربية الحالية لاختبار واطسون – جليسر الصورة المختصرة "S" بخصائص سيكومترية مقبولة في البيئة المصرية، وبالتالي صلاحيته للاستخدام فيها.

بعض الخصائص السيكومترية لاختبار واطسون – جليسر للتفكير الناقد، الصورة "S" لدى عينة من طلاب الجامعة

دكتور/ عماد الدين محمد السكري

مدرس علم النفس التربوي

كلية التربية – جامعة المنوفية

مقدمة:

في إطار الثورة المعلوماتية والتكنولوجية التي يشهدها العالم الآن لم يعد من المناسب أن يكون الشغل الشاغل لمناهجنا التعليمية حشو رؤوس التلاميذ بأكبر قدر من المعلومات، على حساب كيفية الحصول على المعلومات من مصادرها، وتوظيفها بشكل جيد من أجل حل مشكلات الإنسان، ومواكبة تطورات هذا العصر الذي يتسم بالتحول و التغير السريع.

والتفكير هو وسيلة الإنسان لحل المشكلات، والوصول إلى حلول إبداعية لما يواجهه من أزمات. وبه يواجه الإنسان هذا العالم بثقة، ويحل طلاسم الكون المستعصية، ويتحكم في الظواهر بل ويخضعها لمصلحته ومنفعته ورفاهيته، ويبدع وينتج ويضيف إلى صرح حضارته المزيد والمزيد من اللبنات.

ولأهمية موضوع التفكير في حياة الإنسان فقد نال مكانة خاصة في علم النفس المعاصر، وبخاصة في علم النفس المعرفي.وقد أكد ذلك بعض الباحثين بقولهم: إن سيكولوجية التفكير تمثل منزلة خاصة في علم النفس المعاصر، وإنه منذ بداية النصف الثاني من القرن العشرين شهد علم النفس اهتماماً متزايداً بما يسمى " بالعمليات المعرفية " إلى الحد الذي يدفعنا إلى القول بأن العصر الراهن هو عصر الاهتمام بسيكولوجية التفكير (سيد عثمان وفؤاد أبو حطب، ١٩٨٧، ١).

وقد صار التفكير الناقد في الآونة الأخيرة من العناوين الرنانة في الدوائر التربوية ، ويعلل فيشر (Fisher, 2001, p1) لهذا بقوله: أصبح المربون في غاية الاهتمام بتدريس "مهارات التفكير" في مقابل الاهتمام القديم بتدريس "المحتوى المعرفي" للمواد المختلفة، الذي كان التركيز منصبا عليه في الماضي. وكان المربون يردون على الداعين لتدريس مهارات التفكير بأنهم يفعلون ذلك بشكل ضمني وبطريقة غير مباشرة أثناء تدريسهم المحتوى المعرفي أو المادة التي يقومون بتدريسها. وبالتدريج بدأ المربون يشككون في فاعلية تدريس مهارات التفكير بهذه الطريقة التقليدية، لأن معظم الطلاب لا يمكنهم التقاط مهارات التفكير بهذه الطريقة. وبهذا تحول اهتمام

معظم المربين وكذلك الكثير من الأنظمة التربوية بتدريس مهارات التفكير بشكل مباشر، حيث تدرس بطريقة تسهل انتقالها إلى أي موضوعات أو محتويات أخرى.

إن التكنولوجيا الحديثة جعلت كم المعلومات المتاح أمام الطلاب من الضخامة بحيث يمكن وصفها بالانفجار المعرفي، وإزاء ذلك يجب أن يطور الطلاب أنفسهم من خلال استعمال مهارات التفكير الناقد وتنميتها حتى يكونوا قادرين على مواجهة المشكلات المعقدة التي سيواجهونها ولا يقفوا أمامها موقف القبول السلبي (Oliver & Utermohlen , 1995, p1).

وتعتبر مهارات التفكير أكثر قابلية للانتقال من محتوى إلى آخر، بما يسهل أمر استكشاف المعرفة وتعلمها على الطلاب بدلا من المعرفة التلقينية التي تتطلب بذل جهود جبارة في تحصيلها، فضلا عن الشك في فاعليتها وجدواها واستمرارها. ويضرب فيشر (Fisher, 2001, p1) مثلا يوضح أهمية تدريس مهارات التفكير الناقد: إذا تعلم شخص ما كيف يكوّن ويناقش ويحكم على مصداقية مصدر ما، أو يكون رأيا موضوعيا في موضوع معين باستخدام مهارات التفكير الناقد، فلن يكون من الصعب على هذا الشخص أن يعرف كيف يستخدم هذه المهارات مع محتويات أخرى كثيرة.

والعملية التعليمية الناجحة يجب أن تركز علي إكساب الأفراد مهارات التفكير الناقد التي تمكنهم من مواجهة التطورات السريعة والمتلاحقة في الحياة المعاصرة وما ينجم عنها من مشكلات تحتاج إلى هذا النوع من التفكير وبخاصة في ظل مفهوم العولمة الذي يجعل أبناءنا في مواجهه مباشرة مع ثقافات وأفكار واتجاهات عديدة لابد أن يكون قادرا علي فهمها والتعامل معها وعدم الوقوع في براثن أي أفكار خاطئة (عماد السكري، ٢٠٠٤، ٤٠).

ويحدد فاروق عثمان (١٩٩٣، ٥٥) أهمية التفكير الناقد في أنه:

(١) ضروري لسلوك حل المشكلات .

(٢) يساعد الفرد على حل المشكلات، دون التعصب لغرض معين أو فكرة سابقة أو رأي معين.

(٣) يساعد على استخدام النقد العلمي وعدم الانقياد للآراء التواترية والمعاني العاطفية.

(٤) يلعب دورا في تنمية القدرات التحليلية والجدلية.

(٥) يؤدي إلى المرونة العقلية للفرد مما يجعله يستخدم التفكير العقلاني المنطقي.

(٦) يهتم التفكير الناقد بأسس التقويم والاستدلال والاستنتاج المنطقي العقلاني.

ويذكر عبد العال عجوة، وعادل البنا (٢٠٠٠، ٥: ٨) عدة فوائد للتفكير الناقد، منها أنه:

(١) يعتبر بمثابة القاطرة المعرفية التي تقود عمليات حل المشكلة واتخاذ القرار.

(٢) يمكن الفرد من الحكم الحصيف والصائب على النظريات والأدلة والمحكات ومعايير الحكم وغيرها.

(٣) يجعل من الأفراد مفكرين مهرة حتى في المواقف ذات الضغط الانفعالي.

(٤) يساعد على النجاح المهني.

ويستخلص عماد السكري (٢٠٠٤، ٧٣: ٧٤) عددا من العوامل التي تزيد من أهمية التفكير الناقد، يجملها فيما يلي:

(١) يجعل الطالب أقدر على التفكير المتروي واستخدام المحكات المنطقية المناسبة للحكم على المعلومات المتاحة وفهم طبيعة المهمة.

(٢) يجعل الطلاب قادرين على إصدار الأحكام السليمة الموضوعية على المواقف والأحداث التي يتعرضون لها.

(٣) يسهم في تكوين شخصية تتسم بالذكاء في مواجهة مشكلات الحياة المعقدة، قادرة على الاعتماد على النفس للوصول إلى حقائق الأشياء.

(٤) يساعد على إحداث نهضة فكرية وثقافية ، كمقدمة للنهضة الشاملة لأي مجتمع من المجتمعات.

(٥) يعود الفرد على الاستدلال المنطقي ، وتجنب الأخطاء ، والمعتقدات الخاطئة، ويبعده عن كل ما له صلة بالعوامل الذاتية ، كما يعوده على الانتباه إلى المغالطات وإدراك المتناقضات.

(٦) يدعم كل الأهداف التربوية الأخرى ويساعد على تحقيقها.

(٧) ضروري لصور التفكير الأخرى، فهو عنصر من عناصر حل المشكلة واتخاذ القرار، كما أنه أساس لعملية التفكير الابتكاري.

(٨) يعد مطلبا ضروريا لجميع أفراد المجتمع وبخاصة الشباب الذين هم أمل الأمة في تطورها وتقدمها.

ونتيجة لما سبق من أسباب، أصبح التفكير الناقد في الوقت الراهن من أهم أولويات برامج التعليم العام عامة، وفي برامج التعليم العالي بصفة خاصة (Geisinger, 1998).

وبقدر أهمية موضوع التفكير الناقد بقدر الحاجة إلى وجود مقاييس مناسبة لقياس هذه القدرة ومكوناتها الفرعية، فما لا يمكن قياسه لا يمكن تحقيقه، ولذلك فقد تعددت المقاييس والاختبارات التي وضعت لقياس القدرة على التفكير الناقد ومهاراته، واختلفت هذه الاختبارات وتنوعت تبعا لتنوع الأطر النظرية التي تستند عليها.

ومن الملاحظ ارتباط قياس التفكير الناقد ارتباطا وثيقا ببرامج تنميته، فالكثير من المحاولات التي بذلت لتحليل مكونات التفكير الناقد، وتحديد المهارات التي ينطوي عليها قد أعانت على توضيح الفكرة الخاصة بإمكانية قياسه، كما ساعدت أيضا على إمكانية وضع البرامج المختلفة لتنمية هذه المهارات أو المكونات التي يشتمل عليها (عزيزة السيد، ١٩٩٥، ٦٩).

فتنمية القدرة على التفكير الناقد إذاً رهن بوجود مقاييس مناسبة لقياس هذه القدرة ومكوناتها الفرعية، وقد لاحظ الباحث أن الكثير من مقاييس التفكير الناقد في البيئة العربية تحتاج إلى إعادة نظر إما بسبب قدمها وحاجتها إلى إعادة نظر في مدى مناسبتها للتطبيق في الوقت الراهن، أو بسبب طولها وحاجتها إلى بذل جهد ووقت مضاعف في تطبيقها. وقد جرب الباحث ذلك شخصيا أثناء تطبيقه لترجمة جابر عبد الحميد ويحيى هندام لاختبار واطسون وجليسر للتفكير الناقد المنشور عام ١٩٦٤ أثناء إعداد أطروحته للدكتوراه نظرا لعدم فهم الطلاب لبعض العبارات التي يشتمل عليها الاختبار بسبب عدم مناسبتها للوقت الراهن، أو بسبب إجهاد الطلاب من طول الاختبار (مائة عبارة).

هذا ويعتبر اختبار واطسون – جليسر للتفكير الناقد من أهم الاختبارات التي وضعت لقياس هذه القدرة وأكثرها انتشارا ومصداقة. (Ayala, 1999; Adams, 1999; Hicks-Moore & Pastirik, 2006)

مشكلة البحث:

يتبنى البحث الحالي اختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" WGCTA – Short Form للأسباب الآتية:

(١) تمتع الاختبار الأصلي بالكفاءة السيكومترية الجيدة (الثبات، الصدق) في البيئة الأجنبية (Berger, 1985; Helmstadter, 1985)، وكذلك الحال بالنسبة للصورة المختصرة التي يعنى بها البحث الحالي (Gadzella et al., 2005).

(٢) الشهرة العالمية والانتشار الواسع اللذان حظي بهما هذا الاختبار منذ صدوره لأول مرة عام

١٩٤٢، مرورا بالإصدارات المختلفة له عام ١٩٦٤ وعام ١٩٨٠، وحتى الإصدار الأخير الذي صدر لأول مرة عام ١٩٩٤، والذي قام الباحث الحالي بترجمته إلى اللغة العربية، ويعتبر هذا الاختبار كما يذكر ماكميلان (McMillan, 1987, p9) أكثر اختبارات التفكير الناقد انتشارا على مستوى العالم. والبحث الحالي بصدد دراسة بعض الخصائص السيكومترية للترجمة العربية للنسخة المختصرة لهذا الاختبار واسع الانتشار.

(٣) الميزة العملية لهذا الإصدار من الاختبار، حيث يوفر الكثير من الوقت والجهد والتكلفة ، فهذا الإصدار يتكون فقط من أربعين مفردة، لا تستغرق الإجابة عنها أكثر من ثلاثين دقيقة إضافة إلى خمس إلى عشر دقائق أخرى إضافية لقراءة التعليمات، أما الإصدارات الأقدم من هذا الاختبار فقد كانت تتكون من مائة مفردة في إصدار عام ١٩٦٤ (وهو الإصدار الذي ترجمه إلى العربية جابر عبد الحميد ويحيى هندام)، وثمانين مفردة في إصدار عام ١٩٨٠، مما يزيد من صعوبة وتكلفة ووقت إجرائه الذي يزيد عن ضعف وقت الإصدار الحالي، كما يقلل من الميزة العملية لتلك الإصدارات الأقدم وبخاصة عند استخدامه لأغراض الاختيار الوظيفي.

مما سبق يمكن تحديد مشكلة البحث الحالي في دراسة بعض الخصائص السيكومترية للترجمة العربية لاختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة"، من خلال الإجابة عن التساؤلات الآتية:

(١) ما طبيعة البناء العاملي لاختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" لدي طلاب الجامعة ؟

(٢) هل تتوافر لاختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" درجة مقبولة من الثبات لدي طلاب الجامعة ؟

(٣) هل تتوفر لاختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" درجة مقبولة من الصدق لدي طلاب الجامعة ؟

أهداف البحث:

يهدف البحث الحالي إلى التعرف على بعض الخصائص السيكومترية لاختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" في البيئة العربية (المصرية تحديدا)، وبيان مدى صلاحيته لقياس القدرة على التفكير الناقد ومهاراته، ومن ثم صلاحيته للاستخدام في هذه البيئة.

أهمية البحث:

تكمن أهمية البحث الحالي في توفير نسخة عربية لأحد أهم اختبارات التفكير الناقد، وأكثرها انتشارا على المستوى العالمي، واختبار الخصائص السيكومترية التي يتمتع بها هذا الاختبار، مما يوفر للباحثين في البيئة العربية أداة هامة يمكن الاعتماد عليها والثقة فيها لما لها من خصائص سيكومترية مناسبة في هذه البيئة.

وقد تم اختيار هذا الاختبار للأسباب التي ذكرها الباحث سابقا في مشكلة البحث، بالإضافة إلى أهمية هذه الصورة المختصرة التي بدأت تستخدم بشكل واسع في البيئات المختلفة لما تتمتع به من صفات أجملها جيسينجر (Geisinger, 1998) في ثلاث صفات هي: القصر، وسرعة التطبيق، وسهولة الاستخدام بما يتلاءم مع روح العصر وإيقاعه السريع.

مصطلحات البحث:

التفكير الناقد:

يعرف واطسون وجليسر (Watson & Glaser, 2008, p3) التفكير الناقد بأنه: " مزيج من الاتجاهات، والمعارف، والمهارات يشتمل على:

- القدرة على التعرف على المشكلات القائمة، وتقبل الاحتياج العام إلى إيجاد الدليل الذي يدعم ما يعتقد أنه صحيح.
- المعرفة بطبيعة الاستنتاجات والتجريدات والتعميمات الصالحة، والتي يحدد على أساسها وزن ودقة مختلف أنواع الأدلة بطريقة منطقية.
- المهارة في توظيف وتطبيق الاتجاهات، والمعارف السابقة.

الإطار النظري:

اختبار واطسون – جليسر للتفكير الناقد:

تم تصميم اختبار واطسون – جليسر للتفكير الناقد من أجل قياس القدرة على التفكير الناقد، والمهارات الفرعية المتضمنة فيها. ووفقا لمؤلفي الاختبار فإن جذور هذا الاختبار تمتد إلى الجهود التي قام بها كل منهما على حدة قبل عام ١٩٦٤، حيث بدأ جودوين واطسون جهوده في تصميم اختبار للتفكير الناقد عام ١٩٢٥ حيث نشر اختبارا بعنوان Measure of Fair mindedness

(Moore, 1976, p73)، كما قام إدوارد جليسر بجهود مشابهة منذ عام ١٩٣٧ وحتى حصوله على درجة الدكتوراه من جامعة كولومبيا ١٩٤٠ بعنوان An Experiment In The Development Of Critical Thinking (Gordon, 1994, p755).وتم تطوير هذه الجهود من خلال العمل المشترك بينهما إلى أن قاما بإصدار النسخة الأولى المعتمدة لهذا الاختبار عام ١٩٦٤.

وكان هذا الإصدار يتكون من صورتين متكافئتين للاختبار، الصورة "YM" والصورة "ZM"، وكل منهما يتكون من مائة سؤال موزعة على خمسة اختبارات فرعية كما يلي:

١ – اختبار الاستنتاج Inference ويتكون من عشرين مفردة.

٢ – اختبار التعرف على الافتراضات Recognition of Assumptions ويتكون من ست عشرة مفردة.

٣ – اختبار الاستنباط Deduction ويتكون من خمس وعشرين مفردة.

٤ – اختبار التفسير Interpretation ويتكون من أربع وعشرين مفردة.

٥ – اختبار تقويم الحجج Evaluation of Arguments ويتكون من خمس عشرة مفردة.

ويستغرق الفرد حوالي خمسين دقيقة للإجابة عن مفردات هذا الاختبار (Watson & Glaser, 2008,p3).

ويذكر كرايتز (Crites, 1965, p335) أنه لا توجد معلومات متاحة عن ثبات هذا الاختبار عبر الزمن، ولا عن مدى قدرته على التنبؤ (الصدق التنبؤي)، ولكن ثبات التكوين الداخلي لهذا الاختبار كان مرتفعا، وكذلك كان الصدق التلازمي له مقبولا.

وقد قام كل من جابر عبد الحميد ويحيى هندام بترجمة الصورة "YM" لهذا الاختبار إلى اللغة العربية، ويعد من أكثر اختبارات التفكير الناقد انتشارا في البيئة العربية حتى وقتنا الراهن.

ومنذ نشر الاختبار عام ١٩٦٤ استمرت جهود واطسون وجليسر في تعديل هذا الاختبار إلى أن أصدرا النسخة المعدلة منه عام ١٩٨٠، تلك النسخة التي قام فيها الباحثان بتغيير بعض العبارات، وإعادة صياغة البعض الآخر، وحذف العبارات التي تتضمن أي إشارات للنواحي الجنسية أو العرقية. وقد تكون هذا الإصدار من صورتين متكافئتين "A"، و "B" بدلا من الصورتين "YM" ، و"ZM"، وأصبحت كل صورة منهما تتكون من ثمانين مفردة بدلا من مائة

مفردة في الصورتين القديمتين، كما يستغرق الفرد حوالي أربعين دقيقة بدلا من خمسين للإجابة عن مفرداته (Watson & Glaser, 2008,p3).

أما الصورة المختصرة لهذا الاختبار والتي يعنى بها الباحث في بحثه الحالي لما لها من أهمية سيتم توضيحها في سياق آخر فقد تم نشرها لأول مرة عام ١٩٩٤، حيث أظهرت الدراسات التي عنيت بهذا الاختبار أن كثرة عدد مفردات الإصدارات القديمة منه (مائة مفردة في إصدار عام ١٩٦٤، وثمانون مفردة في إصدار عام ١٩٨٠) تزيد من صعوبة وتكلفة ووقت إجرائه، كما تقلل من الميزة العملية لهذا الاختبار، وبخاصة عند استخدامه لأغراض الاختيار الوظيفي. لذلك فقد تم انتقاء أربعين مفردة من مفردات الصيغة "A"، وأطلق عليها الصورة المختصرة أو الصورة "S"، والتي لا تزيد مدة أدائها عن ثلاثين دقيقة إضافة إلى خمس إلى عشر دقائق أخرى إضافية لقراءة تعليمات هذا الاختبار (Watson & Glaser, 2008,p4).

مكونات اختبار واطسون – جليسر للتفكير الناقد:

وفقا لما أورده واطسون وجليسر (Watson & Glaser, 2008, pp 3-4) فإن اختبار واطسون – جليسر للتفكير الناقد في جميع صوره يتكون من خمسة اختبارات فرعية يقيس كل منها مهارة من مهارات التفكير الناقد، والتي حددها واضعا الاختبار فيما يلي:

١ – الاستنتاج Inference ويعني قدرة الفرد على تمييز درجات احتمال صحة أو خطأ استنتاج معين بناء على بعض المعلومات المتاحة له.

٢ – التعرف على الافتراضات Recognition of Assumptions ويعني قدرة الفرد على فحص بعض المعلومات المتاحة له بحيث يستطيع أن يحكم على فرض ما بأنه وارد أو غير وارد بناء على هذه المعلومات.

٣ – الاستنباط Deduction ويعني قدرة الفرد على الحكم على مدى صحة اشتقاق نتيجة معينة من معلومات متاحة له بغض النظر عن مدى صحة أو خطأ تلك المعلومات.

٤ – التفسير Interpretation ويعني قدرة الفرد على الموازنة بين الأدلة والوصول إلى قرار على درجة معقولة من اليقين بناء على بعض المعلومات المتاحة.

٥ – تقويم الحجج Evaluation of Arguments ويعني قدرة الفرد على التمييز بين الحجج القوية المتعلقة بقضية معينة، والحجج الضعيفة التي ليس لها صلة بنفس القضية.

وكل من الاختبارات الخمس السابقة يتكون من عدد من الفقرات القرائية أو السيناريوهات

Scenarios التي تشتمل على مشكلات وأحكام ومناقشات وتفسيرات لمعلومات مشابهة لما نواجهه في حياتنا اليومية: في العمل، وفي الفصول الدراسية، وعلى صفحات الجرائد والمجلات. وكل سيناريو من هذه السيناريوهات متبوع بعدد من المفردات التي يطلب من الفرد المفحوص الحكم عليها.

وهناك نوعان من السيناريوهات في هذا الاختبار: النوع الأول سيناريوهات حيادية Neutral تتعامل مع الموضوعات والحداث التي لا تسبب أي تحيزات أو مشاعر من قبل الفرد مثل الموضوعات المتعلقة بالطقس، والحقائق العلمية، والتجارب العملية. أما النوع الثاني من السيناريوهات فهو سيناريوهات خلافية أو جدلية Controversial وهي التي تتعامل مع الموضوعات ذات المحتوى الجدلي مثل الموضوعات المتعلقة بالسياسة، والاقتصاد، والقضايا الاجتماعية التي تستثير استجابات انفعالية قوية من قبل الفرد.

البحوث والدراسات السابقة:

أجرى لو، وثورب (Loo & Thorpe, 1999) دراسة للتعرف على الخصائص السيكومترية والتكوين العاملي لاختبار واطسون وجليسر للتفكير الناقد، وقد طبق الباحثان هذه الدراسة على الصورة S من الاختبار والتي صدرت عام 1994. بلغت عينة الدراسة ٢٧١ طالبا وطالبة (١٧٣ من الإناث، و٩٨ من الذكور)، من كليتين مختلفتين هما كلية الإدارة (١٤٢)، وكلية التمريض (١٢٣) وهناك ستة مفحوصين لم يذكروا تخصصاتهم. وتم حساب معاملات الارتباط البينية للدرجة الكلية للاختبار ودرجات الاختبارات الفرعية للعينة الكلية، ثم للمجموعتين الفرعيتين (الإدارة – التمريض)، وقد تراوحت معاملات الارتباط بين درجات الاختبارات الفرعية والدرجة الكلية للاختبار بين ٠,٥٣، و٠,٦٤ للعينة الكلية ، كما تراوحت من ٠,٤٩ إلى ٠,٦٨ للعينتين الفرعيتين، وكل هذه الارتباطات دالة إحصائيا عند مستوى دلالة ($p< 0.01$).

وقام الباحثان بحساب ثبات الاختبار (الدرجة الكلية والأبعاد) بطريقة الاتساق الداخلي Internal Consistency Reliability، وبلغت درجة معاملات ألفا كرونباك كما يلي: بالنسبة للعينة الكلية، ٠,٧١ للدرجة الكلية للاختبار، وتراوحت من ٠,٢٥ إلى ٠,٧٣ للاختبارات الفرعية. وبالنسبة لمجموعة الإدارة، ٠,٧٤ للدرجة الكلية للاختبار، وتراوحت من ٠,٣٠ إلى ٠,٧٢ للاختبارات الفرعية. أما بالنسبة لمجموعة التمريض فكان معامل ألفا ٠,٦٥ للدرجة الكلية للاختبار، وتراوحت من ٠,١٧ إلى ٠,٧٤ للاختبارات الفرعية. كما قام الباحثان بإجراء تحليل عاملي توكيدي Confirmatory Factor Analysis للتأكد من البنية العاملية للاختبار ولكن نتائج

هذا التحليل لم تدعم البنية العاملية المفترضة لهذا الاختبار، لذلك قام الباحثان بإجراء تحليل عاملي استكشافي Exploratory Factor Analysis ولكن نتائج هذا التحليل أيضا لم تدعم البنية العاملية حيث إن بعض المفردات قد تشبعت على غير العوامل التي تنتمي إليها، وبعضها لم يتشبع على أي عامل. وبشكل عام فإن هذا النوع من التحليل العاملي قد أكد تطابقا كبيرا للعامل الثاني فقط من العوامل الخمس المقترحة، أما في بقية العوامل فالتطابق ضعيف إلى متوسط.

ونتائج الدراسة السابقة تشير إلى ضعف الخصائص السيكومترية للاختبار الحالي متمثلة في ضعف مؤشرات ثبات وصدق الاختبار مع أفراد العينة الحالية، وبخاصة الاختبارات الفرعية، حيث إن معاملات ثبات الاختبارات الفرعية ضعيفة إلى متوسطة. كما أن نتائج التحليل العاملي التوكيدي والاستكشافي لم تقدم دعما للبناء العاملي الذي يقوم الاختبار على أساسه.

أما دراسة ويليامز (Williams, 2003) فقد هدفت إلى تحقيق هدفين، أولهما المقارنة بين نوعين من اختبارات التفكير الناقد (الاختبارات العامة – الاختبارات محددة الموضوع) في التنبؤ بالأداء الأكاديمي لدى عينة من طلاب الجامعة الدارسين لمقرر علم النفس التربوي. وثانيهما تحديد إلى أي مدى تتأثر درجة الفرد على اختبار التفكير الناقد من بداية الفصل الدراسي إلى نهايته.

وبلغت العينة الكلية للدراسة ٤٢٨ من طلاب وطالبات الجامعة الدارسين لمقرر علم النفس التربوي طبق عليهم الباحث ثلاثة اختبارات: اختبار واطسون وجليسر للتفكير الناقد – الصورة S، واختبار كاليفورنيا لمهارات التفكير الناقد، واختبار التفكير الناقد النفسي. والاختباران الأول والثاني من اختبارات التفكير الناقد العامة، أما الاختبار الثالث فهو من الاختبارات المحددة الموضوع.

ودلت نتائج الدراسة على أن اختبار واطسون وجليسر للتفكير الناقد – الصورة S يعد أقوى الاختبارات في التنبؤ بالأداء الأكاديمي عند مقارنته بكل من اختبار كاليفورنيا، واختبار التفكير الناقد النفسي. كما دلت النتائج على وجود فروق بين الطلاب المرتفعين في الأداء الأكاديمي والطلاب المنخفضين في الأداء الأكاديمي في درجاتهم على اختبارات التفكير الناقد لصالح الطلاب المرتفعين في الأداء الأكاديمي. كما وجد الباحث أن نتائج أن هناك فروقا في درجات الطلاب على اختبار واطسون وجليسر ، واختبار التفكير الناقد النفسي بين التطبيق قبل وبعد الفصل الدراسي لصالح التطبيق البعدي.

وتدل نتائج الدراسة السابقة على صدق الاختبار من خلال إثباتها لنوعين من الصدق: أولهما الصدق التنبؤي (حيث كان منبئا بالأداء الأكاديمي)، والصدق التمييزي من خلال تمييزه بين مرتفعي ومنخفضي الأداء الأكاديمي.

وقام جادزيلا وبالوجلو (Gadzella & Baloglu, 2003) بدراسة هدفت إلى التعرف على بعض الخصائص السيكومترية لاختبار واطسون وجليسر للتفكير الناقد. وطبق الباحثان هذه الدراسة على الصورة A من الاختبار والتي صدرت عام ١٩٨٠ مع الصورة B. وقد بلغت عينة الدراسة ١٣٥ طالبا من الطلاب المختارين لمادة علم النفس التربوي في جامعة Southwestern State، حيث كان عدد الذكور ٣٠ طالبا بنسبة (٢٢,٢%)، و١٠٥ طالبة بنسبة (٧٧,٨%). وقد تراوحت أعمار هؤلاء الدارسين من ١٩ إلى ٥٤ عاما بمتوسط حسابي قدره ٢٩,٧، وانحراف معياري قدره ٩,٠٠.

وقام الباحثان بحساب ثبات الاختبار بطريقتين، الأولى: طريقة ثبات الاتساق الداخلي Internal Consistency Reliability، وقد بلغت درجة معاملات ألفا كرونباك ٠,٨٦ للدرجة الكلية للاختبار، و٠,٥٨ لاختبار الاستنتاج، و٠,٨٣ لاختبار التعرف على الافتراضات، و ٠,٥٧ لاختبار الاستنباط، و٠,٦٣ لاختبار التفسير، و ٠,٧٠ لاختبار تقويم الحجج. أما الطريقة الثانية لحساب الثبات فهي طريقة التجزئة النصفية Split-Half Reliability، عن طريق حساب مقارنة مفردات الأسئلة الفردية والزوجية للاختبار ككل، وقد جاء معامل الثبات بهذه الطريقة للدرجة الكلية للاختبار ٠,٧٥.

وقام الباحثان بحساب الصدق التلازمي للاختبار Concurrent Validity عن طريق حساب معاملات الارتباط بين درجات الطلاب على الاختبار (الدرجة الكلية، والأبعاد)، ودرجات الأفراد على مادة علم النفس التربوي التي يدرسونها. وقد بلغت معاملات ارتباط درجات الأفراد على هذه المادة والدرجة الكلية للاختبار ٠,٤٢، وبينها ودرجة اختبار الاستنتاج ٠,٣٢، وبينها ودرجة اختبار التعرف على الافتراضات ٠,٢٠، وبينها ودرجة لختبار الاستنباط ٠,٣٩، وبينها ودرجة اختبار التفسير ٠,٣٧، وبينها واختبار تقويم الحجج ٠,١٩. وكلها ارتباطات دالة عند مستوى دلالة ($p< 0.05$)، ما عدا ارتباط الدرجة الكلية للاختبار فهو دال عند مستوى دلالة ($p< 0.01$).

وتدل نتائج الدراسة السابقة على صدق وثبات اختبار واطسون وجليسر للتفكير الناقد الصورة A، إلا أن الملاحظ أن درجات ثبات الاختبارات الفرعية كانت منخفضة نسبيا عن الدرجة الكلية للاختبار وبخاصة لاختباري الاستنباط، والاستنتاج، كما أن درجات الصدق التلازمي للدرجة الكلية والاختبارات الفرعية رغم أنها جميعا دالة إحصائيا إلا أن أقوى الارتباطات كانت بين الدرجة الكلية واختبار مادة علم النفس التربوي، بما يؤكد أن درجة صدق الدرجة الكلية أعلى من درجات صدق الاختبارات الفرعية. وقد يكون هذا طبيعيا في ظل هذا المفهوم متعدد الأبعاد للتفكير الناقد.

فالتكوينات متعددة الأبعاد عادة ما يكون صدق وثبات درجاتها الكلية أعلى من صدق وثبات الاختبارات الفرعية لها[1].

كما أجرى جادزيلا وزملاؤه (Gadzella et al., 2005) دراسة أخرى هدفت إلى التعرف على بعض الخصائص السيكومترية لاختبار واطسون وجليسر للتفكير الناقد- الصورة S على عينة قوامها ١٣٧ طالبا من الطلاب المختارين لمادة علم النفس التربوي في جامعة Southwestern State، حيث كان عدد الذكور ٢٨ طالبا ، وعدد الإناث ١٠٩ طالبة.

وقام الباحثون بحساب ثبات الاختبار بطريقتين، الأولي: طريقة الاتساق الداخلي Internal Consistency Reliability، و بلغت درجة معاملات ألفا كرونباك ٠,٧٦ للدرجة الكلية للاختبار، و٠,٢٩ لاختبار الاستنتاج، و٠,٨٠ لاختبار التعرف على الافتراضات، و ٠,٤٢ لاختبار الاستنباط، و٠,٣١ لاختبار التفسير، و ٠,٣٢ لاختبار تقويم الحجج. أما الطريقة الثانية لحساب الثبات فهي طريقة التجزئة النصفية Split-Half Reliability، و جاء معامل الثبات بهذه الطريقة للدرجة الكلية للاختبار ٠,٤٤.

أما صدق الاختبار فقد قام الباحثون بحساب الصدق التلازمي للاختبار Concurrent Validity عن طريق حساب معاملات الارتباط بين الدرجة الكلية للاختبار، ودرجات الاختبارات الفرعية، ودرجات الأفراد على مادة علم النفس التربوي التي يدرسونها. وبلغت معاملات ارتباط درجات الأفراد على هذه المادة والدرجة الكلية للاختبار ٠,٣١، وبينها ودرجة اختبار الاستنتاج ٠,١٥، وبينها ودرجة اختبار التعرف على الافتراضات ٠,٠٤، وبينها ودرجة اختبار الاستنباط ٠,٣٣، وبينها ودرجة اختبار التفسير ٠,١٠، وبينها واختبار تقويم الحجج ٠,٣٢. ويوجد ثلاثة من الارتباطات السابقة دالة عند مستوى دلالة ($p< 0.01$) وهم ارتباطات الدرجة الكلية، واختبار الاستنباط، واختبار تقويم الحجج، كما يوجد ارتباط وحيد دال عند مستوى دلالة ($p< 0.05$) وهو الارتباط الخاص باختبار الاستنتاج، أما الارتباطان الخاصان باختباري التعرف على الافتراضات، والتفسير فهما غير دالين إحصائيا.

وتشير نتائج الدراسة السابقة إلى أن درجات ثبات وصدق الصورة S من اختبار واطسون جليسر للتفكير الناقد أقل من نظيرتها التي أجريت باستخدام الصوررة A سنة ٢٠٠٣، والتي سبق عرضها، إلا أن النتائج بشكل عام مشجعة وتدل على صلاحية هذه الصورة لقياس التفكير الناقد،

---

[1] هذه الملحوظة سمعتها شفاهة من Professor William Bart، أستاذ علم النفس التربوي بجامعة مينيسوتا في لقاء بيننا أثناء مهمتي العلمية بالولايات المتحدة عام ٢٠٠٩.

وخاصة الدرجة الكلية منه. ويوصي الباحثون بتكرار هذه الدراسة على عينات أكبر للتأكد من الخصائص السيكومترية للاختبار.

وقام جادزيلا وزملاؤه (Gadzella et al., 2006) بإجراء دراسة أخرى هدفت إلى التعرف على صدق وثبات اختبار واطسون وجليسر للتفكير الناقد، الصورة S على عينة نهائية قوامها ٥٦٥ طالبا وطالبة من مختلف تخصصات علم النفس في جامعة Southwestern State، منهم ٤٨٦ من طلاب وطالبات المرحلة الأولى الجامعية، و ٧٩ من طلاب وطالبات الدراسات العليا.

و قام الباحثون بحساب ثبات الاختبار (الدرجة الكلية فقط) بطريقة الاتساق الداخلي Internal Consistency Reliability، و بلغت درجة معاملات ألفا كرونباك للدرجة الكلية للاختبار ٠,٩٢ لدرجات جميع أفراد العينة، و٠,٧٥ لدرجات طلاب وطالبات علم النفس العام، و٠,٧٤ لدرجات طلاب وطالبات علم النفس التربوي، و ٠,٧٦ لدرجات طلاب وطالبات التربية الخاصة، و٠,٩٢ لدرجات طلاب وطالبات المرحلة الجامعية الأولى، و ٠,٧٨ لدرجات طلاب وطالبات الدراسات العليا.

أما صدق الاختبار فقد قام الباحثون بحساب الصدق التلازمي للاختبار Concurrent Validity عن طريق حساب معاملات الارتباط بين الدرجة الكلية للاختبار، ودرجات الأفراد في امتحان نهاية الفصل الدراسي في تخصصاتهم. و بلغت معاملات الدرجة الكلية للاختبار ٠,٣٠ مع درجات جميع أفراد العينة، و٠,٦٢ مع درجات طلاب وطالبات علم النفس العام، و٠,٣٨ مع درجات طلاب وطالبات علم النفس التربوي، و ٠,٢٤ مع درجات طلاب وطالبات التربية الخاصة، و٠,٢٠ مع درجات طلاب وطالبات المرحلة الجامعية الأولى، و٠,٣٦ مع درجات طلاب وطالبات الدراسات العليا. وجميع الارتباطات السابقة دالة عند مستوى دلالة ($p< 0.01$) عدا ارتباط درجات المجمع الكلي للطلاب فهو دال عند مستوى دلالة ($p< 0.05$).

ونتائج الدراسة السابقة تدل على صدق وثبات اختبار واطسون وجليسر للتفكير الناقد – الصورة S، ولكننا نلاحظ أن الباحثين قد اهتموا بصدق وثبات الدرجة الكلية للاختبار متجاهلين درجات الاختبارات الفرعية، ويبررون ذلك بأن واضعي الاختبار قد نوها إلى أن الدرجة الكلية للاختبار أكثر دقة من درجات مقاييسه الفرعية.

تعليق على الدراسات السابقة:

لاحظ الباحث الحالي الاختلاف الواضح إلى حد التناقض بين نتائج الدراسات السابقة، فبينما أكدت نتائج دراسة (Loo & Thorpe, 1999) ضعف الخصائص السيكومترية للاختبار الحالي

متمثلة في ضعف مؤشرات ثبات وصدق الاختبار وبخاصة الاختبارات الفرعية، كما أن نتائج التحليل العاملي التوكيدي والاستكشافي لم تقدم دعما للبناء العاملي الذي يقوم الاختبار على أساسه مثل ؛ أكدت معظم الدراسات على تميز الاختبار بخصائص سيكومترية جيدة، مثل دراسات كل من (Williams, 2003; Gadzella et al., 2005; Gadzella et al., 2006). هذا التناقض في النتائج بالإضافة إلى عدم وجود ترجمة عربية سابقة لهذه الصورة من الاختبار – وبالتالي فلا توجد دراسات عربية تدعم الخصائص السيكومترية للاختبار الحالي – قد دفعا الباحث الحالي إلى بحث الخصائص السيكومترية لهذا الاختبار بعد ترجمته إلى اللغة العربية.

وقد استفاد الباحث من تنوع الطرق المستخدمة لحساب الصدق والثبات في الدراسات السابقة أثناء إجرائه لبحثه الحالي واختياره لطرق حساب الصدق والثبات فيها.

فروض البحث:

الفرض الأول: يتشبع اختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" على عدد من العوامل لدي عينة الدراسة من طلاب الجامعة.

الفرض الثاني: تتوافر لاختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" درجة مقبولة من الثبات لدي عينة الدراسة من طلاب الجامعة.

الفرض الثالث: تتوافر لاختبار واطسون وجليسر للتفكير الناقد "النسخة المختصرة" درجة مقبولة من الصدق لدي عينة الدراسة من طلاب الجامعة.

إجراءات البحث:

أولا – العينة:

تكونت العينة النهائية للبحث من (٥٠٣) طالبا وطالبة من طلاب وطالبات الفرقة الرابعة شعبة اللغة الإنجليزية بكلية التربية جامعة المنوفية، و تراوحت أعمار العينة ما بين (١٩) عاما، و(٢٣) عاما بمتوسط قدره (٢١,٧٣)، وانحراف معياري قدره (١,٥٣).

ثانيا – أداة البحث:

استخدم الباحث في بحثه الحالي اختبارا واحدا هو اختبار واطسون – جليسر للتفكير ، الصورة المختصرة "S" الناقد (ترجمة الباحث الحالي)، وقد سبق التعريف بالاختبار وتطوره في الإطار النظري للدراسة الحالية، كما سيتم عرض أهم خصائصه السيكومترية في نتائج الدراسة.

ثالثا – الأساليب الإحصائية:

تم إجراء التحليل الإحصائى للبيانات عن طريق الحاسوب باستخدام الحزمة الإحصائية للعلوم الاجتماعية (SPSS,Version15)، وكانت الأساليب الإحصائية المستخدمة هى: التحليل العاملي – معامل الارتباط – معامل ثبات ألفا كرونباك – معامل ثبات جتمان– معامل ثبات سبيرمان–براون.

نتائج البحث:

نتائج الفرضين الأول والثاني (صدق الاختبار، والبناء العاملي له):

تم التأكد من صدق الاختبار في هذا البحث من خلال الطرق التالية:

١ – الصدق العاملي Factorial Validity:

استخدم الباحث الحالي التحليل العاملي Factor Analysis كوسيلة من وسائل التحقق من صدق الاختبار الحالي، عن طريق تحديد العوامل المكونة للاختبار وتحديد مدى اتفاقها أو اختلافها مع العوامل التي يفترض أن يقيسها هذا الاختبار. و استخدم الباحث التحليل العاملي الاستكشافي Exploratory factor Analysis بطريقة المكونات الأساسية Principal Components Method وتحديد عدد العوامل Number of Factors مع تدوير المحاور بطريقة الفاريماكس Varimax Method لدرجات الطلاب باستخدام بالبرنامج الإحصائي SPSS -Version 15 ، كما استخدم الباحث الاختبار الركامي (Scree Test) الذي وضعه Catell لتحديد عدد العوامل عن طريق دراسة الرسم البياني للجذور الكامنة، والإبقاء على العوامل التي تظهر في الجزء شديد الانحدار من المنحنى قبل أن يبدأ المنحنى في الاعتدال (رجاء أبو علام، ٢٠٠٣، ٢٧١)، كما تم قبول التشبعات الدالة للعوامل بناء على محك جيلفورد الذي يقبل التشبعات إذا كانت قيمتها أكبر من أو تساوي ٠,٣(بالانت، ٢٠٠٧، ١٩٨: ١٩٩).

و نتج عن التحليل العاملي تشبع جميع مفردات الاختبار بعد تدويرها على خمسة عوامل تفسر في مجملها ٨١,٠٣% من التباين الكلي، وبعض هذه المفردات تشبعت على أكثر من عامل، وقد نسبت هذه المفردات إلى العوامل التي تشبعت عليها بدرجة أعلى.

والجدولان التاليان رقم (١)، ورقم (٢) يوضحان قيم تشبع مفردات الاختبار وقيم الجذر الكامن ونسبة التباين المفسرة للعوامل بعد التدوير بطريقة Varimax.

جدول رقم (١) يبين قيم الجذور الكامنة للعوامل المستخرجة ونسب التباين الناتجة عنها

العوامل	الأول	الثاني	الثالث	الرابع	الخامس
الجذر الكامن	١٤,٥٩	٦,١٠	٤,٣١	٣,٩٢	٣,٤٩
نسبة التباين	٣٦,٤٨	١٥,٢٥	١٠,٧٨	٩,٧٩	٨,٧٢
النسبة التجمعية للتباين	٣٦,٤٨	٥١,٧٤	٦٢,٥١	٧٢,٣٠	٨١,٠٣

جدول رقم (٢) يبين تشبعات مفردات الاختبار على العوامل المستخرجة بعد التدوير

المفردات	العوامل			المفردات	العوامل	
	الأول	الثاني	الثالث		الرابع	الخامس
١	,٧٠٤			٢٥	,٧٥١	
٢	,٦١٩			٢٦	,٧٥١	
٣	,٧٩٤			٢٧	,٧٥١	
٤	,٨٦٥			٢٨	,٨١٠	
٥	,٨٤٣			٢٩	,٧٢٥	
٦	,٨٦٥			٣٠	,٨٢٦	
٧	,٥٩٧			٣١	,٥٩٢	
٨		,٦٦٤		٣٢		,٨٢٨
٩		,٤٤٧		٣٣		,٩٠٤
١٠		,٦١٨		٣٤		,٨٢٨
١١		,٦٣٠		٣٥		,٦٣٩
١٢		,٥٦٥		٣٦		,٨٢٨
١٣		,٥٥٣		٣٧		,٦٨٤
١٤		,٦٧٥		٣٨		,٨٠٥
١٥		,٧٢١		٣٩		,٦٥٣
١٦			,٩٢٧	٤٠		,٥٢٥
١٧			,٩٢٧			
١٨			,٧٤٥			
١٩			,٦٦٠			
٢٠			,٧٨٧			
٢١			,٧٩٠			
٢٢			,٩٢٧			
٢٣			,٩٢٧			
٢٤			,٧٤٥			

ويتضح من النتائج السابقة للتحليل العاملي ما يلي:

- تشبع مفردات الاختبار على خمسة عوامل فسرت مجتمعة معا ٨١,٠٣% من التباين الكلي، وهذه العوامل هي:

العامل الأول: وجذره الكامن (١٤,٥٩) وفسر حوالي (٣٦,٤٨%) من التباين الكلي للمصفوفة، وتشبع على هذا العامل سبع من مفردات الاختبار.

العامل الثاني: وجذره الكامن (٦,١٠) وفسر حوالي (١٥,٢٥%) من التباين الكلي للمصفوفة، وتشبع على هذا العامل ثماني من مفردات الاختبار.

العامل الثالث: وجذره الكامن (٤,٣١) ويفسر حوالي (١٠,٧٨%) من التباين الكلي للمصفوفة، وتشبع على هذا العامل تسع من مفردات الاختبار.

العامل الرابع: وجذره الكامن (٣,٩٢) ويفسر حوالي (٩,٧٩%) من التباين الكلي للمصفوفة، وتشبع على هذا العامل سبع من مفردات الاختبار.

العامل الخامس: وجذره الكامن (٣,٤٩) وفسر حوالي (٨,٧٢%) من التباين الكلي للمصفوفة، وتشبع على هذا العامل تسع من مفردات الاختبار.

- تتفق نتائج التحليل العاملي مع التكوين الفرضي للقدرة على التفكير الناقد عند واطسون وجليسر، حيث وضع الاختبار لقياس كل من الدرجة الكلية للتفكير الناقد، وخمس مهارات فرعية لهذه القدرة، وهو ما كشفت عنه نتائج التحليل العاملي السابقة بما يؤكد الصدق العاملي للاختبار الحالي.

- نتائج التحليل العاملي السابقة تؤكد صحة الفرضين الأول والثاني.

٢ – الصدق الظاهري Face Validity :

حيث قام الباحث بعرض النسخة الأصلية باللغة الإنجليزية والنسخة المترجمة إلى العربية على عدد من المحكمين من أعضاء هيئة التدريس بأقسام علم النفس بكليات التربية في جمهورية مصر العربية (خمسة محكمين)، وذلك للحكم على مدى وضوح صياغة العبارات، وسلامتها، وملاءمتها للعينة المصرية، وأيضا على مدى دقتها وصلاحيتها لقياس ما وضعت لقياسه. وقد أجاز المحكمون جميع العبارات مع بعض التعديلات البسيطة في الصياغة قام الباحث بمراعاتها في الصياغة النهائية للاختبار.

٣ – صدق الاتساق الداخلي Internal Consistency Validity :

صدق الاتساق الداخلي من أهم أنواع الصدق التي يمكن استخدامها للتحقق من صدق الأداة، ويرتبط هذا النوع من الصدق بالتحقق من الاتساق بين مفردات المقياس ومدى ارتباطها بالدرجة الكلية، مع مراعاة عدم التداخل بين المفردات المكونة لكل بعد، أو التداخل بين مفردات المقياس ككل (بندر العتيبي، ٢٠٠٤، ١٩: ٢٠). وللتحقق من صدق الاتساق الداخلي، قام الباحث بإجراء ثلاثة أنواع من الارتباطات، هي:

(أ) ارتباط كل عبارة من عبارات الاختبار بالدرجة الكلية للاختبار الفرعي: والجدول رقم (٣) يبين معاملات الارتباط الناتجة.

جدول رقم(٣) يبين معاملات الارتباط بين كل عبارة من عبارات الاختبار والدرجة الكلية للاختبار الفرعي

المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط
١	٠,٣٥٥**	١١	٠,٤٤٤**	٢١	٠,٤٠٩**	٣١	٠,١٨٨**
٢	٠,٦٤٣**	١٢	٠,٧١٩**	٢٢	٠,٣٧٦**	٣٢	٠,٢٩٨**
٣	٠,٣٧٤**	١٣	٠,٤٣٨**	٢٣	٠,٥٦٥**	٣٣	٠,٦٧٧**
٤	٠,٧٢٠**	١٤	٠,٣٥٦**	٢٤	٠,٧٨٧**	٣٤	٠,٧٦٢**
٥	٠,٤٩٥**	١٥	٠,٣٩٥**	٢٥	٠,٢٥٤**	٣٥	٠,٧٠٨**
٦	٠,٥٤٣**	١٦	٠,٥٦٧**	٢٦	٠,٣١٦**	٣٦	٠,٣٠٠**
٧	٠,٥٢٣**	١٧	٠,٥٤٦**	٢٧	٠,٢٢٣**	٣٧	٠,٣٨٨**
٨	٠,٦١١**	١٨	٠,٧٦٥**	٢٨	٠,١٦٦**	٣٨	٠,٤٨٩**
٩	٠,٤١٧**	١٩	٠,٤٨٨**	٢٩	٠,٣٨٧**	٣٩	٠,٤٢٢**
١٠	٠,٨٧٤**	٢٠	٠,٥٦٩**	٣٠	٠,٣٠٩**	٤٠	٠,٠٩٣*

** دال عند مستوى دلالة ٠,٠١ * دال عند مستوى دلالة ٠,٠١

ويتضح من الجدول السابق أن جميع معاملات الارتباط بين كل عبارة من عبارات الاختبار والدرجة الكلية للاختبار الفرعي موجبة ودالة إحصائيا عند مستوى دلالة ٠,٠١ ما عدا العبارة الأخيرة فهي دالة عند مستوى دلالة ٠,٠٥

(ب) ارتباط كل عبارة من عبارات الاختبار بالدرجة الكلية للاختبار ككل: والجدول رقم (٤) يبين معاملات الارتباط الناتجة.

جدول رقم(٤) يبين معاملات الارتباط بين كل عبارة من عبارات الاختبار الدرجة الكلية للاختبار ككل

المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط
١	٠,٧٥٥**	١١	٠,٥١١**	٢١	٠,٤١١**	٢١	٠,١٨٨**
٢	٠,٤٨٧**	١٢	٠,٦٨٩**	٢٢	٠,٣٩٨**	٣٢	٠,١٤٦**
٣	٠,٤٥٦**	١٣	٠,٤٦٧**	٢٣	٠,٥٦٦**	٣٣	٠,٢٠٩**
٤	٠,٦٥٤**	١٤	٠,٢٣٢**	٢٤	٠,٤٧٦**	٣٤	٠,٠٩٩*
٥	٠,٥٢٢**	١٥	٠,٢٧٧**	٢٥	٠,١٠٤*	٣٥	٠,٥٨٧**
٦	٠,٤٦٧**	١٦	٠,٣٤٥**	٢٦	٠,٢٧٨**	٣٦	٠,٢١١**
٧	٠,٥٦٤**	١٧	٠,٥٣٤**	٢٧	٠,٢٩٦**	٣٧	٠,٢٣٥**
٨	٠,٥٦٤**	١٨	٠,٣٦٧**	٢٨	٠,١٧٦**	٣٨	٠,٣٥٤**
٩	٠,٢٩٩**	١٩	٠,٢٨٨**	٢٩	٠,٢٦٥**	٣٩	٠,٢٤٣**
١٠	٠,٦٨٩**	٢٠	٠,٤٢٣**	٣٠	٠,٤٠٠**	٤٠	٠,٢٣٠**

** دال عند مستوى دلالة ٠,٠١ * دال عند مستوى دلالة ٠,٠١

ويتضح من الجدول السابق أن جميع معاملات الارتباط بين كل عبارة من عبارات الاختبار والدرجة الكلية للاختبار الفرعي موجبة ودالة إحصائيا عند مستوى دلالة ٠,٠١ ما عدا العبارتين ٢٥، و ٣٤ فهما دالتان عند مستوى دلالة ٠,٠٥.

(ج) ارتباط الدرجة الكلية لكل اختبار من الاختبارات الفرعية بالدرجة الكلية للاختبار ككل:

والجدول رقم (٥) يبين معاملات الارتباط الناتجة.

جدول رقم (٥) يبين ارتباط الدرجة الكلية لكل اختبار من الاختبارات الفرعية بالدرجة الكلية للاختبار ككل:

الأبعاد	معامل الارتباط بالدرجة الكلية
الاستنتاج	٠,٧١٢**
التعرف على الافتراضات	٠,٥٩٧**
الاستنباط	٠,٨٢٤**
التفسير	٠,٨٣٤**
تقويم الحجج	٠,٨٨٨**

** دال عند مستوى دلالة ٠,٠١

- يتضح من الجدول السابق أن جميع معاملات الارتباط بين كل من درجات الاختبارات الفرعية والدرجة الكلية للاختبار ككل موجبة ودالة إحصائيا عند مستوى دلالة ٠,٠١.

- كما يتضح من نتائج حساب الصدق بطرقه المختلفة (الصدق العاملي- صدق المحكمين- صدق الاتساق الداخلي) تمتع الاختبار الحالي بدرجة مناسبة من الصدق، وبذلك يتأكد صحة الفرض الثاني.

نتائج الفرض الثالث (ثبات الاختبار):

تم التأكد من ثبات الاختبار في هذا البحث من خلال الطرق التالية:

١ – طريقة إعادة تطبيق الاختبار Test Retest:

حيث تم تطبيق الاختبار على نفس العينة مرتين، وخلال فترة زمنية لا تقل عن أسبوعين، يُقاس بعد ذلك العلاقة بين درجات الطلاب على كلا التطبيقين. ولتطبيق هذا الأسلوب من قياس الثبات، أعاد الباحث تطبيق الاختبار على نفس عينة الدراسة بعد ثلاثة أسابيع من التطبيق الأول. والجدول رقم (٦) يبين ثبات الاختبار عند استخدام طريقة إعادة التطبيق

جدول رقم (٦) يبين معاملات ثبات الدرجة الكلية للاختبار ودرجات الاختبارات الفرعية باستخدام طريقة إعادة التطبيق

الاختبار	معامل الثبات
الاستدلال	٠,٨٩**
التعرف على المسلمات	٠,٧٣**
الاستنتاج	٠,٩٠**
التفسير	٠,٨٦**
تقويم الحجج	٠,٩١**
الدرجة الكلية	٠,٩٣**

**** دالة إحصائيا عند مستوى دلالة ٠,٠١**

ويتضح من الجدول السابق وجود معاملات ارتباط مرتفعة بين درجات التطبيقين الأول والثاني، وذلك في كل من الدرجة الكلية للاختبار، ودرجات الاختبارات الفرعية، حيث تراوحت معاملات الارتباط من ٠,٧٣ إلى ٠,٩٣، بما يدل على ثبات الاختبار.

٢ – طريقة الاتساق الداخلي[2] Internal Consistency:

عن طريق حساب معامل ألفا كرونباك Cronbach's Alpha ، وتعتمد هذه الطريقة على حساب قيمة معامل الثبات بطريق معرفة متوسط معاملات الارتباطات الداخلية بين الدرجة الكلية للاختبار ودرجات العبارات المكونة له. أو بين الدرجة الكلية للبعد ودرجات العبارات المكونة له. والجدول رقم (٧) يوضح معاملات ثبات ألفا كرونباك لكل من الدرجة الكلية للاختبار، ودرجات الاختبارات الفرعية.

جدول رقم (٧) يبين معاملات ثبات الدرجة الكلية للاختبار
ودرجات الاختبارات الفرعية باستخدام طريقة معامل ألفا كرونباك

الاختبار	معامل ثبات ألفا
الاستدلال	٠,٨٨
التعرف على المسلمات	٠,٦٥
الاستنتاج	٠,٩٦
التفسير	٠,٨٥
تقويم الحجج	٠,٩١
الدرجة الكلية	٠,٩٥

ويتضح من الجدول السابق أن قيم معامل ألفا كرونباك مرتفعة لكل من الدرجة الكلية للاختبار، ودرجات الاختبارات الفرعية عدا اختبار التعرف على المسلمات حيث بلغت قيمة ألفا كرونباك ٠,٦٥، وهي تدل على ثبات منخفض نسبيا لهذا الاختبار الفرعي، بالمقارنة بقيم ألفا لدرجة الاختبار الكلية، وباقي الاختبارت الفرعية. وهذه الدرجة وإن كانت أقل من درجة معامل ثبات الاختبار عن طريق إعادة التطبيق، ولكنه من الملاحظ رغم الدلالة المرتفعة لمعامل ثبات اختبار التعرف على المسلمات بطريقة إعادة التطبيق إلا أنه كان منخفضا بشكل ملحوظ عن باقى

[2] آثر الباحث استخدام طريقة الاتساق الداخلي في حساب الثبات رغم استخدامه نفس الطريقة في حساب الصدق وذلك لاختلاف ناتج كل منهما حيث يظهر في حساب الصدق على شكل معاملات ارتباط، أما في ثبات الاختبار فيظهر في شكل رقم واحد يعبرعن درجة ثبات الاختبار ككل، والجمع بين الطريقتين يؤكد الاتساق الداخلي للاختبار سواء أكان ذلك دليلا على صدق الاختبار أو على ثباته.

معاملات الثبات لبقية الاختبارات الفرعية والدرجة الكلية للاختبار بنفس الطريقة، وهذا ربما يدعو إلى إعادة النظر في مفردات هذا الاختبار الفرعي، ومدى تفاوتها في السهولة والصعوبة، أو عدم مناسبة بعض هذه المفردات لما وضعت لقياسه بما أدى إلى ضعف الاتساق الداخلي لها كما ظهر من قيمة ألفا كرونباك.

٣ – طريقة التجزئة النصفية Split Half :

وتقوم هذه الطريقة على تجزئة مفردات الاختبار إلى نصفين متكافئين، وحساب معاملات الارتباط بينهما، و قام الباحث بحساب هذه المعاملات بطريقتين هما: معامل جوتمان، ومعامل سبيرمان–براون، والجدول رقم (٨) يوضح معاملات ثبات جوتمان، وسبيرمان–براون لكل من الدرجة الكلية للاختبار، ودرجات الاختبارات الفرعية.

جدول رقم (٨) يبين معاملات ثبات الدرجة الكلية للاختبار ودرجات الاختبارات الفرعية باستخدام معاملات ثبات جوتمان، وسبيرمان– براون

الاختبار	جوتمان	سبيرمان– براون
الاستدلال	٠,٨٢	٠,٨٤
التعرف على المسلمات	٠,٧٥	٠,٧٥
الاستنتاج	٠,٩٥	٠,٩٦
التفسير	٠,٦٥	٠,٧٩
تقويم الحجج	٠,٨١	٠,٩٢
الدرجة الكلية	٠,٨٧	٠,٩٠

ويتضح من الجدول السابق أن قيم معاملات ثبات جوتمان، وسبيرمان–براون مرتفعة لكل من الدرجة الكلية للاختبار، ودرجات الاختبارات الفرعية، إلا أنه من الملاحظ أيضا أن معامل الثبات بطريقة التجزئة النصفية لاختبار التعرف على المسلمات أقل من نظيراتها في باقي الاختبارات الفرعية حيث بلغت قيمة معامل ثبات جوتمان، ٠,٧٥، وكذلك معامل سبيرمان–براون، وهي قيمة تدل على ثبات منخفض نسبيا لهذا الاختبار الفرعي، بالمقارنة بقيم معاملات ثبات جوتمان، وسبيرمان–براون لدرجة الاختبار الكلية، وباقي الاختبارت الفرعية. إلا ان درجات التجزئة النصفية بشكل عام تدل على ثبات الاختبار الحالي.

مناقشة النتائج:

هدف البحث الحالي إلى تحقيق ثلاثة أهداف رئيسة:

(١) الكشف عن البناء العاملي للترجمة العربية لاختبار واطسون – جليسر للتفكير الناقد، الصورة "S" لدى عينة من طلاب الجامعة، ومدى مطابقة هذا البناء للتكوين النظري الذي يقوم عليه الاختبار.

(٢) التحقق من صدق الترجمة العربية لاختبار واطسون – جليسر للتفكير الناقد، الصورة "S" لدى عينة من طلاب الجامعة.

(٣) التحقق من ثبات الترجمة العربية لاختبار واطسون – جليسر للتفكير الناقد، الصورة "S" لدى عينة من طلاب الجامعة.

وقد حقق البحث ما هدف إليه، فبالنسبة للهدف الأول كشفت الدراسة العاملية بطريقة المكونات الرئيسية عن تشبع مفردات الاختبار الأربعين جميعها على خمسة عوامل هي نفسها العوامل المفترضة في البناء النظري الذي يقوم عليه الاختبار وهو ما يدعم صدق البناء العاملي للاختبار الحالي. وبذلك تختلف النتائج التي توصل إليها البحث الحالي عن نتائج دراسة (Loo & Thorpe, 1999) التي لم تدعم البناء العاملي للاختبار.

وبالنسبة للهدف الثاني فقد أكدت نتائج البحث الحالي صدق الترجمة العربية لاختبار واطسون – جليسر للتفكير الناقد، الصورة "S" لدى عينة من طلاب الجامعة. و تحقق صدق الاختبار عن طريق ثلاث طرق من طرق حساب الصدق، وهي: (١) الصدق العاملي. (٢) صدق المحكمين. **(٣) صدق الاتساق الداخلي.**

وجميع الطرق السابقة أكدت تمتع الاختبار بدرجة مقبولة من الصدق، وهذه النتائج تتفق مع ما توصلت إليه دراسات كل من: (Williams, 2003; Gadzella et al., 2005; Gadzella et-al., 2006)

أما بالنسبة للهدف الثالث فقد أكدت نتائج البحث الحالي ثبات الترجمة العربية لاختبار واطسون – جليسر للتفكير الناقد، الصورة "S" لدى عينة من طلاب الجامعة. وقد تحقق ثبات الاختبار عن طريق ثلاث طرق من طرق حساب الثبات، وهي: (١) طريقة إعادة تطبيق الاختبار. (٢) طريقة الاتساق الداخلي (معامل ألفا كرونباك). (٣) طريقة التجزئة النصفية (معامل جتمان، ومعامل سبيرمان براون).

وجميع الطرق السابقة أكدت تمتع الاختبار بدرجة مقبولة من الثبات وبخاصة للدرجة الكلية للاختبار، وهذه النتائج تتفق مع ما توصلت إليه دراسات كل من: (Williams, 2003; Gadzella et al., 2005; Gadzella et- al., 2006)

مما سبق يمكن تأكيد أن الترجمة العربية الحالية لاختبار واطسون – جليسر للتفكير الناقد، الصورة "S" يتمتع بخصائص سيكومترية مرتفعة لدى عينة من طلاب الجامعة، بما يؤكد صلاحية هذا الاختبار للاستخدام في البيئة المصرية.

بحوث مقترحة:

على ضوء متغيرات البحث، وما تم التوصل إليه من نتائج يقدم الباحث بعض التوصيات وعدداً من البحوث المقترحة فيما يلي:

١ – الاهتمام بتعليم التفكير الناقد في مدارسنا، لأنه يعد أساسا للمواطنة الصالحة، به يستطيع الفرد التمييز بين الصواب والخطأ، وأن يختار ما يصلح له ولمجتمعه.

٢ – الاهتمام باختبارات التفكير الناقد وتقنينها، لأن ما لا يمكن قياسه لا يمكن تحقيقه. فتنمية التفكير الناقد والاهتمام باختبارات التفكير الناقد وجهان لعملة واحدة لا يمكن فصل أحدهما عن الآخر.

٣ – عمل برامج لتنمية التفكير الناقد ومهاراته لدى عينات مختلفة.

٤ – إجراء البحث الحالي على عينات مختلفة في البيئة المصرية، وخاصة على عينات من مراحل ما قبل الجامعة، لبيان مدى صلاحية هذا الاختبار لقياس التفكير الناقد لدى هذه العينات.

المراجـــع

بلانت، جولي (٢٠٠٧). *التحليل الإحصائى باستخدام برامج SPSS* (ترجمة خالد العامرى)، القاهرة، دار الفاروق للنشر والتوزيع، ط الأولى.

بندر بن ناصر العتيبي (٢٠٠٤). الخصائص السيكومترية لصورة سعودية من مقياس فينلاند للسلوك التكيفي (دراسة استطلاعية). *مجلة أكاديمية التربية الخاصة*، العدد الخامس، الرياض.

جابر عبد الحميد، ويحيى هندام (بدون تاريخ)*كراسة تعليمات اختبار التفكير الناقد*. دار النهضة العربية، القاهرة .

رجاء أبو علام (٢٠٠٣). التحليل الإحصائي للبيانات باستخدام برنامج spss . دار النشر للجامعات، القاهرة.

سيد عثمان، وفؤاد أبو حطب (١٩٨٧). *التفكير دارسات نفسية*، ط٢، مكتبة الأنجلو المصرية، القاهرة.

عبد العال عجوة ، عادل البنا (٢٠٠٠). *اختبار كاليفورنيا لمهارات التفكير الناقد*. المكتبة المصرية للنشر والتوزيع، الإسكندرية .

عزيزة السيد (١٩٩٥). *التفكير الناقد، دراسة في علم النفس المعرفي*. دار المعرفة الجامعية، الإسكندرية.

عماد الدين محمد السكري (٢٠٠٤). استراتيجيات معالجة المعلومات ومهارات التفكير الناقد. رسالة دكتوراة غير منشورة. كلية التربية، جامعة المنوفية.

فاروق السيد عثمان (١٩٩٣). *التفكير الناقد وعلاقته بتخفيض مستوى التعصب لدى عينة من طلاب الجامعة*. مجلة علم النفس ، السنة السابعة ، العدد ٢٧ ، الهيئة المصرية العامة للكتاب ، القاهرة.

Adams, B.L. (1999). Nursing Education for Critical Thinking: An Integrative Review. *Journal of Nursing Education*, Vol. 38, No.3, 111-119.

Ayala, A.E. (1999). Critical Thinking Skills and knowledge of English of Baccalaureate Nursing Students in Six Schools in Puerto Rico Accredited by the National League for Nursing Accrediting Commission. A Doctor of Philosophy Thesis Submitted to the School of Nursing, Wideener University.

Berger, A. (1985). Watson-Glaser Critical Thinking Appraisal. In *The Ninth*

Mental Measurement Yearbook. Edited by James, V. & Mitchell, Jr. Retrieved July 2, 2009, from EBSCO Host Mental Measurements Yearbook database.

Crites, J.O. (1965). Test Reviews. *Journal of Counseling Psychology*, Vol 12, No. 3, 328- 330.

Fisher, A. (2001). *Critical Thinking: An Introduction*. Cambridge University Press. New York.

Gadzella, B.M. & Baloglu, M. (2003). Psychometric Properties of Watson-Glaser Critical Thinking Appraisal for A Sample of Education Majors. *Psychological Reports*, 92, 1249-1254.

Gadzella, B.M., Stacks, J., Stephens R.C., and G. Masten W.G. (2005). *Watson-Glaser Critical Thinking Appraisal, Form-S For Education Majors*. Journal of Instructional Psychology, Vol. 32, No. 1, 9-12.

Gadzella, B., Hogan, L., Masten, W., Stacks, J., Stephens, R., & Zascavage, V. (2006). Reliability and validity of the Watson-Glaser critical thinking appraisal-form-s for different academic groups. *Journal of Instructional Psychology*, 33(2), 141- 143.

Geisinger K.F. (1998). Watson-Glaser Critical Thinking Appraisal, Form S. In *The thirteenth Mental Measurement Yearbook*. Edited by James C. Impara and Barbara S. Plake. Retrieved July 2, 2009, from EBSCO Host Mental Measurements Yearbook database.

Gordon, B.R. (1994). Edward Maynard Glaser (1911-1993). *American Psychologist*, Vol. 49, No. 8, 755.

Helmstadter, G.C. (1985). Watson-Glaser Critical Thinking Appraisal. In *The Ninth Mental Measurement Yearbook*. Edited by James, V. & Mitchell, Jr. Retrieved July 2, 2009, from EBSCO Host Mental Measurements Yearbook database.

Hicks-Moore, S.L. & Pastirik, P.J. (2006). Evaluating Critical Thinking in Clinical Concept Maps: A Pilot Study. *International Journal of Nursing Education Scholarship*, Vol. 3, No. 1, Article 27, 1-15.

Loo, R. & Thorpe, K. (1999). A Psychometric Investigation of Scores on the Watson-Glaser Critical Thinking Appraisal New Form S. *Educational and Psychological Measurement*, Vol. 59 No. 6, 995 -1003.

McMillan, J.H. (1987). Enhancing college student critical thinking: a review of studies. *Research of Higher Education*, 26, 3-29.

Moore M.L. (1976). Effects of Value Clarification on Dogmatism, Critical

Thinking, and Self-Actualization. Unpublished Doctoral Dissertation. Arizona State University.

Oliver, H. & Utermohlen, R. (1995): An Innovative Teaching Strategy: Using Critical Thinking to Give Students A Guide to the , *Eric Document Reproduction Services,* NO. 389702.

Watson, G. & Glaser, E.M. (2008). *Watson-Glaser Critical Thinking Appraisal: Short* Form Manual. Pearson Education, Inc., USA.

Williams. R. L. (2003) Critical thinking as a predictor and outcome measure in a large undergraduate educational psychology course. ERIC. TMO35016.1-19.

Some Psychometric Properties of WGCTA – Form S Among a Sample of University Students

By

Dr. Emad Eldin Mohamed Elsokkary

Lecturer of Educational Psychology
Faculty of Education – Menoufiya University

Abstract:

The current study aimed to investigate some psychometric properties of Watson – Glaser Critical Thinking Appraisal – form S. The final sample of the study consisted of (503) students (males & females) from the Faculty of Education, Menoufiya University, 4th graders enrolled in English Department, between (19-23) years of age, with mean age (21.73), and standard deviation (1.53).

The results revealed that the WGCTA form "S" is a valid and reliable test. Validity was determined by three methods in the current study: Factorial Validity, face Validity, and Internal Consistency Validity. Reliability was determined by three methods also: Internal Consistency Reliability, Split-Half Reliability and Test Retest Reliability. The Exploratory Factor Analysis (Principal Components) confirmed the factorial validity of the test. All the test items were loaded on five factors. The findings of the study were discussed in the light of the theoretical framework and previous studies

In total, the results proved that the WGCTA-S has acceptable psychometric properties in the Egyptian environment.