

أثر تعليم استراتيجيات التخمين في اكتساب الطلبة بالمستوى الجامعي لمهارة حكمة الاختبار «دراسة تجريبية»

الدكتور أحمد سليمان عودة*
كلية التربية - جامعة اليرموك

الملخص

هدفت هذه الدراسة إلى تعرف مدى اكتساب الطالب الجامعي لمهارة حكمة الاختبار، من خلال خبرته بالاختبارات، خلال المراحل التعليمية، ومدى إمكانية رفع هذه المستوى بتعليمه استراتيجيات حكمة الاختبار. ولتحقيق هذا الهدف تم اختيار مجموعتين (ضابطة، تجريبية)، وأعطيت المجموعة التجريبية محاضرتين (ساعة لكل محاضرة) غطت استراتيجيات التخمين، حسب تصنيف فردركسون. وبعد أسبوع أعطيت المجموعتان إختباراً من جزئين، يقيس جزء منه مهارة حكمة الاختبار، (وتكون من ٢٤ فقرة، غطت استراتيجيات التخمين)، ويقاس الجزء الآخر التحصيل الأكاديمي، وتكون من ٤٠ فقرة .

أشارت النتائج إلى أن متوسط علامات المجموعة الضابطة على اختبار حكمة الاختبار أعلى من التوقع، مما يدل على أن الطالب الجامعي يتعلم حكمة الاختبار من تكرار موقف الاختبار، كما أشارت نتائج تحليل التباين المصاحب إلى فرق دال إحصائياً بين المجموعتين، لصالح المجموعة التجريبية، مما يشير إلى أنه لا يمكن الاعتماد على خبرة الطالب لتعلم حكمة الاختبار، وأنه لا بد من إكساب الطلبة هذه المهارة بالتعليم، ما دام في الإختبارات المدرسية جوانب ضعف تسمح بممارستها. ولذلك توصي هذه الدراسة بإجراء دراسات مسحية؛ لتعرف جوانب الضعف في الاختبارات المدرسية ذات العلاقة بحكمة الاختبار .

خلفية الدراسة :

يتعرض الطالب خلال المراحل التعليمية المختلفة إلى عدد كبير من الإختبارات بعد أشكال من الفقرات، ويفترض المدرس أن الفروق بين الطلبة في نتائجهم في الإختبارات تعزى إلى قدراتهم في مجال المادة الدراسية، إلا أنه من السهل أن يكتشف أي مدرس من

* أستاذ مشارك، دكتوراه في البحث والتقييم، جامعة ولاية أيوا، إيمز الولايات المتحدة الأمريكية،

خلال التغذية الراجعة للإختبارات أن حكمة الطالب في الإجابة (حكمة الإجابة) أو حكمة الإختبار Testwiseness (أنظر التعريف الإجرائي) يمكن أن يكون عاملاً من العوامل التي تساهم في الحيلولة دون تحقيق هذا الافتراض؛ فبعض الطلبة يتدمرون من عدم حصولهم على علامات تكافئ فهمهم واستعدادهم، بينما يبدي آخرون ارتياحهم للعلامة (الدرجة) مقارنة بتحضيرهم واستعدادهم، خاصة في الاختبارات من نوع الاختيار من متعدد. وقد أشار دولي، ووليمز (Dolly and Williams, 1986) إلى أهمية هذا العامل كمصدر تباين في العلامات، بمعنى أن جزءاً من التباين الحقيقي في علامات الطلبة يعزى إلى حكمة الإختبار، كما أن الإرتفاع النسبي في علامات بعض الطلبة يمكن أن يعزى إلى مهارة حكمة الإختبار، وبالتالي أن بعض القرارات التربوية تبنى على معلومات غير صادقة .

ويعرف ميللمان ورفيقاه (Millman, et. al., 1965) حكمة الإختبار بأنها قدرة المفحوص على استغلال خصائص الاختبار، ونوع الفقرات، وإدارة (إجراء) الاختبار في رفع علامته، دون معرفته لمحتوى الإختبار. وبالمعنى نفسه يعرف سارناكي (Sarnacki, 1979) حكمة الاختبار بأنها قدرة معرفية، أو مجموعة مهارات، تمكن المفحوص من رفع علامته، بصرف النظر عن محتوى فقرات الاختبار. وقد صنف ميللمان ورفيقاه استراتيجيات حكمة الإختبار إلى عناصر لا تعتمد على بازي (مطور) الاختبار، والغرض منه، وهي تلك المتعلقة باستراتيجية استغلال الوقت، وإستراتيجية تجنب الأخطاء، وإستراتيجية التخمين، واستراتيجية التفكير الاستدلالي. وعناصر أخرى تعتمد على بازي الإختبار، والغرض منه، وهي تلك المتعلقة باستراتيجية معرفة القصد، واستراتيجية الإستفادة من المفاتيح .

وقد أعاد فردركسون (Fredrickson, 1984) تصنيف هذه العناصر في أربع استراتيجيات، تحت عنوان استراتيجيات تعليم حكمة الإختبار، وهي : استراتيجية تهيئة الطلبة للاختبار، واستراتيجية إستخدام الوقت بفعالية، وإستراتيجية تجنب الخطأ، واستراتيجية التخمين. وتتضمن الإستراتيجية الأخيرة التقنيات (العناصر) التالية : أسلوب صياغة الإجابة مقارنة بالموهات، ملاحظة الإتساق النحوي أو عدمه بين المتن والبدائل (كلمات مفتاحية)، حذف البدائل التي تبدو غير مقبولة منطقياً والإختيار من البدائل الباقية، وجود بدائل تتضمن بدائل أخرى في الفقرة نفسها، وجود فقرات كاشفة عن الإجابة في فقرات أخرى، وجود كلمات تفيد التأكيد أو الإحتمال .

وحول أهمية تدريس حكمة الاختبار يقول إبل (Ebel, 1965, P.3) إن التربويين المثاليين (Idealistic) يشكّون في أهمية صرف الوقت في تدريب الطلبة على مهارة حكمة الاختبار ويعتقدون بأهمية صرف الوقت في تعليم الطلبة محتوى المادة الدراسية، ولكن التربويين الواقعيين (Realistic) يحثون على تطوير هذه المهارة؛ لأن واقع الاختبارات يسمح بممارستها، وأن أخطاء القياس عند الطلبة الذين تتوفر لديهم هذه المهارة أقل منها عند من لا يملكونها. ومن المتحمسين الذين يحثون على تدريس حكمة الاختبار ريتشاردسون (Richardson, 1984) فهو يعدّ الاختبارات مواقف مثالية للتعامل والتفاعل بأقصى حكمة، ويحث على تدريس هذه الحكمة كجزء من منهاج الدراسات الاجتماعية؛ لأنها توفر مهارة في التواصل الاجتماعي، وهذا يعني إمكانية تعميم مهارة حكمة الاختبار على مجال أوسع من مجال الاختبارات، بمعنى أن الحكمة التي يتعلمها الطالب من موقف الاختبار يمكن أن تنعكس على الحكمة في التصرف في مواقف الحياة المختلفة أو ما يشار إليه بانتقال أثر التعلم.

أجري العديد من الدراسات حول مهارة حكمة الاختبار وإمكانية تعليمها، وقد تعددت أهدافها وإجراءاتها، واتبعت في معظمها أسلوب البحث التجريبي المتمثل في مقارنة أداء مجموعة تجريبية بمجموعة ضابطة، كما تنوعت استراتيجيات حكمة الاختبار، واختلفت فترة أو طريقة تدريسها للمجموعات التجريبية، واختلفت العينات من حيث العمر أو المستوى أو المعرفة. ولذلك، فإن فهم نتائج الدراسات السابقة لا يمكن أن يتم دون معرفة الإجراءات المتبعة فيها.

يتركز اهتمام الدراسات المتعلقة بالقياس الصفّي بالكشف عن جوانب الضعف في صياغة الفقرات التي تنهى الفرصة لممارسة حكمة الاختبار (See McMorris., et. al., 1977, Mentzer, 1982) ويرافق ذلك الإهتمام دراسة أثر تعليم استراتيجيات حكمة الاختبار في أداء الطلبة. فقد هدفت إحدى التجارب في دراسة سمث (Smith, 1982) إلى معرفة أثر تعليم حكمة الاختبار بإستراتيجيات التخمين التقاربي Guessing strategies على علامات طلبة المرحلة الثانوية في اختبار الاستعداد الدراسي المقنن (SAT) بجزئيه اللفظي والرياضي، وقد أشارت النتائج إلى فرق ذي دلالة إحصائية لصالح المجموعة التجريبية على الجزء اللفظي، بينما كان الفرق على الجزء الرياضي غير دال إحصائياً، وخسرت النتيجة بعدم وجود مفاتيح تسمح بممارسة استراتيجيات التخمين؛ هذا وسيوضح المقصود بالمفاتيح واستراتيجيات التخمين في التعريفات الإجرائية.

استخدمت دراسة دولي ووليمز (Dolly and Williams, 1986) استراتيجيات التخمين على مجموعتين (تجريبية وضابطة) من المرحلة الجامعية الأولى، حيث أعطيت المجموعة التجريبية محاضرة ومناقشة مع نشرات وتدريبات لمدة ساعة، غطت استراتيجيات التخمين، وبعد أسبوع من التجربة خضعت المجموعتان لاختبار حكمة الاختبار من تصميم الباحثين ويتكون من ٢٤ فقرة، يحتوي نصفها على مفاتيح حكمة الإجابة، ونصفها الآخر متحرر من المفاتيح. وقد أشارت النتائج إلى فرق دال إحصائيا لصالح المجموعة التجريبية على الجزء المتضمن مفاتيح حكمة الاختبار. وتتفق هذه النتيجة مع نتيجة دراسة دياموند وايفنز (Diamond and Evans, 1972)، التي أشارت إلى أن أثر حكمة الاختبار يظهر في الفقرات التي تم التدريب على الإستراتيجيات والمفاتيح الواردة فيها.

أظهرت دراسات أخرى مزيداً من التنوع في الإجراءات، فقد بحثت دراسة كلستين ورفيقاه (Kalechstien, et. al, 1981) في أثر تعليم الطلبة السود استراتيجيات حكمة الاختبار الواردة في تصنيف ميللمان ورفيقه في علاماتهم على اختبار ستانفورد في القراءة، وكانت عينة الدراسة عن طلبة الصف الثاني الابتدائي. وقد أجرى دريسباخ وكوج (Dreishbach and keogh, 1982) دراسة مماثلة على الطلبة الأميركيين، من أصل مكسيكي، في رياض الأطفال بمتغير تابع هو اختبار الإستعداد المدرسي. أما دراسة رايتز وايدول - ميستس (Ritter and Idol - Meastas, 1986) فقد هدفت إلى معرفة أثر مهارات حكمة الاختبار حسب الإستراتيجيات الواردة في تصنيف ميللمان ورفيقاه في أداء الطلبة في الاختبارات التحصيلية من إعداد المدرس، وكانت العينة من طلبة الصف السادس. وخلافاً لما ورد في الدراسات السابقة فقد انفرد ايفانز (Evans, 1984) في طريقة تدريبيه الطلبة على استراتيجيات حكمة الاختبار، فقد بحثت في إمكانية تعلم حكمة الاختبار من خلال الفة المفحوص باني الاختبار. كانت العينة من المهنيين الذين يعملون في وحدات للإسعاف الفوري، وقد وجد أن الفرق بين المتوسطين دال إحصائيا لصالح المجموعة التجريبية على الفقرات التي تتضمن المفاتيح الواردة في التجربة. كما انفرد سمث (١٩٨٢) في التجربة الثانية من دراسته في الكشف عن إمكانية استخدام الطالب الجامعي لاستراتيجيات التخمين في رفع علامته فوق علامة التخمين العشوائي.

وقد أجرى سكرجز ورفيقاه (Scruggs et.al., 1986) دراسة تحليلية للدراسات التي بحثت في تعليم مهارة حكمة الاختبار، ووجدوا أنها تركز على الاختبارات المقننة في المراحل التي تسبق المرحلة الجامعية؛ كما وجدوا أن أثر تعلم حكمة الاختبار في العلامة

التي تعكس تحصيله أو قدرته يختلف باختلاف خصائص المجموعة، وطول فترة التدريب. وجاء في معرض مناقشة نتائج الدراسة التوصية بأجراء المزيد من الدراسات حول هذا الموضوع قبل أن نقبل بأهمية تعليم حكمة الإختبار، سواء من حيث تأثيرها في العلامات والقرارات المترتبة عليها، أو في علاقتها بقلق الإختبار، ومفهوم الطالب لذاته، واتجاهاته نحو الإمتحان بشكل عام .

يتضح من الدراسات السابقة أهمية موضوع حكمة الإختبار كمجال للبحث، وذلك لأثرها في القرارات التربوية، التي يتم إتخاذها في ضوء نتائج الإختبارات، ولعلاقتها ببعض الخصائص التي تؤثر في هذه النتائج، مثل قلق الإختبار والاتجاهات نحو الإختبارات، وإمكانية تعدي أثرها إلى تطوير مستوى الذكاء المكتسب، لكون موقف الإختبار من المواقف التي يمارس فيها الطالب درجة عالية من عمق التفكير المستقل المنظم، والتي ما زالت كغيرها من الموضوعات في هذا المجال مجرد فروض تتطلب المزيد من البحوث لفحصها، وإذا كان الطالب بالمستوى الجامعي قد أمضى عشرات بل مئات الساعات في الإجابة على الإختبارات خلال المراحل التعليمية والمباحث المتنوعة، فهل نتوقع أن يكون قد تعلم حكمة الإختبار بالخبرة؟ وهل يمكن رفع مستواها عند الطالب الجامعي إلى مستوى أعلى مما توفره خبرته خلال السنوات الدراسية؟ إن الإجابة على هذين السؤالين تحدد الإطار الذي يوجه البحث في الآثار السابقة لحكمة الاختبار، وتأتي هذه الدراسة لتساهم في تحديد هذا الإطار .

مشكلة الدراسة وأهميتها

قد لا يخلو كتاب في القياس والتقويم، خاصة تلك التي تعتمد مراجع أساسية أو كتباً مقررة لمسافات القياس والتقويم في برامج تأهيل المعلمين في الجامعات والكليات المتوسطة من إرشادات تتعلق بصياغة فقرات الإختبارات التحصيلية من نوع الإختبار من متعدد، ومن ضمن هذه الإرشادات تلك المتعلقة بالمنبهات أو مفاتيح الإجابة، التي توفر الفرصة لممارسة استراتيجيات التخمين المعنية في هذه الدراسة، بمعنى أن توظيف هذه الإستراتيجيات يعتمد على جوانب ومظاهر التحيز في صياغة الفقرات .

تتميز هذه الدراسة في أنها تحاول تعرّف مدى إمكانية الاعتماد على الخبرة الزمنية في اكتساب الطالب مهارة حكمة الاختبار، ولذلك كان الاختبار للطلاب الجامعي الذي يتوقع أن يكون قد وصل إلى درجة الاتقان، التي تساوي ٨٠ من مئة حسب اقتراح الباحث، وأن المجال مفتوح لاقتراح أي معيار آخر يمكن أن يكون مقبولاً منطقياً في

دراسات لاحقة. كما تتميز هذه الدراسة في معالجة أثر استراتيجية التخمين في اختبار تحصيلي غير مقنن (أي من إعداد المدرس)، وذلك لأهمية الدور الذي تلعبه هذه الاختبارات في العملية التعليمية التعلمية. وبما يستحق الإشارة إليه هنا أن الدراسات السابقة أجريت في مجتمع يخضع فيه الطلبة عند وصولهم إلى الجامعة إلى عدد أكبر مما يخضع له الطلبة في المدارس والجامعات العربية، وبالتالي يتوقع أن يكون الطلبة في ذلك المجتمع أكثر اهتماماً في توظيف الاستراتيجيات الممكنة للوصول إلى أعلى علامة ممكنة كونها اختبارات من نوع أقصى أداء. ولذلك جاءت هذه الدراسة كأول دراسة تبحث هذا الموضوع في البيئة العربية، خاصة أن المزايا التي تتمتع بها أسئلة الاختيار من متعدد، تشجع المدرسين على استخدامها في المراحل التعليمية المختلفة. وفي ضوء هذه المعطيات، بالإضافة إلى ما جاء في الفقرة التي ربطت خليقة الدراسة بمشكلتها، فإن الهدف من هذه الدراسة يتلخص في فحص الفرضيتين التاليتين عند مستوى الدلالة (٠,٠٥) :

الفرضية الأولى : لا يختلف متوسط علامات الطلبة على مقياس حكمة الاختبار عن العلامة المتوقعة طبقاً للتخمين العشوائي .

الفرضية الثانية : لا يوجد فرق ذو دلالة إحصائية بين متوسط علامات الطلبة، الذين تعلموا حكمة الاختبار ومتوسط علامات الطلبة الذين لم يتعلموا حكمة الاختبار في مقياس حكمة الاختبار .

التعريفات الإجرائية

ورد في ثنايا البحث عدة مصطلحات متداخلة ارتبطت بمفهوم أساسي هو مفهوم «حكمة الاختبار» ويقصد بهذا المفهوم هنا أنه النشاط العقلي الذي يقوم به الطالب لرفع علامته في الاختبار أكثر من العلامة التي تعكس معرفته بمحتوى الاختبار، سواء من حيث استغلال الوقت، أو الاستفادة من خصائص الاختبار، ونوع الفقرات فيه، أو من ظروف التطبيق للاختبار، أو التهيئة النفسية والجسمية والعقلية. وقد حدد ميللمان ورفيقاه مجموعة عناصر ضمن هذه المجالات، وصنفت في عدة استراتيجيات، ومنها استراتيجية التخمين. وبما أن هذا النشاط يبرز في أثناء الإجابة على أسئلة أو فقرات الاختبار، فيمكن أن تكون «حكمة الإجابة» رديفاً «لحكمة الاختبار»، ويعبر عن «مهارة حكمة الاختبار» بمدى توظيف هذه الاستراتيجيات في الوصول إلى الإجابة الصحيحة على أكبر عدد ممكن من الفقرات .

وفي هذه الدراسة إن الاستراتيجية المعنية بالبحث هي «استراتيجية التخمين» التي تستخدم فيها المؤشرات الدالة على الإجابة سواء كانت في الفقرة نفسها مثل : طول البديل

الصحيح، ووجود كلمة في متن (أصل) الفقرة، وأحد البدائل، ووجود بديل يتضمن بديلاً آخر؛ أو في فقرة أخرى في الاختبار نفسه، التي تسمى بالفقرة الكاشفة؛ أو ملاحظة الخصائص التي تربط البدائل مع بعضها، والذي يسمى «بالتخمين التقاري». وتقاس حكمة الاختبار الخاصة باستراتيجيات التخمين في هذه الدراسة بمقدار الزيادة في علامة الطالب فوق «علامة التوقع» (وهي العلامة التي يتوقع حصوله عليها بالتخمين العشوائي) على اختبار تحتوي كل فقرة من فقراته على مفاتيح توفر استخدام استراتيجية أو أكثر من استراتيجيات التخمين المشار إليها في «موقف اختبار تحصيلي حقيقي»، ويقصد بذلك موقف اختبار يشعر فيه الطالب بأهمية الحصول على أعلى علامة ممكنة، وهو ما وفرته هذه الدراسة من خلال وضع فقرات تتضمن المفاتيح (كمقياس فرعي) ضمن فقرات اختبار تحصيلي لمساق جامعي (انظر أداة الدراسة).

الطريقة والإجراءات

عينة الدراسة

اشترك في الدراسة (٧٤) طالباً في شعبتين من شعب مادة القياس والتقويم في دائرة التربية بجامعة اليرموك، (منهم ٣٦ طالباً في المجموعة التجريبية، ٣٨ طالباً في المجموعة الضابطة)، ومن محددات هذه الدراسة عدم إمكانية التوزيع العشوائي لأفراد العينة في المجموعتين، (الضابطة، والتجريبية). وكان البديل للضبط العشوائي في هذه الدراسة هو الضبط الإحصائي، فقد استخدم تحصيل الطالب في المساق (معبراً عنه بالعلامات على الاختبار الأول، حسب نظام العلامات في جامعة اليرموك)، كمتغير مصاحب. ولعله من الصعب تصور علاقة نتائج هذه الدراسة بنتائج الدراسات السابقة إلا بمعرفة أوجه الشبه، والاختلاف في الطريقة والإجراءات، ومنها المتعلقة في العينة. فقد استخدم بعضها أسلوب التقسيم العشوائي، ومنها : دراسة دولي ووليمز (١٩٨٦)، حيث تكونت العينة من ٥٤ طالباً (٢٩ في المجموعة التجريبية، ٢٥ طالباً في المجموعة الضابطة)، وبعضها أسلوب التقسيم بالأزواج المتكافئة، ومنها دراسة رايتز وإيدول - ميسنس (١٩٨٦)، حيث تكونت العينة من ٥٦ طالباً، واتباع أغلبها أسلوب الضبط الإحصائي، ومنها : دراسة سمث (التجربة الثانية) (١٩٨٢)، وكان المتغير المصاحب لعلامات الطلبة في اختبار الاستعداد الدراسي التمهيدي المقنن، Primary Scholastic Aptitude Test حيث كان عدد الأفراد ٣٧ طالباً (٢٠ في التجريبية، ١٧ في الضابطة). ودراسة كلستين ورفيقه (١٩٨١)، حيث كان عدد أفراد الدراسة ٤٣ طالباً (٢٣ في التجريبية، ٢٠ في الضابطة)،

والتغير المصاحب نتائج الطلبة في اختبار ستانفورد في القراءة. ولذلك تشابه الدراسة الحالية مع بعض الدراسات السابقة في أسلوب اختيار العينة والتصميم الإحصائي .

أداة الدراسة

تكونت أداة الدراسة من ٦٤ فقرة في الاختبار الأول لمادة القياس والتقييم لطلبة البكالوريوس والدبلوم، منها ٢٤ فقرة (الملحق رقم ١) تقيس حكمة الاختبار باستخدام عناصر استراتيجية التخمين، أو ما يشار إليها باستراتيجيات التخمين في تصنيف فردركسون، وتتضمن جميع المفاتيح (المنبهات) المتعلقة باستراتيجيات التخمين، وتتلخص في : وجود الكلمة نفسها أو كلمة بالمعنى نفسه في المتن، وأحد البدائل، ووجود بدائل غير منطقية، لا تنسجم مع المتن، والترابط النحوي بين المتن والبدائل، وطول البديل الذي يشكل الإجابة الصحيحة مقارناً بالموهات، ووجود بدائل تتضمن بدائل أخرى في الفقرة نفسها، أو وجود فقرات كاشفة عن فقرات أخرى، ووجود كلمات تفيد التأكيد أو الإحتمال. وقد يرد في الفقرة الواحدة أو تستخدم فيها أكثر من استراتيجية، وبالتالي إن قدرة الطالب في الوصول إلى الإجابة تعتمد على ترجيحه للاستراتيجية الأقوى (أنظر الهامش في الملحق ٢) .

وقد كانت هذه الفقرات من مادة القياس والتقييم التي لا تتوفر أي معلومات لدى الطلبة عنها، فبعض هذه الفقرات من مادة غير واردة في خطة المساق، وهي الفقرات المتعلقة بأدوات الملاحظة في القياس والتقييم، ونظم العلامات المدرسية وتركيبها؛ وبعضها الآخر يغطي وحدة مقدمة في القياس والتقييم، وهي الخصائص السيكمترية للاختبار، حيث لا يعطى الطلبة في هذا المساق إلا مقدمة بسيطة عنها. ولذلك، فإن الفقرات صادقة ظاهرياً في انتمائها لمادة القياس والتقييم، خاصة أن الوحدة الأولى من هذا المساق مخصصة لتوضيح المفاهيم والمصطلحات الأساسية في القياس والتقييم الصفي. وهذا يوفر موقفاً طبيعياً في أثناء الاختبار، كونها اختباراً تحصيلياً يتطلب تقديم الطالب أفضل ما عنده، والحصول على أعلى علامة ممكنة؛ كما أن توظيف مهارة حكمة الاختبار تتطلب إثارة الدافعية والاهتمام، وهذا ما يوفره موقف الاختبار الذي تساهم نتائجه في اتخاذ قرارات أكاديمية بحق الطالب في المساق. ومن حيث ترتيب الفقرات في كراسة الاختبار، فقد جاءت مرتبة عشوائياً، كونها جميعاً من نوع الاختيار من متعدد (أربعة بدائل لكل منها). أما معامل ثبات مقياس حكمة الاختبار، فكان ٤٣,٠ محسوباً بالطريقة النصفية على عينة من ٤٦ طالباً في شعبة من المساق نفسه لم تتعلم استراتيجيات حكمة

الاختبار، ولكن كان لديها معرفة أعلى من التخمين، بدلالة إحصائية (٠,٠٥). وقد تم حساب معامل الثبات على المجموعة التي تعلمت حكمة الاختبار حيث كان ٠,٦٤، وهذه نتيجة مبكرة تشير إلى تحسن خصائص الاختبار باكتساب الطلبة لمهارة حكمة الاختبار، حيث يتوقع أن لا يختلف معامل الثبات عن الصفر، عندما لا يزيد متوسط الأداء عن التوقع.

استخدمت علامات الطلبة على فقرات الجزء الخاص بتحصيل الطالب في المساق كمتغير مصاحب (ضابط) Covariate، حيث كان معامل ثبات هذا الاختبار (٨٥،)، في المجموعة الضابطة والتجريبية، كما أن معامل الارتباط بين العلامات على اختبار التحصيل ومقياس حكمة الاختبار (٦٣،، ٧٦)، في المجموعة الضابطة والتجريبية، مما يسوغ استخدام نتائج اختبار التحصيل كمتغير مصاحب، ولذلك تفترض هذه الدراسة عدم تأثر هذه العلامات بحكمة الاختبار. ومن المؤشرات على تحقق هذا الافتراض مراجعة مدرسي المساق للفقرات وتأكدهم من خلو الفقرات من مفاتيح حكمة الاختبار عدا ما يندرج تحت المعرفة الجزئية، وهو ما تفترض الدراسة تكافؤه في المجموعتين الضابطة والتجريبية.

أما بالنسبة للأدوات التي استخدمت في الدراسات السابقة، فقد كان بعضها من إعداد الباحثين مثل دراسة وولي وليمز (١٩٨٢)، حيث كان عدد الفقرات ٢٤ (تحتوي ١٢ فقرة منها على مفاتيح الإجابة)؛ ودراسة إيفانز (١٩٨٤) حيث كان عدد الفقرات ٦٧ فقرة، ١٥ فقرة منها تتضمن مفاتيح حكمة الاختبار وباقي الفقرات عبارة عن اختبار تحصيلي. وقد استخدمت دراسات أخرى اختبارات مقننة، يتوفر فيها مفاتيح حكمة الإجابة، أو تنطبق عليها استراتيجيات حكمة الاختبار الواردة في تصنيف ميللمان ورفيقيه، مثل دراسة كلستين (١٩٨١)، ودراسة سمث (١٩٨٢). وقد اهتمت الدراسة التي اعتمدت أدوات قياس من إعداد الباحثين في تغطية الأداة للاستراتيجيات الواردة في التجربة، وعلى أي حال فإن موضوع حكمة الاختبار ما زال بحاجة إلى دراسات في مجال تطوير اختبارات لهذه المهارة، بحيث تتمتع بخصائص سيكومترية مقبولة لمراحل عمرية ومستويات أكاديمية وفتات عرقية مختلفة؛ لتمكين الباحثين من إجراء دراسات متعددة الأغراض على هذه المهارة. فقد تراوحت معاملات الثبات في الدراسات السابقة بين ٣٠، - ٧٠، ويعود هذا الانخفاض إلى ضيق المدى (أي تجانس أداء مجموعات الثبات) في علامات مجموعات الثبات، وهذا يشير ضمناً إلى ضرورة اكتساب مجموعات الثبات لمهارة

حكمة الاختبار قبل إصدار حكمٍ على دقة نتائجه، وهذا ما لاحظناه من تحسن في معامل الثبات عند تغيير خصائص المجموعة بعد تعليمها بعض عناصر حكمة الاختيار خلال فترة التجربة. حيث ارتفع معامل الثبات النصفى المصحح من ٠,٤٦ إلى ٠,٦٤ .

التجربة والتصميم التجريبي

ورد في الحديث عن عينة الدراسة أن هناك مجموعة ضابطة ومجموعة تجريبية، كما جاء في الحديث عن أداة الدراسة أن هناك متغيراً مصاحباً بغرض ضبط الفروق المحتملة بين المجموعتين في التحصيل مع الافتراض بعدم تأثر المتغير المصاحب بتعلم استراتيجيات حكمة الاختبار، ولذلك استخدم تحليل التباين المصاحب الذي يناسب هذا التصميم، وهناك عدة مؤشرات تدل على تحقق هذا الافتراض، أهمها التأكد من خلوفقرات الاختبار التحصيلي من المفاتيح التي توفر فرصة استخدام استراتيجيات التخمين، وقد تمت الإشارة إليها في الحديث عن أداة الدراسة، كما أنه لم يتم مناقشة الموضوعات المتعلقة بصياغة الفقرات المختلفة، والإرشادات الخاصة بفقرات الاختبار من متعدد مع الطلبة في المساق، سواء في المجموعة التجريبية أو الضابطة إلا بعد انتهاء التجربة، وأنه لا توجد أي خلفية سابقة عن هذه الموضوعات عند أي من طلبة المجموعتين .

أعطيت المجموعة التجريبية محاضرتين (ساعة لكل محاضرة في يومين متتاليين) عن مفهوم حكمة الاختبار والاستراتيجيات المختلفة، التي توفر هذه الحكمة (وهي: استراتيجيات التحضير والتهيئة، واستراتيجيات استغلال الوقت، واستراتيجيات تجنب الأخطاء، واستراتيجيات التخمين) دون الخوض في تفاصيل هذه الاستراتيجيات؛ وقد خصص الجزء الأكبر من المحاضرتين للحديث والنقاش حول عناصر استراتيجية التخمين في تصنيف فردركسون (Fredrickson) مع إعطاء بعض الأمثلة التوضيحية. وقد كانت الأمثلة من نوعين، الأول: أمثلة مكشوفة غير متخصصة، أو لا تحتاج عمقا في التفكير، مثل: الجذر التربيعي للعدد ١٦ هو

أ.	٨	جـ.	٤
ب.	٤/٨	د.	٢

أما النوع الثاني فكان متطرفاً في صياغته من حيث غموض المطلوب في الفقرة، حيث كانت فقرات بدون جذر (أي البدائل فقط)، وذلك لإظهار أهمية حكمة الاختبار كمهارة يمكن استغلالها في التوصل إلى إجابات ذكية في مواقف غامضة نسبياً، وأهمية هذه المهارة في المواقف الحياتية المختلفة، ومن هذه الأمثلة المثال التالي :

- أ - حل الواجبات المدرسية .
ب - قبل تحضير الدروس .
ج - القراءة .
د - قبل الذهاب إلى السينما .

حيث تتطلب الاجابة عن مثل هذه الفقرات مهارة التخمين التقاربي (انظر التعريفات الإجرائية) وفي هذا المثال تتطلب الإجابة استخدام البعد الزمني واستغلال الزمن في المذاكرة والواجبات المدرسية. أما المجموعة الضابطة فقد استمرت بالمحاضرات العادية حسب مخطط المساق، وبعد أسبوع من التجربة أعطي الطلبة في المجموعتين الاختبار التحصيلي في المساق، الذي يتضمن فقرات حكم الاختبار، كما اتضح في الحديث عن أداة الدراسة. ويلاحظ في الدراسات السابقة الاختلاف البسيط في فترة التدريب مع اختلاف الاستراتيجيات الداخلة في التجربة فقد تراوحت فترة التدريب بين ساعة إلى خمس ساعات عدا تلك التي اعتمدت الألفة بياني الاختبار، وتكرار الاختبارات كاستراتيجية تدريب. ولم يكن هدف أي من الدراسات السابقة الوصول إلى مستوى إتقان مهارة حكمة الاختبار، وإنما الكشف عن إمكانية تعلمها، وأثر البرنامج التدريسي في تطوير هذه المهارة من خلال مقارنة أداء مجموعة ضابطة بمجموعة تجريبية.

روعي في إدارة الاختبار توفير أقصى درجات الضبط للتدخل الأداء بين الطبة، حيث وضعت كل مجموعة (شعبة) في قاعة خاصة، وكان المراقب هو مدرس المساق لتلك الشعبة. كما كانت القاعات واسعة، بحيث توفر جلوس الطلبة في أماكن متباعدة. كما حرص مدرسو المساق أن لا يشعروا الطلبة بما فيهم المجموعة التجريبية أنهم يخضعون للتجربة، حتى لا يتم التدخل بينهم في مجال حكمة الاختبار على الرغم من قصر فترة التدريب، بالإضافة إلى قلة الفرص المتوفرة بين طلبة المجموعتين للتدخل نتيجة للبرنامج المكثف في الفصل الصيفي الجامعي، ووجود محاضرات المجموعتين في مبنيين متباعدين من مباني الجامعة. وقد طلب من كل طالب في تعليمات الاختبار أن يجيب على جميع الأسئلة، ولا يترك منها أي فقرة دون إجابة، حيث إن العلامة لم تصحح لأثر التخمين؛ كما لم يسمح بالأسئلة في أثناء الإجابة حتى لا يتم إسترجار المراقب إلى وضع يوحي فيه بوجود بعض الفقرات من مادة لم يتم إعطاؤها لهم. وعلى الرغم من الرضا الذي يشعر فيه الباحث عن إجراءات الضبط المشار إليها، إلا أن الاقتراح باستخدام متغيرات ضابطة أخرى مثل اختبارات القدرات العقلية يبقى وارداً على الرغم من ارتباط نتائج اختبارات القدرة العقلية ونتائج اختبارات التحصيل بدرجة عالية نسبياً.

النتائج ومناقشتها

هدفت هذه الدراسة إلى معرفة مستوى مهارة حكمة الاختبار لدى الطالب الجامعي، وإمكانية رفع مستوى هذه المهارة بتعليم الطلبة استراتيجيات التخمين لحكمة الاختبار. ولتحقيق هذه الأهداف حددت علامتان لكل طالب في المجموعتين الضابطة والتجريبية، علامة تعبر عن تحصيل الطالب، وعلامة تعبر عن مهارة حكمة الاختبار. ومن هذه العلامات أمكن فحص فرضيتي الدراسة السابقة الذكر.

نصت الفرضية الأولى على أن متوسط علامات الطلبة على اختبار حكمة الاختبار قبل تعليمهم استراتيجيات حكمة الاختبار لا يزيد عن التوقع طبقاً للتخمين العشوائي. وبما أن التوقع يساوي ٦,٠٠، حيث إن عدد الفقرات يساوي ٢٤، ويحتوي كل منها على أربعة بدائل، فإن المتوسط الحسابي لعلامات الطلبة في المجموعة الضابطة على متغير حكمة الاختبار (وهو هنا ٨,٧٩ والانحراف المعياري للعلامات ٢,٧٦) أعلى من التوقع عند مستوى الدلالة ٠,٠٥، حيث كانت قيمة الإحصائي ت تساوي ٦,٢٣٦ بينما القيمة الحرجة عند مستوى الدلالة ٠,٠٥ ودرجات حرية ٣٧ تساوي ٠,٢٧٢. وتتفق هذه النتيجة مع نتيجة دراسة سميث (Smith, 1982)، حيث كان المتوسط الحسابي في تلك الدراسة ٧,٤ بانحراف معياري، يساوي ٢,٨، والتوقع يساوي ٥. ويمكن عزو الفرق الإحصائي في المجموعة الضابطة في هذه الدراسة إلى خبرة الطالب الجامعي في الاختبارات خلال المراحل التعليمية السابقة، ومعرفتهم بجوانب الضعف المحتملة في صياغة فقرات الاختبار من متعدد. وإخراج كراسة الاختبار من قبل المدرسين. على الرغم من أن خبرة الطالب في الاختبارات ترفع مهارة الحكمة بقيمة دالة إحصائية، إلا أنه من الناحية العملية يبقى هذا المتوسط منخفضاً مقارنة بأقصى علامة ممكنة، وهي هنا ٢٤، أو مقارنة بعلامة الإلتقان المقترحة في مشكلة الدراسة، وهي ٨٠ من مئة، أي ما يعادل ١٩,٢ وهي أعلى (رقمياً) من ضعف المتوسط المشار إليه (٨,٧٩).

أما الفرضية الثانية، فقد هدفت إلى الكشف عن إمكانية رفع مستوى مهارة حكمة الاختبار بالتعليم، وبين الجدول رقم (١) المتوسطات الحسابية والانحرافات المعيارية لعلامات الطلبة في المجموعتين الضابطة والتجريبية على اختبار التحصيل (المتغير المصاحب)، ومقياس حكمة الاختبار (المتغير التابع). كما بين الجدول رقم (٢) نتائج تحليل التباين المصاحب.

جدول رقم (١)

المتوسطات الحسابية والانحرافات المعيارية لعلامات الطلبة في المجموعتين التجريبية والضابطة على اختبار التحصيل ومقياس حكمة الاختبار والدلالة الإحصائية للفرق بين المتوسطين لكل من المتغيرين

المتغير	المجموعة	العدد	الوسط الحسابي	الانحراف المعياري	ت	الدلالة الإحصائية
التحصيل حكمة الاختبار	التجريبية	٣٦	٢٧,٠٠	٥,٠٥	١,٤٧	,١٤٦
	الضابطة	٣٨	٢٥,٣٧	٤,٥٠		
	التجريبية	٣٦	١٠,٤٢	٢,٣٧	٢,٧١	,٠٠٨
	الضابطة	٣٨	٨,٧٩	٢,٧٦		

* وهي أعلى من الدلالة الإحصائية ٠,٠٥ التي اعتمدتها الدراسة .

جدول رقم (٢)

نتائج تحليل التباين المصاحب لعلامات الطلبة على مقياس حكمة الاختبار

مصدر التباين	درجات الحرية	متوسط المربعات	ف	الدلالة الإحصائية
المجموعات الخطأ	١ ٧١	٣٦,٢١٢ ٦,٣٧٦	٥,٦٧٩	*,٠٢٠

* وهي أعلى من الدلالة الإحصائية ٠,٠٥ التي اعتمدتها الدراسة

تشير هذه النتائج إلى فرق دال إحصائياً بين متوسط علامات المجموعتين على حكمة الاختبار بمستوى دلالة ٠,٢ , لصالح المجموعة التي تعلمت استراتيجية التخمين لحكمة الاختبار، مع ملاحظة أن الدلالة الإحصائية للفرق كانت ٠,٠٠٨ , قبل ضبط أثر المتغير المصاحب، وهذا يشير إلى أهمية اختيار التصميم الذي يوفر درجة عالية من الضبط، خاصة أن الدراسات السابقة ركزت على إمكانية رفع مستوى مهارة حكمة الاختبار بالتعليم، مما

يشير إلى تعذر تحسين هذه المهارة بسهولة، وهذا ما تؤكدته نتائج هذه الدراسة، إذ تتحدد الدلالة العملية للفرق بين المتوسطين بمدى فعالية البرنامج التعليمي لحكمة الاختبار، والاستراتيجيات الداخلية في هذا البرنامج، ودرجة التشابه بين مفاتيح حكمة الاختبار الواردة في فقرات المقياس، وتلك الواردة في البرنامج التعليمي. فالبرنامج التعليمي في الدراسة الحالية عبارة عن مقدمة في تعليم استراتيجيات التخمين الواردة في تصنيف فردركسون (١٩٨٤)، وقد تضمن البرنامج جميع استراتيجيات التخمين، مما يضيف صعوبة في اختيار الطالب لاستراتيجية التخمين المناسبة، أو صعوبة في ترجيح دور استراتيجية على أخرى عندما تتضمن الفقرة الواحدة مفاتيح ترتبط بأكثر من استراتيجية، وهو ما أشار إليه ميللمان ورفيقاه (١٩٦٥) في مراجعتهم للأدب التربوي في مجال حكمة الاختبار. كما أن الأمثلة في البرنامج التعليمي متطرفة من حيث التعقيد (انظر الطريقة والإجراءات)، وبالتالي فإن الفقرات أعلى من مستوى المعرفة في تصنيف بلوم للأهداف في المجال المعرفي، طالما أنها قدرة معرفية، كما ورد في تعريف سارناكي (١٩٧٩) لحكمة الاختبار.

الخلاصة

يتضح من خلال المتوسط الحسابي لعلامات كل من المجموعتين على مقياس حكمة الاختبار وموقع كل منهما بالنسبة للعلامة القصوى على المقياس أنه من الصعب الاعتماد على خبرة الطالب وألفته بالاختبارات في اكتساب مهارة حكمة الاختبار بالمستوى المقبول، كما أن الدور الذي تلعبه مهارة حكمة الاختبار في الذكاء الاجتماعي، والخصائص السيكومترية. للاختبارات، (كما بينته الدراسات السابقة) يشير إلى أن الواقع التربوي لا يحتمل تأجيل اكتساب هذه المهارة للمرحلة الجامعية أو إهمالها كلياً، طالما أنه يمكن تعليمها للطلبة في رياض الأطفال. كما يتضح أيضاً أنه ليس من السهولة اكتساب مهارة حكمة الاختبار، وأن المواقف الحياتية تختلف في درجة تعقدها، وبالتالي فهي تختلف في درجة حاجتها لمستوى الاتقان لهذه المهارة، وهو ما أوضحه كل من فردركسون (Frderickson, 1984) ومدلتن (Middleton, 1985)، ومكفيل (Mc Phail, 1981). وهي كأي مهارة معرفية، يمكن أن تكون هدفاً تربوياً عاماً يتطلب تحقيقه ما تتطلبه الأهداف الأخرى من بيئة تربوية غنية. ولذلك يمكن أن يكون تطوير برامج تعليم حكمة الاختبار بجميع استراتيجياتها، وتقويم فعاليتها، وانتقال أثر تعلم هذه المهارة في الذكاء الاجتماعي، والخصائص السيكومترية للاختبارات، وجوانب القصور في صياغة فقرات الامتحانات المدرسية التي تسمح بممارسة هذه المهارة، وأثر اتقان مهارة استخدام استراتيجيات

التخمين والاستراتيجيات الأخرى في تصنيف ميللمان ورفيقه أو تصنيف فردركسون في قلق الاختبار واتجاهات الطلبة نحو الامتحانات مجالا الدراسات أخرى .

المراجع

- 1 - Diamond, J. and Evans, W. "An investigation of the cognitive correlates of test-wisenes." **Journal of Educational Measurement**, 1972, 9: 145-150.
- 2 - Dolly, J. and williams, K. Using TEst-Takeing strategies to Maximize Multiple-Choice Test-Scores. **Educational and Psychological Measurement**, 1986, 46 : 619-625.
- 3 - Dreisbach, M. and Keogh, b. "Testwiseness as a factor in Readiness Test Performance of Young Mexican-American Children." **Journal of Educational Psychology**, 1982, 74: 224-229.
- 4 - Ebel, R. Measuring Educational Achievement, Englewood Cliifs, New Jersty., Prentice-Hall, 1965.
- 5 - Evans, W. "Testwiseness: An Examination of Cue-Using Strategies." **Journal of Experimental Education**, 1984, 52: 141-144.
- 6 - Fredrickson, L. "Teaching Test-taking Skills." **Social Studies Review**, 1984, 23: 23-28.
- 7 - Kalechstein, P., Kalechstein, M. and Doctor, R. "The effect of Instruction on test-taking skills in second-grade black children." **Measurement and Evaluation in Guidance**, 1981, 13 : 198-202.
- 8 - McMorris, R., Brown, J., Snyder, G. and Pruzek, R. "Effect of violating item construction principles." **Journal of Edcuational Measurement**, 1972, 9 : 287-295.
- 9 - Mcphail, I., "Why teach testwiseness." **Journal of Reading**, 1981, 25: 23-38.
- 10 - Mentzer, J. "Response Biases in Multiple-choice Test Item Files" **Educational and Psycholgical Measurement**, 1982, 42, 437-447.
- 11 - Middleton, J. "I Just Can't take tests." **The Science Teacher**, 1985, 52: 34-35.
- 12 - Milman, J., Bishop, C. and Ebel, R. "An Analysis of Test **Wiseness**." **Educational and Psychological Measurement**, 1965, 25: 707-726.
- 13 - Ritter, S. and Idol-Meastas, L. "Teaching Middle School students to use a test-taking strategy." **Journal of Educational Research**, 1986, 79: 350-357.
- 14 - Samacki, R. "An Examination of Testwiseness in the cognitive Test Domain." **Review of Educational Research**, 1979, 2: 252-279.
- 15 - Scruggs, T., White, K., and Bennion, K. Teaching Test-taking skills to Elementary Grade Students: A **Meta-analysis**. **The Elementary School Journal**; 1986, 87: 69-80.
- 16 - Smith, J. "Converging on Correct answers: A peculiarity of Multiple choice items." **Journal of Educational Measurement**, 1982, 19: 211-220.