المملكة الأردنية الهاشمية

المركز الوطني لتنمية الموارد البشرية *

تقرير رقم (١٤)

تحليل أخطاء أداء طلبة الصف الثامن الأساسي في الأردن على اختبار العلوم

إعـــداد

الدكتور موسى النبهان

تموز / ١٩٩٥

سلسلة منشورات المركز (٣٩)

* المركز الوطني لتنمية الموارد البشرية اعتباراً من أيار/مايو ١٩٩٥، وكان الإسم قبل ذلك هو المركز الوطني للبحث والتطوير التربوي.

بسم الله الرحمن الرحيم

كلمة شكر

بعد أن ظهر هذا العمل إلى حيّز الوجود ، لابد من تقديم الشكر والامتنان إلى الزملاء الباحثين في المركز الوطني للتنمية الموارد البشرية لما أبدوه من ملاحظات كان لها كبير الأثر في إثراء هذا العمل وتطويره.

كما أشكر كلاً من السيد رمزي أبوغزالة، لمساهمته في إجراء التحليلات الإحصائية المطلوبة، والآنسة هالة المسعود ، على طباعتها الأنيقة لهذه الدراسة. وان نسيت فيجب أن أذكر جهود السيد عزمي جرار المشهودة في تدقيق وتحرير نص الورقة.

تقديـــــــم

تأتي هذه الدراسة في سلسلة منشورات المركز الوطني لتنمية الموارد البشرية، التي كانت تهدف بمجموعها إلى إلقاء الضوء على مسائل تربوية هامة، والعمل بعد ذلك على تطوير الواقع التربوي في الأردن، ومن ثم النهوض به إلى مستوى أمثل.

عالجت هذه الدراسة الوصف الكمي والنوعي لأداء طلبة الصف الثامن الأساسي على اختبار العلوم الذي شملته الدراسة من زاوية فنية وتكنيكية، قد لايرتاح إلى الخوض بها إلاّ من توافرت لديه مهارات بمستوى معين في قراءة، وفهم الخصائص السيكومتوية لفقرات الاختبار، وما لهذه الخصائص من آثار على درجة دقة ومصداقية ذلك الاختبار كأداة تستخدم في الحكم على مستوى الأداء الأكاديمي للمتعلّمين.

ومهما يكن من أمر، فإن نتائج هذه الدراسة ستنير الطريق أمام مطوّري الاختبار ومستخدميه وذلك لتهذيبه وإخراجه بشكل أفضل للوصول إلى قرار يتميز بالدقة والحكمة والفاعلية.

رئيس المركز

د. منذر المصري

Error Analysis of Eighth Grade Students' Performance on the National Science Test

Abstract

The main goal of this study was to analyze types of errors the eighth grade students have committed on the science test which was developed by the National Center for Educational Research and Development.

This analysis has been accomplished by scanning the psychometric properties of the test items like item difficulty, item discrimination, distractor attractiveness and blind guessing.

At the begining, the scores of 1740 students were analyzed. These were the stratified study sample selected from schools at the end of the school year 1993/1994.

Results of the study showed that the performance level of the eighth graders was low (46%) and the score distribution was positively skewed.

Scanning the students' performance errors revealed that the mistakes could be classified into conceptual errors, item construction and test instructions.

Relatively, the items on the 2nd half of the whole test showed lower quality compared with the items on the first part. This reflects the effect of test length on the quality of student performance.

The results also showed that the items which represent the unit of "living creatures and resisting diseases" were of low psychometric level. However, these itmes were well-constructed.

The study recommended that the test length should be suitable to students age. It further suggested the deletion of some weak items to end up with an efficient measurement instrument. Finally, the study stressed the importance of conducting a training session for the science supervisors.

تحليل أخطاء أداء طلبة الصف الثامن الأساسي في الأردن على اختبار العلوم

الخلاصـــة

هدفت هذه الدراسة إلى تحليل أنواع الأخطاء التي ارتكبها طلبة الصف الثامن الأساسي على اختبار العلوم الذي تم اعداده وتطويره في المركز الوطني للبحث والتطوير التربوي، وذلك من خلال مسح الخصائص السيكومترية لفقرات الاختبار كمعاملات الصعوبة والتمييز وجاذبية الموهات.

تم استعراض نوعية أداء عينة الدراسة العشوائية المؤلفة من ١٧٤٠ طالباً/طالبة تم اختيارهم من مدارس المملكة في نهاية العام الدراسي ١٩٩٣/١٩٩٤ من خلال مشروع مراقبة الأداء المدرسي الذي رعاه ونفذه "المركز"، بالتعاون مع وزارة التربية والتعليم، وعدد من المؤسسات التربوية الأخرى.

أشارت نتائج هذه الدراسة إى ضعف مستوى أداء الطلبة على الاختبار عموماً، فقد بلغ متوسط أدائهم (٤٦٪)، وبين توزيع علاماتهم أنها ملتوية التواءً موجباً.

وعند مسح خصائص فقرات الاختبار، أمكن تصنيف أنواع الأخطاء التي ارتكبها الطلبة المفحوصون إلى كونها أخطاء بسبب عدم استيعاب مفاهيم المادة الدراسية وأخطاء بسبب ضعف صياغة الفقرات واخراجها، وكذلك بسبب عوامل تتعلق بإدارة واخراج الاختبار. ولم يظهر وجود أخطاء في أداء الطلبة بسبب الضعف في المعالجات الكمية أو التطبيقات.

إن ظهور ضعف الأداء النسبي على الفقرات الواقعة في النصف الثاني من الاختبار يشير إلى أن الجزء الثاني من الكتاب لم يأخذ الاهتمام الكافي من عملية التدريس، وبخاصة إذ كان ذلك الجزء يتطلب استخدام وسائل تعليمية أكثر.

كما بينت النتائج تحديداً، بأن معظم الفقرات الضعيفة كانت تنتمي إلى وحدة "الكائنات الحية ومقاومة الأمراض" وهذه تقع في النصف الثاني من اختبار العلوم. وقد أشارت النتائج إلى أن الخلل في هذه الفقرات كان بسبب ندرة استخدام الوسائل التعليمية اللازمة لتدريس المفاهيم المتضمنة، إذ أنها لم تتضمن ضعفاً واضحاً في الصياغة.

وقد أوضحت هذه الدراسة ضرورة أن يتناسب حجم الاختبار مع مستوى الطلبة العمري، وذلك بضرورة حذف بعض الفقرات ضعيفة المؤشرات. كما بدت الحاجة ملحة لعقد دورة توجيهية لمشرفي العلوم.

هدفت هذه الدراسة إلى التعرّف على نوعية أداء طلبة الصف الثامن الأساسي في الأردن على اختبار العلوم الذي تم تطويره في نهاية العام الدراسي ١٩٩٣/١٩٩٤ لأغراض مشروع مراقبة وتقييم الأداء المدرسي الذي تم تنفيذه بإشراف المركز الوطني للبحث والتطوير التربوي، وذلك من خلال مسح أنواع الأخطاء التعلّمية التي ارتكبها الطلبة من أفراد عينة الدراسة البالغ عددهم ١٧٤٠ طالباً وطالبة. إذ تم تحديد هذه الأخطاء من خلال استقصاء المؤشرات السيكومترية لفقرات الاختبار، وتحليل المحتوى الدراسي للمادة التعليمية.

أشارت نتائج هذه الدراسة إلى أن معظم الأخطاء التي وقع فيها طلبة الصف الثامن الأساسي كانت من نوع أخطاء المفاهيم، والتي ربما يتم إرجاعها إلى طرق التدريس، أو كفاية الكتاب المدرسي في عرض الأفكار والمفاهيم، إضافة إلى أن السبب الرئيسي لبعض الأخطاء كان طول الاختبار نفسه، ووجود ضعف ما في الصياغة الفنية لبعض الفقرات.

كما أشارت نتائج الدراسة تحديداً إلى أن معظم الفقرات الضعيفة كانت تنتمي إلى وحدة الكائنات الحية ومقاومة الأمراض، وهذه تقع في النصف الثاني من الاختبار.

فهرس المحتويات

كلمة شكر ..... أ

تقديـــم ..... ب

ملخص الدراسة باللغة الانجليزية (Abstract) ..... ج

الخلاصة باللغة العربية ..... د

فهرس المحتويات ..... و

فهرس الجداول والأشكال ..... ز

مقدمة الدراسة وأهميتها ١

هدف الدراسة ٤

منهجية الدراسة ٥

الاختبار التحصيلي في العلوم للصف الثامن الأساسي ٦

نتائج الدراسة ٨

أ. وصف أداء الطلبة على فقرات الاختبار ٨

ب. تصنيف فقرات الاختبار حسب خصائصها السيكومترية ١٠

جـ. تصنيف الفقرات حسب الأخطاء التعلّمية ١٣

مناقشة النتائج والتوصيات ١٩

المراجـــع ٢٢

فهرس الجداول والأشكال

أولاً: الجداول

جدول رقم (١): لائحة المواصفات لاختبار العلوم للصف الثامن الأساسي ٦

جدول رقم (٢): معاملات ثبات اختبار العلوم والاختبارات الفرعية ٧

جدول رقم (٣): التوزيع التكراري لتحصيل طلبة الصف الثامن الأساسي في العلوم على الاختبار ٩

جدول رقم (٤): توزيع معاملات صعوبة فقرات اختبار العلوم للصف الثامن الأساسي ١٠

جدول رقم (٥): توزيع أعداد فقرات الاختبار حسب معاملات صعوبتها بالنسبة لأفراد مجموعتي الأداء ١١

جدول رقم (٦): مقارنة الخصائص السيكومترية لفقرات اختبار العلوم حسب مجموعتي الأداء ١٢

جدول رقم (٧): معاملات صعوبة وتمييز فقرات اختبار العلوم للصف الثامن الأساسي ومفتاح الإجابة ١٣

جدول رقم (٨): قيم جاذبية مموهات فقرات اختبار العلوم للصف الثامن الأساسي ١٥

جدول رقم (٩): توزيع المموهات ذات الجاذبية الموجبة حسب رقم الفقرة والوحدة الدراسية ١٦

جدول رقم (١٠): توزيع تكرارات استجابات الطلبة المفحوصين من مجموعة الأداء المرتفع على بعض من فقرات الاختبار ١٨

ثانياً الأشكال

شكل رقم (١): التمثيل البياني لتوزيع علامات طلبة الصف الثامن الأساسي في العلوم على الاختبار كلّه ٩

شكل رقم (٢): التمثيل البياني لأعداد فقرات الاختبار حسب صعوبتها بالنسبة لمجموعة الأداء ١١

مقدمة الدراسة واهميتها

تعتمد دقة وقوة القرار المتخذ بشأن مسألة تربوية ما على درجة صدق وموضوعية الإجراءات المتبعة في تقييم تلك المسألة، وتحتل أدوات القياس والتقويم حيزاً كبيراً من تلك الإجراءات، فيمكن بواسطتها جمع البيانات المتعلقة بموضوع البحث والاهتمام ليصار إلى تقديم وصف كمي أو كيفي يساعد بدوره في فهم وتفسير ذلك الموضوع، مما يؤدي في الأغلب إلى اتخاذ قرار مفصلي بشأنه (Mehren and Lehman, 1974; Gronlund, 1988, 1983).

من هنا تبرز أهمية أن تكون أدوات القياس على درجة عالية من الصدق والثبات والموضوعية. ويتم التأكد من مدى تحقق صدق وثبات أداة القياس بتطبيقها على عينة من المستهدفين ليصار إلى تطويرها وتعميمها على مجتمع الدراسة موضوع البحث. يعد كل هذا أمراً أساسياً حين يطلب وصف أداء مجموعة من المفحوصين على أدوات القياس بدقة عالية.

تعدّ عمليتا اختبار Testing وتقويم Evaluation مستوى أداء التلاميذ من أصعب المهام التي تواجه المعلم في عمله، لأن كلاً من هاتين العمليتين، يتطلّب تطوير أدوات قياس فعّالة (Lee, 1969. Allen and Yen, 1979)، وهذا بدوره يتطلّب مستوىً رفيعاً في:

أ) امتلاك المعرفة الكافية بالموضوع المنوي قياس أداء التلاميذ فيه.

ب) المقدرة على تحديد النتاجات التعلّمية المتوقع من المتعلم اكتسابها.

جـ) المقدرة على تحديد وصياغة الأهداف التعليمية المنوي قياس تحققها لدى المتعلّمين.

د) تفسير نتائج أداء المتعلمين على أداة القياس.

يمكن تصنيف اختبارات التحصيل في المراحل الدراسية المتوسطة والمتقدمة على أنها اختبارات معيارية المرجع (Norm Referenced Tests) يتم الاهتمام فيها بمدى تباين أداء المفحوصين في الاستجابة، إذ ينصب الاهتمام برتبة المفحوص وموقعه مقارنة مع علامات زملائه في الصف (Mehren and Lehman, 1974; Gronlund, 1988). وبذلك يمكن اعتبار اختبار العلوم للصف الثامن الأساسي على أنه اختبار معياري المرجع. وعلى الرغم من الاعتقاد بوجود فروق فردية بين المفحوصين في قدراتهم على الأداء، واختلافهم في درجة الفهم والاستيعاب للمفاهيم، وتباينهم في القدرة على التطبيق والتحليل وغيرها من القدرات، جاءت النتيجة حتمية باحتمال ارتكاب الطلبة المفحوصين لأخطاء

تعلّمية متنوعة أثناء أدائهم على الاختبار، إذ ليس من المتوقع أن يكون أداء الطلبة متجانساً بدرجة عالية، وبدون أخطاء على فقرات الاختبار. على عكس الهدف من الاختبارات محكية المرجع (Criterion Reference Tests) التي لا يكون تباين أداء المفحوصين على فقراتها أمراً يحتاج إلى تحقيق واكتشاف، إذ تمتاز فقرات هذا النوع من الاختبارات بسهولة نسبية تغطي أهدافاً تدريسية تتسم بالتفصيلية والشمول، ولايقارن أداء المفحوص عليها مع زملائه في نفس المجموعة التي ينتمي إليها، وانما يقارن بمعيار أو أداء.

لما كان واجب المعلمين يتلخص في تعليم التلاميذ اكتساب معارف ومهارات تساعدهم في الوصول إلى المستوى التعلّمي المطلوب، كان لابدّ من القيام بتقويم ما يتعلّمه هؤلاء التلاميذ بوسائل ذات فاعلية عالية تتمع بدرجة مقبولة من الثبات والصدق. من هنا جاءت فكرة الاختبار التحصيلي وضرورته في الوضع الصفي (Nitko, 1989; Mehren and Lehman, 1974). وتبرز أهمية الاختبار التحصيلي الصفي لكونه أداة تساعد في:

- تحديد مستوى المفحوصين في إتقان نشاطات صفية معينة.
- زيادة دافعية الطلبة، إذ يقوم المفحوص عادة بالتركيز على الموضوعات الدراسية التي يشعر بأنها مطلوبة في الامتحان.
- توفير فرصة للمعلم لتحديد نقاط قوة ومواطن ضعف الطلبة في التمكن من استيعاب مادة دراسية معينة.
- توفير تغذية راجعة آنية للمعلّم من أجل إعادة تدريس مفهوم معيّن، أو شرح وحدة دراسية للمرّة الثانية.
- مكافأة المبدعين، أو توفير فرص خاصة للطلبة بطيئي التعلّم.
- تطوير استراتيجيات تدريس فعّالة.
- إجراء التقويم المستمر.

تتكون علامة الطالب المفحوص الظاهرية Observed Score (x) على أي اختبار تحصيلي من جزءين، حقيقي True Score (T) و خطأ Error Score (E)، إذ أن X = T + E ويكون القياس على درجة عالية من الدقة والثبات عندما يتلاشى الجزء الخطأ من علامة المفحوص. وما تقليل الجزء الخطأ من علامة المفحوص وامكانية تحديده إلاّ أحد المهام الرئيسية للقياس.

تتنوع أخطاء القياس في طبيعتها، فقد تكون عشوائية أو منتظمة. كما تتنوع الأخطاء في مصادرها أو أسبابها، فقد تكون راجعة لإجراءات إدارية، أو لطريقة الصياغة، أو لوجود خلل عيني، أو ربما يبرز الخطأ نتيجة لوجود خلل في طرق التدريس (Nunally, 1978; Lord and Novick, 1968).

تأخذ عملية تحليل الأخطاء مجالين من الاهتمام، فإما أن يكون البحث منصباً على تحليل أنواع الأخطاء التي تساعد في عملية تفسير علامة المفحوص والتي عرّفها كيللي (Kelley, 1927) بثلاثة أنواع رئيسية هي:

- الأخطاء الناتجة عن حجم الاختلاف بين العلامة الحقيقية للمفحوص وعلامته الظاهرية على اختبار معين، ويمكن تقدير ذلك الاختلاف بدلالة الخطأ المعياري للقياس.

- الأخطاء الناتجة عن مدى التباين بين العلامة الظاهرية للمفحوص على اختبار وعلامته المقدّرة على اختبار آخر مكافئ، ويمكن تقدير ذلك المدى بدلالة مقدار الخطأ المعياري للتقدير.

- الأخطاء الناتجة عن مدى الاختلاف بين العلامة الحقيقية للمفحوص وعلامته الحقيقية المقدّرة.

أو قد يكون الاهتمام متركزاً على تحليل أنواع الأخطاء التعليمية المرتبطة بكيفية أداء المفحوصين على فقرات الاختبار (Crocker and Algina, 1986). إذ أن الاهتمام بخطأ تفسير علامات المفحوصين يعتبر أمراً تكنيكياً لا مجال للخوض فيه في هذه الدراسة، في حين يأخذ البحث في أنواع أخطاء الأداء منهجاً تطبيقياً إجرائياً يساعد في النهاية في تقدير المؤشر الفعلي على حجم الأداء الحقيقي للمفحوص على فقرات اختبار معين، أو في تقدير العلامة الحقيقية لأداء ذلك المفحوص.

وعلى أية حال، فإن مسح أخطاء المفحوصين وتصنيفها سيعود حتماً بالفائدة على:

١. مطوّري أدوات القياس عامة والاختبارات بشكل خاص، ليمكنهم ذلك من تطوير أدوات فعالة في تقدير مستوى الأداء.

٢. المدرسين أنفسهم، ليكتشف كل منهم الجوانب الإيجابية والأخرى السلبية في أنماط وطرق التدريس التي يتبعها، ليصار بعد كل ذلك إلى تعميق معرفة الطلبة بالمفاهيم والأفكار العلمية وإلى تصحيح معرفتهم الخاطئة وتصويب قدراتهم في التطبيق والتحليل وغيرها من المهارات.

لقد برزت أهمية هذه الدراسة في كونها محاولة لتشخيص وتصنيف أخطاء أداء طلبة الصف الثامن الأساسي على الاختبار الوطني لمادة العلوم الذي تم تطويره في المركز الوطني للبحث والتطوير التربوي لتكون

مرجعاً يستنير به التربويون عند تخطيط المادة التعليمية المتضمنة في الكتاب المدرسي، وكذلك عند تدريسها، وفيما يتعلق بالمهارات التي سيحويها دليل المعلم، كما أن ندرة الدراسات الأردنية التي اهتمت بمثل هذا الموضوع وتهدف إلى تحقيق الغرض نفسه، أعطى هذه الجهد تميزاً جاداً.

هدف الدراسة

تهدف هذه الدراسة إلى تحليل أداء طلبة الصف الثامن الأساسي وتصنيف أخطائهم في اختبار العلوم الذي تم إعداده وتطويره في المركز الوطني للبحث والتطوير التربوي، ومن ثم تطبيقه على عينة من ١٧٤٠ طالباً، تم اختيارهم عشوائياً من مدارس المملكة في نهاية العام الدراسي ١٩٩٣/١٩٩٤. وبشكل أدق، فقد حاولت الدراسة الإجابة عن الأسئلة البحثية التالية:

س١: كيف كان أداء طلبة الصف الثامن الأساسي في الأردن على اختبار العلوم؟

س٢: كيف تصنف فقرات الاختبار حسب خصائصها السيكومترية بمقارنة مجموعتي الأداء (مجموعة الأداء العالي ومجموعة الأداء المنخفض)؟

س٣: ما طبيعة أخطاء الأداء التي ظهرت في استجابات طلبة الصف الثامن الأساسي على اختبار العلوم من حيث كونها:

* **مفاهيم** (أخطاء سببها ضعف الطلبة في استيعاب المفاهيم العلمية)؟
* **تطبيقات** (أخطاء سببها ضعف الطلبة في اختيار القوانين أو المعادلات أو المبادئ في عملية التفسير)؟
* **معالجات كمية** (أخطاء سببها ضعف الطلبة في إجراء العمليات الرياضية المطلوبة)؟
* **إدارة وإخراج** (أخطاء سببها ضعف تعليمات الأداء على فقرات الاختبار أو آلية إدارة تطبيق الاختبار أو زمن الاختبار وترتيب الإجابات الصحيحة على البدائل)؟
* **صياغة** (أخطاء سببها ضعف صياغة متن الفقرة واختيار البدائل الجذابة لها)؟

منهجية الدراسة

للإجابة على سؤال الدراسة، وتحقيقاً للهدف من إجرائها، فقد تم تنفيذ الإجراءات التالية:

* الوقوف على الخصائص السيكومترية لاختبار العلوم، وذلك باستقصاء عدد من المؤشرات مثل مكونات الاختبار بدلالة وحدات المحتوى والمهارات المعرفية، ومستوى ثبات الاختبار، وتوزيع علامات الطلبة عليه.

* تصنيف الطلبة المفحوصين في مجموعتين، عالية الأداء، ومتدنية الأداء، على أساس أعلى وأدنى ٢٧٪ من العلامات الكلية على الاختبار، فالطلبة الذين تقع علاماتهم في مدى أعلى ٢٧٪ يشكلون المجموعة عالية الأداء، والطلبة الذين تقع علاماتهم في مدى أدنى ٢٧٪ من مجموع العلامات يشكلون المجموعة متدنية الأداء. وقد تم اختيار هذه النسبة بالذات لأنها الأقدر على إبراز مقارنة فاعلة، إضافة إلى الاعتقاد باعتدالية توزيع العلامات.

* الوقوف على الخصائص السيكومترية لفقرات الاختبار في ضوء مقارنة أداء مجموعتي الأداء، وذلك باستقصاء مؤشرات معاملات تمييز الفقرات ودرجة جاذبية كل مموه في كل فقرة، كذلك توزيع تلك الفقرات حسب الوحدة الدراسية التي تمثلها الفقرة. إذ تم تصنيف المموهات حسب درجة جاذبيتها على النحو التالي:

  - مموّه فعّال إذا كان معامل جاذبيته سالباً وكبيراً أقل من (-٠٫٠٣).
  - مموّه يحتاج إلى تعديل إذا كان معامل جاذبيته موجباً محصوراً بين (الصفر) و (+٠٫٠٣).
  - مموّه (غير فعّال) إذا كان معامل جاذبيته موجباً وعالياً، ويتوجب استبداله أو حذفه.

* تصنيف خصائص فقرات الاختبار حسب موقعها في الاختبار (في النصف الأول أو النصف الثاني).

* تصنيف أنواع الأخطاء التعليمية التي ارتكبها الطلبة المفحوصون.

* تفحص درجة التخمين العشوائي في الاستجابات، وكذلك درجة الغموض في البدائل.

وقد تم عرض الجداول والأشكال البيانية اللازمة لإخراج الدراسة بصيغة تسهل على القارئ، ومن هم في صنع القرار، عملية قراءتها واستيعابها، ومن ثم توظيفها في المواقف اللازمة.

الاختبار التحصيلي في العلوم للصف الثامن الأساسي

تكوّن الاختبار التحصيلي في العلوم للصف الثامن الأساسي من ٦٠ فقرة من نوع الاختيار من متعدد بواقع أربعة بدائل لكل فقرة لقياس تحصيل الطلبة في أربع عشرة وحدة من محتويات مادة العلوم هي: الطاقة والحرارة، المادة، أهمية الماء والهواء، الكائنات الحية، التغيرات في القشرة الأرضية، المغناطيسية، المركبات الكيماوية، الآلات البسيطة، الغذاء والصحة، النبات مصدر الغذاء، توليد الكهرباء والتحكم بها، أثر تقدم العلم في مكافحة المرض، الطقس والرصد الجوي، والكون وارتياد الفضاء. وقد توزعت فقرات الاختبار على ثلاثة مهارات معرفية هي المعرفة والفهم والتطبيق. وقد تعذّر إرفاق نسخة من الاختبار في ملحق الدراسة لأسباب فنية تتعلّق بالمحافظة على سريته بسبب التخطيط لتطبيق الاختبار نفسه في دراسات تربوية في وقت لاحق. ويشير الجدول رقم (١) إلى توزيع أرقام الفقرات على وحدات المحتوى والمهارات المعرفية (لائحة المواصفات).

جدول رقم (١)

لائحة المواصفات لإختبار العلوم للصف الثامن الأساسي

الوحدة	المهارة			المجموع
	المعرفة	الفهم	التطبيق	
الطاقة والحرارة	١، ٤، ١٠	٣، ٥، ٦، ٧، ١١	٢، ٩، ١٢	١١
المادة	١٤، ١٧، ١٨، ٢٢	١٥	١٦، ١٩، ٢٠، ٢١	٩
الماء والهواء	٨، ٢٣، ٢٤		٢٦	٤
الكائنات الحية	٢٧، ٢٨، ٢٩، ٣٢، ٣٣، ٣٤		٣٠، ٣١، ٣٥	٩
تغيرات القشرة	٣٦	٣٧		٢
المغناطيسية	٣٨	٣٩		٢
المركبات الكيماوية	٤١، ٤٣، ٤٥		٤٠، ٤٢، ٤٤	٦
الآلات البسيطة			٤٦	١
الغذاء والصحة	٤٧، ٤٨			٢
النبات	٦٠		٤٩	٢
توليد الكهرباء	٥٤، ٥٥	٥٢	٥٠، ٥١، ٥٣	٦
مقاومة الأمراض	٥٦			١
الطقس والرصد	٥٧	١٣	٢٥	٣
الكون والفضاء	٥٨	٥٩		٢
المجموع	٢٩	١١	٢٠	٦٠

والجدير بالذكر أن هناك تفاوتاً في عدد الفقرات التي قاست كل محتوى من المحتويات الأربع عشرة، وقد نالت وحدة الطاقة والحرارة ١٨٪ من مجموع فقرات الاختبار، في حين تضمنت وحدة الآلات البسيطة فقرة واحدة. أما توزيع الفقرات حسب المهارات المعرفية فيشير إلى أن عدد أسئلة المعرفة والفهم والتطبيق كان ٢٩، ١١، ٢٠ فقرة على الترتيب، وكان الزمن المخصص لهذا الاختبار هو ٦٠ دقيقة، وهو اختبار تحريري (ورقة وقلم). وقد تم حساب معاملات الثبات باستخدام طريقة كرونباخ ألفا للاختبار ككل، ولكل جزء من أجزاء المحتوى ولكل مهارة تكونت من فقرتين أو أكثر. ويبين الجدول رقم (٢) معاملات الثبات المحسوبة.

جدول رقم (٢)
معاملات ثبات اختبار العلوم ككل والاختبارات الفرعية

الوحدة	المهارة: المعرفة	الفهم	التطبيق	المجموع
الطاقة والحرارة	٣١ر٠	٤١ر٠	٢٢ر٠	٥٨ر٠
المـــادة	٢٠ر٠		٢٨ر٠	٤٤ر٠
الماء والهواء	١٩ر٠			٢٦ر٠
الكائنات الحية	٢٤ر٠		٠٢ر٠	٣٤ر٠
تغيرات القشرة				١٣ر٠
المغناطيسية				٣١ر٠
المركبات الكيماوية	٠٧ر٠		١٠ر٠	٢١ر٠
الآلات البسيطة				
الغذاء والصحة	٠٩ر٠			٠٩ر٠
النبـــات				٣٢ر٠
توليد الكهرباء	٢٢ر٠		٢٣ر٠	٤٦ر٠
مقاومة الأمراض				
الطقس والرصد				١٧ر٠
الكون والفضاء				٢٢ر٠
المجمـــوع	**٦٩ر٠**	**٥٣ر٠**	**٦١ر٠**	**٨٣ر٠**

يلاحظ من ذلك أن معامل ثبات الاختبار الكلي يساوي ٨٣ر٠، في حين تراوحت معاملات الثبات للمهارات المعرفية بين ٥٣ر٠ و ٦٩ر٠، ولوحدات المحتوى بين ٠٩ر٠ و ٥٨ر٠. وهذا أمر طبيعي إذا ما أخذ بعين الاعتبار مدى اعتماد معامل الثبات على عدد فقرات الاختبار. وتجدر الإشارة هنا إلى أن الأداء على وحدة توليد الكهرباء كان الأكثر استقراراً وثباتاً (٤٦ر٠)، وعدد فقراتها (٦) من بين وحدات المحتوى الأخرى، بعد وحدة الطاقة والحرارة التي تضمنت العدد الأكبر من الفقرات (١١ فقرة) وكان معامل ثباتها (٥٨ر٠).

نتائج الدراسة

تم عرض النتائج التي توصلت إليها الدراسة في ثلاثة مسارات هي:

أ) وصف أداء الطلبة على فقرات الاختبار ككل: ويتضمن ذلك عرضاً لتوزيع فقرات الاختبار على وحدات ومهارات محتوى المادة العلمية، ودرجة ثبات أداء الطلبة المفحوصين على الاختبار، وتوزيع علاماتهم بيانياً، وبعض المؤشرات الإحصائية المتعلقة بذلك التوزيع.

ب) تصنيف فقرات الاختبار في ضوء خصائصها السيكومترية، ويتضمن ذلك خصائص الفقرات حسب مجموعة الأداء (عالية الأداء ومتدنية الأداء).

جـ) تصنيف فقرات الاختبار حسب أنواع أخطاء الأداء التعلمية التي ظهرت في استجابات الطلبة. وقد تم استعراض واقع توزيع الإجابة الصحيحة على بدائل الفقرات، معاملات صعوبة وتمييز فقرات الاختبار، جاذبية المموهات وتوزيع المموهات الضعيفة حسب الوحدات الدراسية، كما تم استقصاء درجة غموض الفقرات والتخمين العشوائي أثناء الاستجابة على الفقرات. وتم كذلك استقصاء توزيع الفقرات حسب خصائصها، وحسب موقعها في الاختبار.

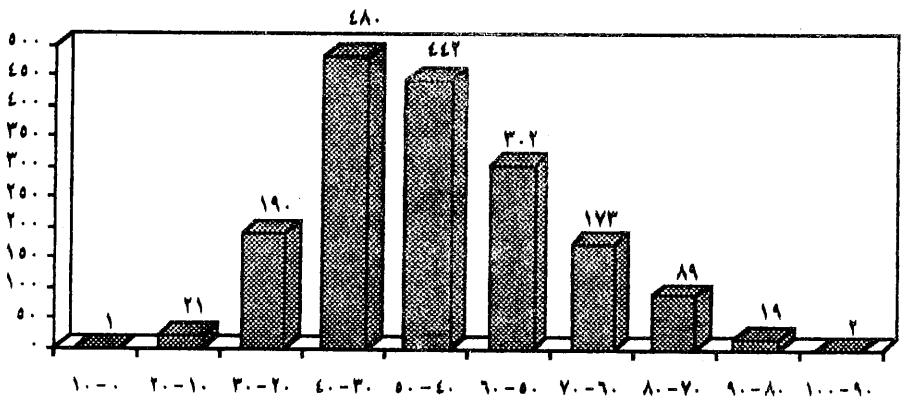
أ) وصف أداء الطلبة على فقرات الاختبار

تم تطبيق هذا الاختبار على عينة عشوائية تتألف من ١٧٤٠ طالباً وطالبة تم اختيارهم من ٢٤٥ مدرسة توزعت على محافظات المملكة روعي في اختيارها عوامل الجنس والسلطة المشرفة على المدرسة وموقع المدرسة (مدن، وريف). وبعد إدخالها إلى الحاسوب تم تصحيح الإجابات، وتم تحويل العلامات على الاختبار إلى علامات من مئة من أجل تسهيل قراءتها وتمثيلها جدولياً وبيانياً. أشارت البيانات إلى أن المتوسط الحسابي لعلامات مجموع الطلبة المفحوصين بلغ ٤٦٪، وأن مدى العلامات تراوح بين الصفر و ٩٣. ويبين الجدول رقم (٣) والشكل رقم (١) توزيعاً تكرارياً وتمثيلاً بيانياً لنسب الإجابة الصحيحة على كل الاختبار على الترتيب.

جدول رقم (٣)

التوزيع التكراري لتحصيل طلبة الصف الثامن الأساسي في العلوم على الاختبار

الفئـــة	العـــدد	النسبة المئوية
صفر - ١٠	١	٠ر٠٥٪
١٠ - ٢٠	٢١	١ر٢٪
٢٠ - ٣٠	١٩٠	١١ر٠٥٪
٣٠ - ٤٠	٤٨٠	٢٨٪
٤٠ - ٥٠	٤٤٢	٢٥ر٧٪
٥٠ - ٦٠	٣٠٢	١٧ر٦٪
٦٠ - ٧٠	١٧٣	١٠٪
٧٠ - ٨٠	٨٩	٥ر٢٪
٨٠ - ٩٠	١٩	١ر١٪
٩٠ - ١٠٠	٢	٠ر١٪



الشكل رقم (١): التمثيل البياني لتوزيع علامات طلبة الصف الثامن الأساسي في العلوم على الاختبار كله

من هذا كلّه يمكن القول بأن مستوى أداء طلبة الصف الثامن الأساسي على الاختبار كان متدنياً، كما يشير الشكل إلى أن توزيع علامات الطلبة ملتوٍ إلتواءً موجباً، الأمر الذي يشير إمّا إلى ضعف أداء الطلبة بسبب صعوبة فقرات الاختبار، أو أن تعلم الطلبة أساساً لم يكن بالمستوى المطلوب، إذ أن أكثر من ثلثي الطلبة لم تتجاوز علاماتهم الخمسين.

ب) تصنيف فقرات الاختبار حسب خصائصها السيكومترية

أشارت نتائج الدراسة إلى أن أكثر من نصف فقرات الاختبار كانت تتمتع بمعامل صعوبة يقل عن ٥٠ر٠، أي أن أقل من نصف الطلبة المفحوصين لم يتمكن من الإجابة على هذه الفقرات إجابة صحيحة. ويبين الجدول رقم (٤) توزيع معاملات صعوبة فقرات اختبار العلوم للصف الثامن الأساسي.

الجدول رقم (٤)

توزيع معاملات صعوبة فقرات اختبار العلوم للصف الثامن الأساسي

معامل الصعوبة	العدد	النسبة المئوية
≤ ٢٥ر٠	٨	١٣٪
٢٦ر٠ - ٤٠ر٠	١٥	٢٥٪
٤١ر٠ - ٥٠ر٠	١٢	٢٠٪
٥١ر٠ - ٦٠ر٠	١٢	٢٠٪
٦١ر٠ - ٧٩ر٠	١٣	٢٢٪
≤ ٨٠ر٠	صفر	صفر٪
	٦٠	١٠٠٪

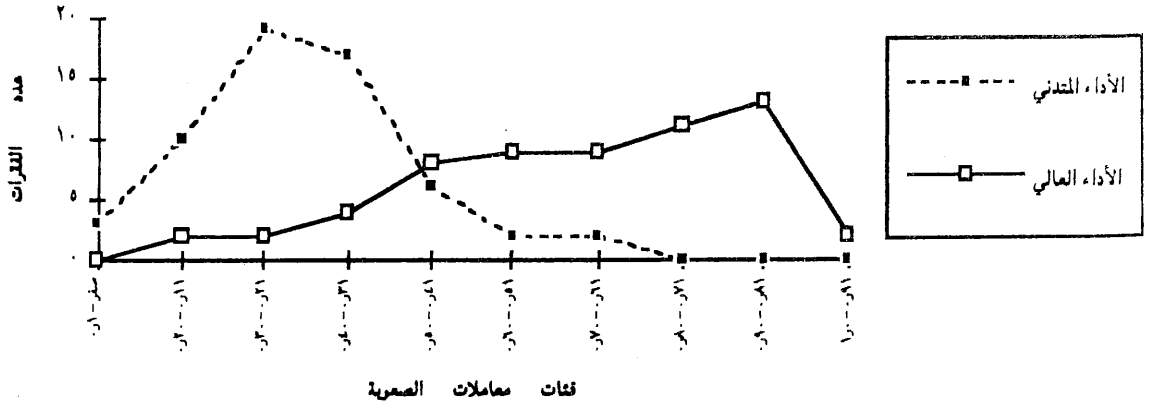
من الجدول رقم (٤)، يمكن ملاحظة أن ثماني فقرات (١٣٪) من فقرات الاختبار صعبة جداً، إذ تراوحت صعوبتها بين الصفر و ٢٥ر٠ وأن قرابة ٦٠٪ من فقرات الاختبار ككل أظهرت معاملات صعوبة أقل من ٥٠٪، وهذا يشير إلى الصعوبة النسبية للاختبار.

بعد ترتيب العلامات الكلية لطلبة الصف الثامن الأساسي الذين طبّق عليهم اختبار العلوم ترتيباً تنازلياً، تم تصنيف مجموعة الطلبة الذين أحرزوا أعلى ٢٧٪ من العلامات على أنها مجموعة الأداء العالي، في حين صنفت مجموعة الطلبة الذين أحرزوا أدنى ٢٧٪ من العلامات على أنها مجموعة الأداء المتدني، وكانت كل مجموعة تتألف من ٤٧٠ طالباً مفحوصاً.

كان متوسط علامات أفراد المجموعة عالية الأداء (٦٥٪)، وتراوحت علاماتهم من ٥٥٪ إلى ٩٣٪، بينما كان متوسط علامات أفراد المجموعة متدنية الأداء (٣٠٪)، ومدى علاماتهم من صفر إلى ٣٥٪، وقد تم اختبار دلالة الفرق بين متوسطي مجموعتي الأداء (٥٦٪، ٣٠٪) وذلك باستخدام الإحصائي (ت)، وقد بلغت قيمته المحسوبة (٨٠ر٥) والتي أشارت بدلالة عالية إلى الفرق بين أداء المجموعتين. وللتعرّف على عدد فقرات الاختبار مصنّفة حسب معاملات صعوبتها بالنسبة لأفراد مجموعتي الأداء يمكن الاطلاع على محتوى الجدول رقم (٥) والشكل رقم (٢).

الجدول رقم (٥)
توزيع أعداد فقرات الاختبار حسب معاملات صعوبتها
بالنسبة لأفراد مجموعتي الأداء

فئات معاملات الصعوبة	مجموعة الأداء	
	الأداء المتدني	الأداء العالي
صفر - ١ر٠	٣	صفر
١١ر٠ - ٢٠ر٠	١٠	٢
٢١ر٠ - ٣٠ر٠	١٩	٢
٣١ر٠ - ٤٠ر٠	١٧	٤
٤١ر٠ - ٥٠ر٠	٦	٨
٥١ر٠ - ٦٠ر٠	٢	٩
٦١ر٠ - ٧٠ر٠	٣	٩
٧١ر٠ - ٨٠ر٠	صفر	١١
٨١ر٠ - ٩٠ر٠	صفر	١٣
٩١ر٠ - ١ر٠	صفر	٢
المجمـــوع	٦٠	٦٠



الشكل رقم (٢): التمثيل البياني لأعداد فقرات الاختبار حسب صعوبتها بالنسبة لمجموعة الأداء

يمكن الاستنتاج بأن أكثر من نصف فقرات الاختبار (٣٢ فقرة) كانت صعبة، إذ تراوح معامل صعوبتها من صفر إلى ٣٠ر٠ بالنسبة للمفحوصين من مجموعة الأداء المتدني، بينما كانت هناك أربع فقرات فقط صعبة بالنسبة للمفحوصين من مجموعة الأداء العالي، ويبدو ذلك بوضوح أكثر عند

مطالعة الرسم البياني الوارد في الشكل رقم (٢)، فمنحنى مستوى صعوبة الفقرات بالنسبة لمجموعة الأداء المتدني كان ملتوياً إلتواءاً موجباً، في حين كان بالنسبة لمجموعة الأداء العالي ملتوياً إلتواءاً سالباً.

إن ما تم عرضه من خلال الجدول رقم (٥) والشكل رقم (١) يعد أمراً متوقعاً، وليس موضوعاً تم اكتشافه، إنما كان ذلك بهدف إبراز صورة جديدة عن وصف واقع الأداء وإيضاح وجه آخر لضعف التحصيل الذي يشير إلى وجود أخطاء متعددة في أداء طلبة الصف الثامن الأساسي، الأمر الذي شكّل دعامةً جديدة لضرورة إجراء هذه الدراسة.

ولتوضيح مدى اختلاف الخصائص السيكومترية لفقرات الاختبار حسب مجموعة الأداء، يمكن قراءة محتوى الجدول رقم (٦). ولا تفوتنا الإشارة هنا، إلى أن معامل تمييز الفقرة، كان قد تم حسابه بدلالة معامل ارتباط بوينت بايسيريل (الثنائي المتسلسل النقطي) بين الأداء على الفقرة والأداء على كامل الاختبار.

الجدول رقم (٦)

مقارنة الخصائص السيكومترية لفقرات اختبار العلوم حسب مجموعتي الأداء

مجموعة الأداء	الخاصية السيكومترية: مدى صعوبة الفقرات	مدى تمييز الفقرات	معامل الثبات
مجموعة الأداء المتدني (ن = ٤٧٠)	٠٨ر٠ - ٦٢ر٠	(-١٤ر٠) - (٠٨ر٠)	٢٩ر٠
مجموعة الأداء العالي (ن = ٤٧٠)	١٧ر٠ - ٩٣ر٠	(٠١ر٠) - (٢٦ر٠)	٥٢ر٠
مجموعة المفحوصين (ن = ١٧٤٠)	١٣ر٠ - ٧٩ر٠	(-٠٧ر٠) - (٤٧ر٠)	٨٣ر٠

ويمكن تفسير الاختلاف في درجات معامل الثبات للاختبار نفسه باختلاف مجموعة المفحوصين، إذ يمكن عرض ذلك في ضوء اختلاف درجة التباين في مستوى أداء هؤلاء المفحوصين، إذ يبدو أن أداء المفحوصين ذوي الأداء العالي أكثر تجانساً من أداء مجموع المفحوصين، إذ أن معامل ثبات مجموعة الأداء العالي (٥٢ر٠) أقل من معامل ثبات مجموعة المفحوصين (٨٣ر٠). أما أن يكون معامل ثبات الاختبار منخفضاً، فهو إنما يعكس ضعف قدرة الفقرات على التمييز من جهة وعدم ارتباط الأداء على معظم الفقرات والأداء على مجمل الاختبار من جهة أخرى بمستوى ذي دلالة. ومن الجدير ذكره أن ٤١ فقرة من فقرات الاختبار أظهرت تمييزاً سالباً لمستوى الأداء في حالة مجموعة الأداء المتدني، في حين لم يظهر التمييز السالب إلا في ٤ فقرات في مجموعة الأداء العالي.

جـ) تصنيف الفقرات حسب الأخطاء التعلّمية

للتعرّف على مستوى قدرة فقرات الاختبار على التمييز بين أداء الطلبة من مجموعة الأداء العالي وأداء الطلبة من مجموعة الأداء المتدني، تم حساب معاملات تمييز فقرات الاختبار كما ظهرت في الجدول رقم (٧)، وذلك بإيجاد نسبة الطلبة المفحوصين من مجموعة الأداء العالي الذين أجابوا الفقرة إجابة صحيحة مطروحاً منها نسبة الطلبة المفحوصين من مجموعة الأداء المنخفض الذين أجابوا الفقرة إجابة صحيحة.

الجدول رقم (٧)
معاملات صعوبة وتمييز فقرات اختبار العلوم
للصف الثامن الأساسي ومفتاح الإجابة

رقم الفقرة	مفتاح الإجابة	معامل الصعوبة	معامل التمييز	رقم الفقرة	مفتاح الإجابة	معامل الصعوبة	معامل التمييز	رقم الفقرة	مفتاح الإجابة	معامل الصعوبة	معامل التمييز
١	د	٠ر٦٤	٠ر٣٧	٢١	ب	٠ر٣٥	٠ر٤٨	٤١	د	٠ر٤٧	٠ر١٦
٢	جـ	٠ر٧٩	٠ر٣١	٢٢	جـ	٠ر١٤	٠ر٠٦	٤٢	ب	٠ر٤٦	٠ر٢٤
٣	أ	٠ر٤١	٠ر٤٨	٢٣	د	٠ر٢٠	٠ر٢٦	٤٣	أ	٠ر٢٣	٠ر٢٢
٤	د	٠ر٦٣	٠ر٤٢	٢٤	ب	٠ر٥٩	٠ر٤٣	٤٤	أ	٠ر٢٣	٠ر٢٨
٥	أ	٠ر٧٩	٠ر٣٤	٢٥	د	٠ر٦٠	٠ر٥٠	٤٥	ب	٠ر٤٨	٠ر٣٠
٦	أ	٠ر٦٤	٠ر٥٢	٢٦	ب	٠ر٣٩	٠ر٢٩	٤٦	أ	٠ر٢٨	٠ر٣٤
٧	د	٠ر٣٠	٠ر٣٤	٢٧	د	٠ر٥٧	٠ر٦٧	٤٧	د	٠ر١٣	٠ر١٤
٨	أ	٠ر٣٦	٠ر٣٩	٢٨	ب	٠ر٤٧	٠ر٣٨	٤٨	جـ	٠ر٦٣	٩ر٤٩
٩	د	٠ر٤٤	٠ر٥٥	٢٩	د	٠ر٣٦	٠ر١٤	٤٩	ب	٠ر٥٣	٠ر٦٠
١٠	جـ	٠ر٧٥	٠ر٣٠	٣٠	د	٠ر٤٠	٠ر١٦	٥٠	د	٠ر٥٦	٠ر٣٨
١١	ب	٠ر٣٣	٠ر٢٨	٣١	جـ	٠ر٢٥	٠ر٠٦	٥١	أ	٠ر٣٦	٠ر٤٢
١٢	جـ	٠ر٦٤	٠ر٢٦	٣٢	د	٠ر١٥	٠ر٢٢	٥٢	جـ	٠ر٤١	٠ر٤٤
١٣	ب	٠ر٥٤	٠ر٣٥	٣٣	أ	٠ر٣٧	٠ر٣٥	٥٣	ب	٠ر٢٣	٠ر٣٣
١٤	جـ	٠ر٣٦	٠ر٤٤	٣٤	ب	٠ر٢٩	٠ر١٦	٥٤	ب	٠ر٤٧	٠ر٥٤
١٥	ب	٠ر٥٨	٠ر٤٧	٣٥	د	٠ر٥٨	٠ر٥٢	٥٥	د	٠ر٥٧	٠ر٤٥
١٦	د	٠ر٤٤	٠ر٥١	٣٦	جـ	٠ر٦٧	٠ر٣١	٥٦	د	٠ر٥٠	٠ر٢٩
١٧	جـ	٠ر٥٦	٠ر٣٢	٣٧	د	٠ر٥٦	٠ر٣٠	٥٧	ب	٠ر٣٩	٠ر١٨
١٨	أ	٠ر٤٩	٠ر٤٦	٣٨	أ	٠ر٤٠	٠ر٤٨	٥٨	أ	٠ر٦١	٠ر٤٨
١٩	جـ	٠ر٦٩	٠ر٢٠	٣٩	ب	٠ر٤٨	٠ر٤٩	٥٩	أ	٠ر٤٤	٠ر٢٥
٢٠	د	٠ر٢٩	٠ر٣٠	٤٠	ب	٠ر٦٠	٠ر٤٩	٦٠	د	٠ر٦٤	٠ر٥٤

لقد جرى تصنيف الفقرة على أنها فعّالة إذا كان معامل تمييزها موجباً وعالياً وصعوبتها متوسطة. في حين ينصح باستبعاد الفقرة ذات التمييز السالب، وبتعديل وتطوير الفقرة ذات التمييز الموجب القليل أو القريب من الصفر (Crocker and Algina, 1986; Magnuson, 1976; Nitko, 1983; and Nunally, 1978).

يلاحظ من البيانات الواردة في الجدول رقم (٧)، بأن هناك خللاً في توزيع مكونات مفتاح الإجابة، أي أن هناك عدم توازن في توزيع عدد البدائل الصحيحة على فقرات الاختبار، فبدلاً من أن يكون التوزيع متوازناً بواقع ٢٥٪ لكل بديل، توزعت الإجابات الصحيحة على البدائل د، ب، أ، جـ بنسب ٣٤٪، ٢٥٪، ٢٢٪، ٢٠٪، على الترتيب، الأمر الذي سيؤدي إلى اختلاف فرصة المفحوص في الحصول على علامة ما بسبب اكتشافه لنمط الإجابة الصحيحة.

تراوح معامل الصعوبة لفقرات الاختبار بين ١٣ر٠ و ٧٩ر٠، وقد كانت الفقرات ٢، ٥، ١٠ هي الأسهل، في حين كانت الفقرات ٢٢، ٣٢، ٤٧ هي الأكثر صعوبة من بين فقرات الاختبار. وعند أخذ موقع الفقرات في الاختبار بعين الاعتبار، أمكن ملاحظة أن متوسط معاملات صعوبة الفقرات الواقعة في النصف الأول من الاختبار (الفقرات ١، ٢، ...، ٣٠) هو ٤٩ر٠، في حين كان متوسط معاملات صعوبة الفقرات الواقعة في النصف الأخير من الاختبار (الفقرات ٣١، ٣٢، ...، ٦٠) هو ٤٣ر٠، وهذا يشير إلى ضعف أداء الطلبة، أو إلى صعوبة الفقرات الواقعة في نهاية الاختبار.

ويلاحظ من البيانات الواردة في الجدول رقم (٧) أن جميع فقرات اختبار العلوم للصف الثامن الأساسي تتمتع بمعاملات تمييز موجبة وعالية، ما عدا بعض الفقرات التي يقل معامل تمييز كل منها عن (+٢٥ر٠) وهي: ١٩، ٢٢، ٢٩، ٣٠، ٣١، ٣٢، ٣٤، ٤١، ٤٢، ٤٣، ٤٧، ٥٧.

والملفت للانتباه أن ٦٧٪ من مجمل هذه الفقرات الضعيفة تقع في النصف الأخير من الاختبار، وأن ٤٢٪ منها تنتمي إلى وحدة الكائنات (الفقرات ٢٩، ٣٠، ٣١، ٣٢، ٣٤)، وأن ربع هذه الفقرات ينتمي إلى وحدة المركبات الكيماوية (٤١، ٤٢، ٤٣).

وللتعرّف على خصائص فقرات الاختبار بدلالة مقدرة بدائل كل فقرة على جذب الطلبة المفحوصين لاختيارها، تم احتساب جاذبية كل مموه (Distractor) من مموهات كل فقرة. وتعرف جاذبية المموه بأنها نسبة الطلبة من المجموعة عالية الأداء الذين اختاروا المموه مطروحاً منها نسبة الطلبة المفحوصين من المجموعة متدنية الأداء الذين اختاروا ذلك المموه. والجدير بالذكر أن المموه الأكثر فعالية والذي يفضل احتواء الفقرة عليه، هو ذلك الذي يكون معامل جاذبيته كبيراً وسالباً، إنه المموه الذي يجذب أكبر عدد ممكن من الطلبة الضعاف لاختياره، في حين لا يقبل على اختياره إلا عدد قليل من الطلبة المتفوقين (Brown, 1983). ويبين الجدول رقم (٨) قيم جاذبية كل مموه في كل فقرة من فقرات الاختبار.

الجدول رقم (٨)

قيم جاذبية مموهات فقرات اختبار العلوم للصف الثامن الأساسي

جاذبية المموهات / الفقرة	أ	ب	جـ	د	جاذبية المموهات / الفقرة	أ	ب	جـ	د
١	-١٨ر٠	-١١ر٠	-٠٨ر٠		٣١	-٠٨ر٠	-١٦ر٠		١٧ر٠
٢	-٠٤ر٠	-١٢ر٠		-١٤ر٠	٣٢	-٠٤ر٠	-١٧ر٠	-٠١ر٠	
٣		-٢٩ر٠	-١٦ر٠	-٠١ر٠	٣٣		-٠٨ر٠	-١١ر٠	-١٨ر٠
٤	-٠٨ر٠	-١٨ر٠	-١٦ر٠		٣٤	-٠٦ر٠		-٠٣ر٠	-٠٦ر٠
٥		-١٢ر٠	-٠٨ر٠	-١٤ر٠	٣٥	-١٨ر٠	-١٩ر٠	-١٣ر٠	
٦		-٢١ر٠	-٠٧ر٠	-٢٤ر٠	٣٦	-١٠ر٠	-٠٨ر٠		-١٣ر٠
٧	٠١ر٠	-١ر٠	-٢٤ر٠		٣٧	-٠٤ر٠	-١٩ر٠	-٠٨ر٠	
٨		-٢٣ر٠	-٠٦ر٠	-١١ر٠	٣٨		-١١ر٠	-١٥ر٠	-٢٢ر٠
٩	-١٩ر٠	-٢٥ر٠	-١١ر٠		٣٩	-٢٩ر٠		-١١ر٠	-٠٩ر٠
١٠	-١٠ر٠	-٠٣ر٠		-٢١ر٠	٤٠	-٠٨ر٠		-١٧ر٠	-٠٤ر٠
١١	-٢٣ر٠		-٠٤ر٠	-٠١ر٠	٤١	-٠٤ر٠	-١٥ر٠	٠٤ر٠	
١٢	-٠٥ر٠	-١٢ر٠		-١٠ر٠	٤٢	-٠٢ر٠		-١٥ر٠	-٠٧ر٠
١٣	-٠٣ر٠		-٢٠ر٠	-١٠ر٠	٤٣		-٠٨ر٠	١٣ر٠	-٠٣ر٠
١٤	-٣١ر٠	-٠٣ر٠		-٠٩ر٠	٤٤		-٠٥ر٠	-٠٣ر٠	-٢٠ر٠
١٥	-١٢ر٠		-١٨ر٠	-١٧ر٠	٤٥	-٠٩ر٠		-١٢ر٠	-٠٩ر٠
١٦	-١٧ر٠	-٢٤ر٠	-١٠ر٠		٤٦		-٠٩ر٠	-١٤ر٠	-١٠ر٠
١٧	٠٥ر٠	-١٨ر٠		-١٧ر٠	٤٧	-١٠ر٠	-٢١ر٠	١٨ر٠	
١٨		-١٣ر٠	-١٥ر٠	-١٨ر٠	٤٨	-٢٥ر٠	-١٤ر٠		-١٠ر٠
١٩	-٠٨ر٠	-٠٨ر٠		-٢٢ر٠	٤٩	-٤٥ر٠		-٠٦ر٠	-٠٨ر٠
٢٠	١٠ر٠	-٢١ر٠	-١٨ر٠		٥٠	-٠٦ر٠	-١٩ر٠	-١٣ر٠	
٢١	-٠٧ر٠		-١١ر٠	-٢٨ر٠	٥١	-٢٣ر٠	-١٣ر٠		-٠٧ر٠
٢٢	-٠٣ر٠	+١٠ر٠		-١١ر٠	٥٢	-١٣ر٠	-١٤ر٠		-١٧ر٠
٢٣	-٠٣ر٠	-٠٩ر٠	-١٥ر٠		٥٣	-٠١ر٠		-١٥ر٠	-١٢ر٠
٢٤	-٠٤ر٠		-٢٣ر٠	-١٠ر٠	٥٤	-٢١ر٠		-٢٠ر٠	-١٣ر٠
٢٥	-٣٨ر٠	-٠٩ر٠	-٠٣ر٠		٥٥	-٣٠ر٠	-٠٩ر٠	-٠٦ر٠	
٢٦	-٠٥ر٠		-٠٦ر٠	-١٨ر٠	٥٦	-١٦ر٠	-١٨ر٠	٠٤ر٠	
٢٧	-١٣ر٠	-٤٢ر٠	-١١ر٠		٥٧	-١٠ر٠			-٢٢ر٠
٢٨	-٠١ر٠		-٣٠ر٠	-٠٤ر٠	٥٨	-٣٥ر٠		-١٣ر٠	-٠٧ر٠
٢٩	١٩ر٠	-٢٦ر٠	-٠٧ر٠		٥٩	٠٤ر٠		-١٥ر٠	-١٣ر٠
٣٠	٠٩ر٠	-١٩ر٠	-٠٦ر٠		٦٠	-٢٠ر٠	-١٣ر٠	-٢٠ر٠	

ويلاحظ من البيانات الواردة في الجدول رقم (٨) أن معامل جاذبية عددٍ من مموهات عشر الفقرات كان موجباً يقارب الصفر، الأمر الذي يشير إلى أن نسبة الطلبة من المجموعة عالية الأداء التي اختارت

ذلك البديل كانت أكبر من مثيلتها من المجموعة متدنية الأداء، وهذا الخطأ في الإجابة يشير إلى خلل ما في طريقة التدريس لفكرة كل سؤال يحوي مموهاً من هذا النوع أو ضعف في صياغة تلك الفقرة، أو صعوبة من نوع ما لا تتناسب مع مستوى الطلبة المفحوصين أو ضعفاً في ممارسات الطلبة في المهارات الرياضية والتطبيقية.

وللتعرّف على توزيع مموهات الفقرات التي تتمتع بمعاملات جاذبية موجبة، يمكن مطالعة محتوى الجدول رقم (٩).

الجدول رقم (٩)
توزيع المموهات ذات الجاذبية الموجبة حسب رقم الفقرة والوحدة الدراسية

رقم الفقرة	المموه	معامل الجاذبية	الوحدة الدراسية
٢٠	أ	+ ٠١٠ر٠	المادة
٢٢	ب	+ ٠١٠ر٠	المادة
٢٩	أ	+ ٠١٩ر٠	الكائنات الحية
٣٠	أ	+ ٠٠٩ر٠	الكائنات الحية
٣١	د	+ ٠١٧ر٠	الكائنات الحية
٤١	جـ	+ ٠٠٤ر٠	المركبات الكيماوية
٤٣	جـ	+ ٠١٣ر٠	المركبات الكيماوية
٤٧	جـ	+ ٠١٨ر٠	الغذاء والصحة
٥٦	جـ	+ ٠٠٤ر٠	مقاومة الأمراض
٥٩	ب	+ ٠٠٤ر٠	الكون والفضاء

يلاحظ من البيانات الواردة في الجدول رقم (٩) أن عشر فقرات فقط هي التي تتضمن مموهات غير فعّالة (ذات جاذبية موجبة). وتراوح مدى جاذبية المموهات من +٠٠٤ر٠ إلى + ٠١٩ر٠، الأمر الذي يشير إلى أن عدد الطلبة من مجموعة الأداء العالي ممن اختاروا هذه المموهات كان أكبر من عدد الطلبة من مجموعة الأداء المنخفض الذين اختاروا تلك المموهات. والجدير بالذكر أن ٧٠٪ من هذه الفقرات يقع في النصف الثاني من الاختبار، وأن ٤٠٪ منها ينتمي إلى وحدة الكائنات الحية ومقاومة الأمراض، وأن ٦٠٪ منها تقيس مهارة المعرفة. وقد أمكن تصنيف نوع الخطأ الذي ارتكبه المفحوصون في الاستجابة على هذه الفقرات على النحو التالي:

١) أخطاء في صياغة الفقرة، إذ يبدو أن واقع كل من الفقرات ٢٢، ٢٩، ٥٦، ٣٠، ٤١، ٤٣، يتضمن خللاً ما في صياغته، وقد أمكن استعراض الضعف في كل فقرة على النحو التالي:

الفقرة رقم ٢٢: أثار وجود كلمة هيدروكسيد في البديل (ب) منبهاً جذب عدداً غير قليل من الطلبة المفحوصين لاختياره وذلك بسبب احتواء متن الفقرة على كلمة (أكسيد).

الفقرة رقم ٢٩: كان المموه (أ) في هذه الفقرة يتمتع بجاذبية موجبة عالية، إذ أن نسبة كبيرة جداً من الطلبة ذوي الأداء المرتفع اختارته أكثر من أولئك الذين ينتمون إلي مجموعة الأداء المنخفض. والجدير بالذكر أن هذا المموه كان أكثر جاذبية من الإجابة الصحيحة نفسها. وسبب كل ذلك يعود إلى ضعف الصياغة اللغوية المتعلقة باحتواء البديل على حرف الباء، فلأن المموه (بالبروتينات) كان يوحي بسلامة الاختيار من ناحية لغوية، إذ من المفروض أن يكون حرف (الباء) في متن الفقرة.

الفقرة رقم ٥٦: على الرغم من أن الخصائص السيكومترية لهذه الفقرة تشير بقوتها عموماً، إلا أن ضعفاً في صياغتها بدا واضحاً عندما تضمن المتن كلمة (الجسم) كمنبه للإجابة الصحيحة في البديل (د) وهو (الأجسام المضادة)، والذي خلت المموهات الباقية من أي إشارة لكلمة (الجسم). أما أن يكون البديل (ب) مموهاً ذي جاذبية سالبة، فمرد ذلك إلى أنه خطأ يصنف تحت ضعف استيعاب المفاهيم.

الفقرات ٣٠، ٤١، ٤٣: احتوت هذه الفقرات مموهات ضعيفة وغير مستحبة من مثل (جميع ما ذكر، لاشئ مما ذكر، بدائل مركبة)، الأمر الذي يؤدي إلى إحداث اضطراب في اختيار المفحوص للبديل الصحيح، إذ أن نسبة من المفحوصين تختار أول بديل صحيح يقع بعد قراءة متن الفقرة، ولايدري المفحوص إن كان هناك مموه مركب من بديلين أو أكثر.

وعند مراجعة فقرات الاختبار، أمكن تصنيف مصدر الخطأ في استجابة المفحوصين على الفقرتين ٢٣، ٣٠ على أنه خطأ في الصياغة. ففي الفقرة رقم ٢٣ كان مصدر الخطأ هو عدم إبراز (تضليل أو تكبير) كلمة (طبيعياً) في متن الفقرة في حين كان استخدام عبارة (جميع ما ذكر) في الفقرة رقم ٣٠ كإجابة صحيحة هو السبب وراء خطأ الطلبة في الاستجابة. فبدلاً من استخدام هذا البديل كإجابة صحيحة، يفضل البحث عن بدائل أخرى أو أن تكون الفقرة مفتوحة النهاية وليست فقرة موضوعية.

٢) ضعف مستوى فهم التلاميذ للمفاهيم والأفكار المتضمنة الفقرات ٢٠، ٣١، ٤٧، ٥٩، بسبب الصعوبة أو خلل في طريقة التدريس واحتوائها على مموهات تتمتع بجاذبية ضعيفة (موجبة).

ولتفحص درجة التخمين العشوائي (Blind Guessing) في استجابات المفحوصين على فقرات الاختبار ودرجة الغموض (Ambiguity) في البدائل، فقد تمت جدولة تكرارات إجابات الطلبة من

مجموعة الأداء المرتفع على بدائل كل فقرة، تم رصد ذلك فيما يتعلّق ببعض الفقرات التي برز فيها غموض في بعض بدائلها من مثل ٢٣، ٢٩، ٣٠، ٣١، ٣٢، ٤٧، كما في الجدول رقم (١٠).

الجدول رقم (١٠)
توزيع تكرارات استجابات الطلبة المفحوصين من مجموعة الأداء المرتفع على بعض فقرات الاختبار

الفقرة	البديل			
	أ	ب	جـ	د
٢٣	١٠١	٢٥	١٥٨	*١٧٢
٢٩	١٧٥	٥٤	٣٧	*١٩١
٣٠	٢٢٩	٥	١	*٢١٧
٣١	٤٨	٣٣	*١٤٥	٢٢٩
٣٢	٣	١٧٨	٣٢	*١٥٠
٤٧	٣٤	١٢٧	١٩٢	*٩٩

* الإجابة الصحيحة

يمكن ملاحظة درجة نزوع الطلبة المفحوصين إلى التخمين العشوائي في استجابتهم على الفقرات، وبخاصة فيما يتعلق بالاستجابة على الفقرة رقم ٤٧، إذ تساوى عدد المفحوصين النسبي من مجموعة الأداء المرتفع في اختيارهم للبدائل ب، جـ، د. والملفت للانتباه كذلك هو أن عدد المفحوصين من مجموعة الأداء المرتفع الذين اختاروا بعضاً من المموهات كان أكثر من الذين اختاروا الإجابة الصحيحة نفسها في الفقرات ٣٠، ٣١، ٣٢، ٤٧. إذ أشار ساكس (Sax, 1974) إلى أن البدائل تتسم بالغموض عندما لا يستطيع الطلبة المفحوصون التمييز بين البديل الصحيح (الإجابة الصحيحة) والمموهات.

أما بالنسبة للفقرات ٣١، ٣٢، ٤٧، فإن مصدر الغموض في استجابات المفحوصين عليها ربما يكون عدم تمكنهم من استيعاب المفاهيم المتضمنة فيها.

والجدير بالذكر أن الأخطاء التعليمية المتعلقة بالمعالجات الكمية والتطبيقات لم يظهر لها وجود في فقرات اختبار العلوم للصف الثامن الأساسي، إذ أن الفقرات من هذا النوع كانت تتمتع بخصائص جيدة. ولم يبلغ أثناء التطبيق عن أية ملاحظة تتعلق بإدارة عملية الاختبار.

مناقشة النتائج والتوصيات

من خلال استعراض نتائج هذه الدراسة ، يمكن ملاحظة الضعف الواضح في أداء طلبة الصف الثامن الأساسي في اختبار العلوم، إذ لم يرق متوسط أداء طلبة المملكة (٤٦٪) إلى مستوى التمكن (٨٠٪)، أو حتى مستوى النجاح التقليدي (٥٠٪)، الأمر الذي يستدعي ضرورة إمعان النظر في هذا الواقع. والغريب في الأمر، أن معدّل تحصيل الطلبة من مجموعة الأداء العالي (٦٥٪) لم يكن بمستوى متميز كذلك. وقد يكون مرد ذلك إلى جملة من الأسباب هي:

١. صعوبة بعض فقرات الاختبار وعدم تناسبها مع مستوى الطلبة المفحوصين في هذا الصف، أو
٢. خلل ما قد حصل في عملية المعاينة (Sampling)، أو
٣. ضعف واضح في مستوى التلاميذ، أو
٤. خلل ما يعود إلى جوانب العملية التعليمية عموماً (منهاج، كتاب مدرسي، إدارة، مرافق، طرق تدريس)، أو
٥. ضعف ما في عملية صياغة فقرات الاختبار حسب الأصول.

ومهما يكن من أمر، فإن التعمق في تحليل خصائص فقرات الاختبار ومسح الأخطاء التعلمية التي أظهرها أداء المفحوصين على الفقرات قد ساعد في ترجيح السبب وراء هذه النتيجة.

وللموضوعية، فإنه يمكن استبعاد احتمال الخلل في المعاينة كسبب من أسباب ضعف الأداء ومصدر رئيسي للأخطاء التعلمية، فقد تم الرجوع إلى آلية الاختيار المتبعة في تحديد عينة الدراسة، فكانت طريقة عشوائية طبقية راعت عوامل عدة مثل:

١. جنس الطلبة وجنس المدرسة.
٢. موقع المدرسة (مدن، ريف).
٣. المحافظة التي تقع فيها المدرسة.
٤. السلطة المشرفة على المدرسة (وزارة التربية والتعليم، خاصة، وكالة غوث، ثقافة عسكرية).

وبشكل عام، فإنه يمكن إبراز الاستنتاجات التالية:

* تألف اختبار العلوم للصف الثامن الأساسي من ٦٠ فقرة، وبزمن ٦٠ دقيقة متاحة للإجابة. إنه اختبار طويل، وقد لايتناسب مع عمر المفحوصين، إذ أن ضعف أداء الطلبة كان واضحاً على الفقرات الواقعة في النصف الثاني من الاختبار، وهذا واضح من خلال المؤشرات السيكومترية لخصائص الفقرات (صعوبة، تمييز، جاذبية مموهات).

وجدير بالذكر أن مسألة تحديد الوقت الفاعل لأي اختبار تعد مهمّةً جوهرية في قضايا تطوير الاختبارات. فقد أشار ننلي (Nunally, 1978) إلى أن الوقت الفعّال (Comfortable time limit)، الواجب تحديده لأي اختبار (غير اختبار السرعة) هو ذلك الزمن الذي يحتاجه ٩٠٪ من المفحوصين لإتمام الإجابة على كافة فقرات الاختبار. إضافة إلى أن التهاون أو التشدد المتطرّف في مسألة زمن الاختبار ربما يؤدي إلى الخروج بخصائص سيكومترية غير حقيقية لفقرات الاختبار.

* بينت مؤشرات نتائج هذه الدراسة عدم توزيع موقع الإجابة الصحيحة على بدائل الفقرات بطريقة عشوائية متكافئة، إذ أنه والحالة هذه، سيحصل المفحوص على علامة ما باستخدام التخمين غير العشوائي، على أن عشوائية التخمين هي افتراض أساسي من افتراضات آلية الإجابة على اختبارات الاختيار من متعدد.

* يعاني عدد من فقرات الاختبار ضعفاً في صياغتها، كما تم الإشارة إلى ذلك في محتوى نتائج الدراسة، الأمر الذي ساهم في زيادة إمكانية ارتكاب المفحوصين لبعض الأخطاء التعلّمية بسبب هذا الضعف الوارد في بعض الفقرات.

* إن نسبة غير قليلة من الأخطاء التعلّمية التي وقع فيها المفحوصون من طلبة الصف الثامن الأساسي كانت تصنف على أنها أخطاء تتعلق بمدى قدرة التلاميذ على استيعاب وتطبيق المفاهيم الواردة في المحتوى الدراسي، وربما يمكن إرجاع ذلك إلى طرق التدريس وإخراج

الكتاب المدرسي، إذ يبدو أنه لم تعط بعض المفاهيم وقتاً أو اهتماماً كافياً من التدريس الفاعل.

* تكاد تتفق جميع جوانب تحليل الفقرات لمستوياتها المختلفة حول ضعف الأداء والصياغة الذي كانت تعانيه فقرات معينة من الاختبار، الأمر الذي يؤيد مدى ثبات طرق التحليل.

* بدى واضحاً أن عدداً غير قليلٍ من الفقرات التي تنتمي إلى وحدة الكائنات الحية ومقاومة الأمراض تعاني من ضعف المؤشرات السيكومترية. وعند الرجوع إلى موقعها في الكتاب المدرسي، لم يكن ذلك مدعاة للتساؤل، إذ أنها كانت في بداية الفصل الدراسي الثاني. وهي فقرات تخلو من عوامل الضعف في الصياغة. الأمر الذي أدى إلى الاستنتاج بأن خللاً في تدريسها قد حصل، ربما كان قلة استخدام الوسائل التعليمية التوضيحية اللازمة في تدريس المفاهيم المتضمنة.

توصي هذه الدراسة بضرورة حذف أو تطوير الفقرات التي لا تتمتع بقدرة تمييز جيدة، وتلك الفقرات التي تتضمن مموهات غير جذابة للطلبة ضعيفي الأداء. كما توصي بأن يتناسب حجم الاختبار مع المستوى العمري للطلبة في الصف الثامن الأساسي، بحيث لا يزيد عدد فقرات الاختبار عن ٣٥ فقرة. وبشكل دقيق، فإنه حتى لو تم حذف الفقرات الواردة في الجدول رقم ٩، والجدول رقم ١٠، فإن الاختبار سيصبح أداة أكثر فاعلية وموضوعية في تقدير مستوى أداء الطلبة. وإن عقد دورة توجيهية لمشرفي العلوم من أجل تعريفهم بأنواع المفاهيم الواجب التركيز عليها في التدريس يعد أمراً حيوياً.

المراجع

1. Allen, M. J and Yen, W. M. (1979). Introduction to measurement theory. Monterey, CA. Brooks/Cole Publishing Company.

2. Brown, F. G. (1983). Principles or educational and psychological testing. NewYork: Holt, Rinehart, and Winston.

3. Crocker, L. and Algina, J. (1986). Introduction to classical and modern test theory. New York. Holt, Rinehart, and Winston.

4. Gronlund, N. E. (1988). Measurement and evaluation in teaching. (4th edition). New York: Macmillan Publishing Co.

5. Lee, A. E (1969). Testing and Evaluation student success with BSCS Laboratory Block. DC. Heath and Company Lexington, Massachusetts.

6. Lord, F. M. and Novick, M. R. (1968). Statistical theories of mental test scores. Reading Mass: Addison-Wesley.

7. Magnuson, D. (1967). Test theory. Boston: Addison-Wesley.

8. Mehren, W. A. and Lehman, I. J. (1978). Measurement and Evaluation in Education and psychlology. New York. Holt, Rinhart and Winston.

9. Nitko, A. J. (1989). Designing Tests that are Integrated with Instruction in Linn, R. L. (Eds), Educational Measurement. New York. American Ccouncil on Education and Macmillan Publishing Company.

10. Nitko, A. J. (1983). Educational Tests and measurement: An introduction. New York: Harcourt Brace Jovanovich, Inc.

11. Nunally, H. C. (1978). Psychometric theory. New York. McGraw-Hill.

12. Sax, G. (1974). Principles of educational measurement and evaluation. Belmont, Calif: Wadsworth Publishing Company Inc.