

الباب الخامس

خاتمة

أ. الاستنتاج

ملخص نتائج البحث حول تأثير اختبار الصدق، والثبات وقوة التمييز ومستوى الصعوبة لأسئلة تقييم نهاية الفصل على نتائج تعلم الطلاب في الصف الثامن بمدرسة PUI المتوسطة الإسلامية كاستوري ماجالينجكا بناءً على نتائج البحث، يمكن استخلاص عدة نقاط رئيسية كما يلي:

١. مستوى الصلاحية والثبات لأسئلة التقييم النهائي. أظهرت نتائج اختبار الصلاحية أن ٥٣٪ من أسئلة الاختيار من متعدد صالحة، بينما ٤٧٪ غير صالحة. بالنسبة للأسئلة المقالية، تبين أن ٩٠٪ منها صالحة و ١٠٪ غير صالحة. تشير هذه النتائج إلى أن الأسئلة الصالحة توفر قياسًا أكثر دقة للمهارات المتوقعة من الطلاب.

كما أظهر اختبار الثبات أن أسئلة الاختيار من متعدد حصلت على معامل ثبات يبلغ ٠,٦٢٢، ويعتبر ضمن فئة "جيد نوعًا ما" رغم أنه أقل من الحد الأدنى المقبول للثبات (٠,٧٠). بينما كانت درجة الثبات للأسئلة المقالية أقل، إذ بلغت ٠,٢٨٩، مما يدل على أن أسئلة الاختيار

من متعدد أكثر ثباتًا في قياس نتائج تعلم الطلاب، بينما تحتاج الأسئلة المقالية إلى تحسينات إضافية.

٢. مستوى الصعوبة وقوة التمييز لأسئلة التقييم النهائي. وفقًا لنتائج اختبار مستوى الصعوبة، تم تصنيف أسئلة الاختيار من متعدد إلى ٤٧٪ أسئلة سهلة، و ٥٣٪ أسئلة متوسطة الصعوبة، بينما لم يكن هناك أي سؤال صعب. أما بالنسبة للأسئلة المقالية، فقد كانت ٨٠٪ منها سهلة و ٢٠٪ متوسطة الصعوبة، دون وجود أي أسئلة صعبة. يشير هذا إلى أن أسئلة التقييم النهائي كانت سهلة نسبيًا، ولم تقدم تحديًا كافيًا للطلاب ذوي القدرات العالية. أظهرت نتائج اختبار قوة التمييز أن ٢٠٪ من أسئلة الاختيار من متعدد كانت غير جيدة، و ٤٣٪ كانت مقبولة، و ٣٧٪ كانت جيدة. وبالنسبة للأسئلة المقالية، كانت ١٠٪ غير جيدة، و ٤٠٪ مقبولة، و ٥٠٪ جيدة، مما يدل على أن قوة التمييز في الأسئلة المقالية كانت أفضل من أسئلة الاختيار من متعدد.

٣. تأثير اختبارات الصلاحية والثبات وقوة التمييز ومستوى الصعوبة على نتائج التعلم. أظهر اختبار "T" المزدوج (Paired T-Test) وجود فرق معنوي في نتائج تعلم الطلاب قبل وبعد تحسين أدوات التقييم. حيث ارتفع متوسط درجات الطلاب من ٤٥ قبل تحسين الأسئلة إلى ٥٥ بعد إجراء اختبارات الصلاحية والثبات وقوة التمييز ومستوى الصعوبة. وبما أن قيمة

الدلالة بلغت ٠,٠٠٠ (أقل من ٠,٠٥)، تم رفض الفرضية الصفرية (H0) وقبول الفرضية البديلة (H1)، مما يؤكد وجود فرق معنوي بين نتائج تعلم الطلاب قبل وبعد تحسين الأسئلة.

ب. المقترحات والتوصيات

بناءً على النتائج التي تم ذكرها، ينبغي تقديم بعض التوصيات لتحسين جودة أسئلة الاختبار، دون أي نية للمساس بمشاعر أي طرف. التوصيات المقترحة هي كما يلي:

١. يُنصح بأن يولي المعلمون اهتماماً أكبر لجودة الأسئلة قبل استخدامها في الامتحانات، وذلك من خلال التأكد من مرورها باختبارات الصلاحية، والثبات، والتمييز، ومستوى الصعوبة. إن الأسئلة الصالحة والموثوقة ستوفر تقييماً أدق لقدرات الطلاب، مما يساعد المعلمين في تقييم تحقيق أهداف التعلم بشكل أدق.

٢. يُوصى بأن يقوم الباحثون المقبلون بتوسيع نطاق هذا البحث ليشمل عينات أكبر ويشمل صفوفاً أو مدارس أخرى، بهدف الحصول على نتائج أكثر شمولية. علاوة على ذلك، يمكن للباحثين دراسة عوامل أخرى تؤثر في نتائج التعلم، مثل دافعية الطلاب، وأساليب التدريس، أو بيئة التعلم.

٣. يُتوقع من الطلاب الاستعداد بشكل أفضل لخوض الامتحانات، خاصة بعد تحسين جودة الأسئلة. بالإضافة إلى ذلك، يُفضل أن يكون الطلاب أكثر نشاطاً في الأنشطة التعليمية اليومية لزيادة فهمهم للمادة الدراسية.
٤. نظراً لانخفاض مستوى الثبات في أسئلة المقال، يُنصح بتحسين وتعديل هذه الأسئلة لتحقيق نتائج تقييم أكثر اتساقاً. يمكن للمعلمين استخدام أساليب تقييم أكثر تنظيماً في أسئلة المقال، بحيث تعكس نتائج التقييم قدرات الطلاب بشكل شامل.

