صدق الاختبارات التحصيلية المدرسية في مرحلة التعليم قبل الجامعي. "دراسة تقييمية"

محمد حسين سعيد حسين
مدرس علم النفس التعليمي
كلية التربية جامعة بني سويف

ملخص:

هدفت الدراسة الحالية إلى الكشف عن صدق الامتحانات المدرسية الرسمية بمرحلة التعليم قبل الجامعي، وذلك من خلال التعرف على العلاقة بين نتائج الطلاب في هذه الامتحانات من ناحية، وبين كل من مستوى ذكائهم، وتقديرات المعلمين لمستوياتهم التحصيلية من ناحية أخرى. واشتملت عينة الدراسة على (450) طالباً وطالبةً بمرحلة التعليم قبل الجامعي بمحافظة بني سويف، ونتائجهم (نتائج 183 طالباً وطالبةً بالمرحلة الابتدائية، ونتائج 130 طالباً وطالبة بالمرحلة الإعدادية، ونتائج 137 طالباً وطالبة بالمرحلة الثانوية)، ومعلميهم (15معلماً). وتم استخدام المنهج الوصفي لملاءمته لطبيعة وأهداف الدراسة. واعتمدت الدراسة على الأدوات التالية: مقياس رافن للذكاء. تعريب "فؤاد أبوحطب" (1977)، واستمارة تقدير مستوى التلاميذ من وجهة نظر معلميهم. إعداد الباحث، ونتائج الطلاب في نهاية العام الدراسي 2008/2007م. وتم استخدام معامل ارتباط بيرسون ومعامل ارتباط الرتب لسبيرمان، كما تم استخدام اختبار "ت" لدلالة الفروق بين المتوسطات غير المرتبطة. وأظهرت نتائج الدراسة ما يلي: 1- تمتع نتائج امتحانات المرحلة الابتدائية (اللغة العربية، الحساب، العلوم، الدراسات الاجتماعية) والمجموع الكلي بدرجة مناسبة من الصدق (المحك، والمقارنة الطرفية)، بينما كانت المشكلة في مادة اللغة الإنجليزية حيث لم ترتبط درجات التلاميذ فيها بمستوى ذكائهم، أو بمستوى تقدير معلميهم لهم. 2- تمتع نتائج امتحانات المرحلة الإعدادية (اللغة العربية، الحساب، العلوم، الدراسات الاجتماعية، اللغة الإنجليزية) والمجموع الكلي بدرجة مناسبة من الصدق (المحك، والمقارنة الطرفية). 3- تمتع نتائج امتحانات المرحلة الثانوية في مادتي اللغة الإنجليزية والرياضيات فقط بدرجة مناسبة من الصدق (المحك، والمقارنة الطرفية)، بينما لم ترتبط درجات الطلاب في المواد الأخرى؛ اللغة العربية، واللغة الفرنسية، والكيمياء، وعلم النفس، والمجموع الكلي بمستوى ذكائهم أو مستوى تقدير معلميهم لهم.

صدق الاختبارات التحصيلية المدرسية في مرحلة التعليم قبل الجامعي. "دراسة تقييمية"

محمد حسين سعيد حسين
مدرس علم النفس التعليمي
كلية التربية جامعة بني سويف

مقدمة:

يعتبر التقويم التربوي واحداً من أهم المكونات الأساسية في المنظومة التعليمية، فهو نقطة البدء في تحسين وتطوير جميع جوانب تلك المنظومة من أهداف ومقررات وطرائق تدريس وكتب ووسائل معينة. فقياس وتقويم التحصيل تؤثر نتائجه تأثيراً جوهرياً في ممارسات العمل التربوي وفاعليته، وما يتعلق به من قرارات متعددة تمس حاضر الطلاب ومستقبلهم، لذلك حظي هذا المجال باهتمام علماء القياس والتقويم التربوي، (صلاح الدين محمود علام، 1995). ويمكن النظر إلي العملية التعليمية علي أنها منظومة مكونة من مجموعة من العناصر التي ترتبط فيما بينها ويؤثر بعضها في بعض، ويعد التقويم أحد أهم مكونات هذه المنظومة التي تضم أيضاً الأهداف التعليمية والمناهج وطرق التدريس، ومما لا شك فيه أن أي تعديل أو تطوير لأحد هذه المكونات، لابد وأن يؤثر ويتأثر بالمكونات الأخرى.

وتعد الامتحانات أو الاختبارات التحصيلية من أهم وسائل التقويم المستخدمة في مصر والعالم العربي في الوقت الحاضر، والتي بها يتم تقويم الطلاب في أية مؤسسة تعليمية، وعلى الرغم من أن التقويم حلقة من تلك المنظومة، إلا أنه يكاد يكون أكثر هذه الحلقات أثراً في المنظومة كلها. فطبيعة الامتحانات تحدد مسار عملية التعليم ووجهتها، فهي تحدد ما تم تعلمه بالفعل، وتكاد تكون صورة لكل ما تم كما تحدد مسار ما يتم، فالامتحانات هي مرآة النظام التعليمي كله بفلسفته وقيمه وأسسه وأصوله وأهدافه وأساليبه وممارساته ونواتجه، (فؤاد أبوحطب، 1983: 587). كما أصبحت مسألة الاهتمام بالتحصيل الدراسي في المؤسسات التعليمية قضية جوهرية شغلت ومازالت تشغل الكثير من علماء التربية وعلم النفس، فقد أُجريت حوله البحوث والدراسات وعقدت من أجله المؤتمرات والندوات، وذلك لما له من أهمية في حياة الطلاب ومن يحيطون بهم، على اعتبار أن رفع مستوى التحصيل الدراسي سيؤدي بالضرورة إلى رفع مستوى مخرجات التعلم وبالتالي رفع مستوى إنتاجية المجتمع وزيادة تقدمه.

فالتحصيل الدراسي ليس أحد نتائج العملية التربوية فحسب، بل هو من أبرز النتائج لهذه العملية، ويُنظر إليه على أنه معيار أساسي يمكن في ضوئه تحديد المستوى الأكاديمي للطلاب، والحكم على النتائج الكمية والكيفية للعملية التربوية، إضافة إلى ما تحدثه هذه العملية من آثار في تكوين شخصية الطلاب وتشكيلها، (سيدة حافظ محمد، 2005: 53). ويؤكد "عادل العدل" (1996) على أن

التحصيل الدراسي يهدف في المقام الأول إلى الحصول على معلومات وصفية تبين مدى استيعاب الطلاب لما تعلموه من خبرات في المواد الدراسية المقررة وكذلك مدى ما حصله الطلاب من محتوى هذه المواد، كما يهدف إلى التوصل لمعلومات عن مستوى الطالب بالنسبة للمجموعة التي ينتمي إليها، وقد يهدف إلى أبعد من ذلك وهو محاولة رسم الصورة النفسية للطالب.

ولذلك أصبحت الامتحانات في الوقت الحاضر هي الشغل الشاغل لكثير من القائمين علي شئون التربية، لما لها من دور فعَّال في العملية التربوية، حيث تتخذ نتائجها مقياساً للمفاضلة بين الطلاب، وأساساً للانتقال من صف لآخر، كما تمنح الشهادات بموجبها، وتختار الطلاب لمدرسة أو لمهنة معينة بناء علي نتائج هذه الامتحانات، (توما جورج خورى، 1991: 163-164). ولذا فقد كانت عملية التقويم ووسائله المستخدمة كالامتحانات والاختبارات أو غيرها مدخلاً أساسياً لتطوير التعليم، وإصلاح جوانبه المختلفة، وشعار العمل والحركة للمرحلة الراهنة والتي تتناول بالتغيير كافة جوانب العملية التعليمية، إذ أن في تطوير العملية التعليمية وعمليات التقويم علاجاً لكثير من المشكلات التربوية التي ارتبطت بنظم التقويم التقليدية مثل ظاهرة الدروس الخصوصية وظاهرة الغش، (محمود منسى وآخرون، 1994: 11).

والدرجة التحصيلية ليست درجة صماء كما يبدو للبعض، ولكنها تعكس أموراً حيوية كثيرة وهامة بالنسبة للمتعلم، ومن هنا كانت "الدرجة التحصيلية" بمثابة "الحقل التجريبي" للباحثين في مجال علم النفس التربوي، (زين بن حسن ردادي، 2002). ويرى "ايبل" ***Ebel*** أن أي قياس للتحصيل الدراسي لابد أن يكون ثابتاً وصادقاً وموضوعياً، حيث يسهم تفنين تلك المقاييس في حسن تفسير الدرجة التي يحصل عليها التلاميذ، (صفوت فرج، 1989: 110). فالصدق والثبات خاصتان ضروريتان وجديرتان بالاعتبار عند إعداد واستخدام اختبار التحصيل، فهما إذاً ضروريتان للقياس الفعال، ويجب أن يلم بهما كل مشتغل في الاختبارات، وعليه فإن فهم وإدراك هذين المفهومين ينمي المهارة المتطلبة في تصميم الاختبار، وتفسير نتائجه، (محمد رضا البغدادي، 1998: 243). ويرى خبراء التقويم أنه من الممكن تقنين تلك الامتحانات لإبقاء الأسئلة القادرة علي إظهار التمايز بين قدرات الطلاب علي استيعاب المادة المتعلمة، (رمزية الغريب، 1996: 80).

وانطلاقاً من أن عملية التقويم أحد مكونات المنظومة التعليمية والتي تسهم في تحسين وتطوير تلك المنظومة، ونظراً للطبيعة الاجتماعية للامتحانات المدرسية في العالم العربي بصفة عامة ومصر بصفة خاصة، حيث إنها المعيار الوحيد الذي يؤهل الطلاب للانتقال من مرحلة دراسية إلى مرحلة أخرى، أو الانتقال من صف لآخر، كانت فكرة هذه الدراسة لتقدير صدق الامتحانات المدرسية الرسمية، بحيث يمكن تحديد إلى أي مدى يمكن الاعتماد على نتائج هذه الامتحانات في القرارات التي تتخذ بشأن الطلاب، أو الاعتماد عليها في البحوث التربوية والنفسية؛ من حيث تصنيف وتوزيع

الطلاب إلى مجموعات مختلفة (متفوقين، عاديين، متأخرين دراسياً).

مشكلة الدراسة:

تعتبر الامتحانات أحد الوسائل المتعددة التي يقصد بها تبين مواطن القوة والضعف في التعليم أثناء سعيه لتحقيق أهدافه ووظائفه، غير أن الخطر في التعليم أن تتحول الأهداف فيه إلي شعارات تفقد محتواها باختيار وسائل لا تتفق مع قيمه واتجاهاته، ومن ذلك أن أنصار التربية التقليدية، وأيضاً التربية الحديثة يؤكدون أن تنمية العقل هو مجال التربية، إلا أنهم يختلفون بشأن طريقة تنميته، ونجد وراء هذا اختلاف في النظر التربوي اختلافاً في معنى التحصيل المدرسي وحدوده، وهو اختلاف ينعكس علي طبيعة الامتحانات ووظيفتها في عملية التعليم. فالامتحان هو المحك الأساسي الذي يتم به إصدار الحكم علي الطالب بمدى استحقاقه للنجاح أو الرسوب، ولاشك أن النجاح أمر يهم الطالب وأبويه ومعلميه، وليس في هذا خطأ، إلا أن المشكلة هي ما يحدث من مبالغة في هذا الاتجاه مما يؤدي إلي اعتبار نجاح الطالب في الامتحان هدفاً تتضاءل أو تختفي –أحياناً– أمامه الأهداف التربوية الرئيسية، (اللجنة الوطنية القطرية للتربية والثقافة والعلوم، 1984: 16).

وقبل استخدام أي اختبار يجب التحقق من صدقه، بمعنى هل هذا الاختبار أداة صالحة لقياس ما وضع لقياسه ولا شيء غيره؟ فاختبار في الرياضيات قد يقيس بجانب الرياضيات أشياء أخرى، وقد يركز علي جانب ويهمل آخر، ولذلك فإن الصدق من المفاهيم الهامة التي بواسطتها يمكن الحكم علي أي اختبار من حيث جودته وصلاحيته للاستخدام والاعتماد على نتائجه. وللصدق أهمية قصوى في بناء الاختبارات، وذلك بالكشف عن محتوياتها الداخلية، وفي الإفادة من تلك الاختبارات في الاختيار التعليمي والمهني، أي التنبؤ بمستويات الأفراد في حياتهم التعليمية والمهنية وفي ذلك توفير للجهد والمال والتدريب، حتى يطمئن كل فرد إلي أنه يعمل في الميدان الذي يتفق مع استعداداته ومواهبه، (مجدي أحمد محمد، 1996: 64). فقد يفيد الاختبار الصادق في اختيار الطالب ما هو مناسب له في الدراسة أو في العمل، فاختبار صادق في مادة ما (لغة عربية – لغة إنجليزية– رياضيات – علوم–..) قد تفيد درجته في توجيه الطالب في اختبار التخصص أو الكلية المناسبة، كما أن حصول الفرد على درجته في اختبار مهني صادق تمكن من اتخاذ قرار في ما إذا كان سيواصل العمل في هذه المهنة أو يتركها.

وأكد "تايلر" علي أن الاختبار الذي يلعب دوراً في عملية اتخاذ قرارات حيوية يجب أن يتم اختياره بدقة، فمع أن الاختبار يجب أن يكون ثابتاً، فإن الاعتبار المهم في الحقيقة هو الصدق، (علي حامد الثبيتى، 1996: 59). ومما لا شك فيه أن الامتحانات المدرسية الرسمية في أي مرحلة دراسية لها دور مهم في اتخاذ قرارات بشأن الطلاب، ومستقبلهم وتوجيههم، ومن ثم يجب أن تكون مثل هذه الامتحانات علي درجة عالية من الصدق بجانب ثباتها. ولذلك حظيت هذه الامتحانات باهتمام العديد

من الباحثين من حيث تحليل أسئلتها، في المراحل التعليمية المختلفة، وخلصت هذه الدراسات إلى أن أسئلة هذه الامتحانات تركز على الجانب المعرفي، وبصفة خاصة التذكر والفهم، وتهمل الجوانب الوجدانية والمهارية والمستويات العليا من التفكير، (أمينة كاظم وآخرون، 1989؛ فكرى ريان، 1994؛ محمد خيرى محمود، 1994؛ رضا درويش، 1995؛ نادية على مسعود، 1999؛ حسن جامع، 2001؛ حمدى البنا، 2001). كما توصل "عبدالوارث الرازحي" (2001) إلى أن الامتحانات في معظم الدول العربية تفتقر إلى الشمول والموضوعية، وتركز في معظمها على التذكر والحفظ مع إهمال المستويات العليا. وهذه الامتحانات بصورتها السابقة وُجّهت إليها كثير من الانتقادات الفنية التي تتعلق بمدى كفاءتها من حيث مدى ثباتها، وموضوعيتها، وصدقها (محمد حسين سعيد، 2001؛ محمد حسين سعيد، 2005).

وثمة علاقة قوية بين اختبارات التحصيل ونوع آخر من الاختبارات وهو اختبارات الذكاء، فقد أوضحت الدراسات أن معاملات الارتباط بين الاختبارات التحصيلية واختبارات الذكاء عالية، حتى أنه يمكن التنبؤ فيها بمستوى ذكاء الفرد، وقد وجد أن هذه الاختبارات إذا وضعت بدقة كان لها معامل ثبات مرتفع، كما أن تغطيتها لمنهج الدراسة في حد ذاته يبين مدى صدقها، والذي يمكن معرفته أيضاً من معاملات الارتباط بين درجات المعلمين في المواد المختلفة علي الرغم من أن درجات المعلمين تتحكم فيها العوامل الذاتية، غير أنهم هم المرجع الأول والأخير في الحكم علي تلاميذهم، (سعد جلال، 1985: 113-114).

وعلى الرغم من وجود العديد من الدراسات والتي ركزت على التحليل الكيفي لأسئلة الامتحانات المدرسية في أعوام دراسية مختلفة، إلا أنه لا توجد دراسة اهتمت ببحث صدق هذه الاختبارات بطريقة كمية موضوعية غير الدراسة التي قام بها الباحث (2001). بالإضافة إلى ذلك فقد لاحظ الباحث العديد من المشكلات التي تتعلق بشأن جودة الامتحانات المدرسية من حيث مدى صدقها والاعتماد على نتائجها في استخدام طلاب الدراسات العليا (ماجستير ودكتوراه) لها في بحوثهم التي تتناول فئات مختلفة من الطلاب مثل: ذوي صعوبات التعلم، وبطيئي التعلم، والمتفوقين دراسيا، والموهوبين؛ وهؤلاء الطلاب من الضروري الرجوع إلى مستواهم التحصيلي والذي تعبر عنه نتائج امتحاناتهم المدرسية، وهذه النتائج قد لا يثق فيها البعض فيقوم بإعداد اختبار تحصيلي في المادة الدراسية التي يركز عليها، وقد يثق البعض في مثل هذه الامتحانات فيعتمد على نتائجها في تصنيف الطلاب وانتقائهم. ومن هنا نبعت مشكلة الدراسة الحالية والتي تحاول التعرف على درجة صدق هذه الامتحانات المدرسية في مرحلة التعليم قبل الجامعي، بحيث يمكن الاعتماد عليها أو تجنبها، وإعداد اختبارات تحصيلية من قبل الباحثين.

واتساقاً مع ما تقدم يمكن تحديد مشكلة الدراسة الحالية في السؤال الرئيسي التالي: إلى أي مدى

تكون الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي (ابتدائي-إعدادي-ثانوي) صادقة؟.

وينبثق من هذا السؤل الرئيسي التساؤلات التالية:

1- هل توجد علاقة بين نتائج الطلاب في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي وبين درجاتهم في مقياس الذكاء كمحك؟
2- هل توجد علاقة بين نتائج الطلاب في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي وبين تقديرات المعلمين لمستوياتهم التحصيلية كمحك؟
3- هل توجد فروق بين مرتفعي (الإرباعي الأعلى) ومنخفضي (الإرباعي الأدنى) الذكاء في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي؟
4- هل توجد فروق في نتائج الطلاب في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي تبعا لتقديرات معلميهم (مرتفع، منخفض)؟

أهداف الدراسة:

تهدف الدراسة الحالية إلى الكشف عن مدى صدق الامتحانات المدرسية الرسمية، وذلك من خلال التعرف على العلاقة بين نتائج الطلاب في هذه الامتحانات من ناحية وبين كل من: مستوى ذكائهم، وتقديرات المعلمين لمستوياتهم التحصيلية من ناحية أخرى.

أهمية الدراسة:

ترجع أهمية الدراسة الحالية إلى أهمية الموضوع الذي تتناوله وهو الامتحانات المدرسية الرسمية التي تحتل مكانة هامة في المنظومة التربوية، فالعديد من القرارات التربوية التي يتم اتخاذها بشأن الطلاب مثل: انتقالهم من صف دراسي إلى آخر، أو الانتقال من مرحلة دراسية إلى أخرى، أو تصنيفهم إلى فئات مختلفة (متفوقين دراسيا، موهوبين، ذوي صعوبات في التعلم، بطيئي التعلم) تعتمد على نتائج الطلاب في هذه الامتحانات، وبالتالي فإنه من الضروري التعرف على إلى أي مدى يمكننا الثقة في هذه الامتحانات من خلال التعرف على درجة صدقها، الأمر الذي يتيح فيما بعد الاعتماد عليها، ومن ثم توفير الجهد والوقت للباحثين في البحوث التربوية والنفسية حيث يتم الاعتماد على هذه الامتحانات بدلاً من إعداد امتحانات أخرى يقوم بها هؤلاء الباحثين. ومن ثم قد تفيد نتائج هذه الدراسة في توعية المتخصصين في مجال القياس والتقويم بمواطن القوة والضعف في هذه الامتحانات، من خلال مؤشرات كمية وموضوعية من المتوقع الحصول عليها في هذه الدراسة.

كما تكتسب هذه الدراسة أهميتها من ندرة البحوث والدراسات العربية في هذا المجال حيث لم تكن

هناك دراسة عربية واحدة اهتمت بهذه المشكلة إلا دراسة قام بها الباحث (محمد حسين سعيد، 2001) وذلك في حدود علم الباحث. كما أن هذه الدراسة تعطي الفرصة لدراسات أخرى مشابهة لعينات أخرى ممثلة للقطاعات التعليمية المختلفة، ولسنوات دراسية أخرى بمحافظات أخرى.

مصطلحات الدراسة:

الاختبارات التحصيلية:

يشير "سعد جلال" (1985: 109) إلى أن الاختبارات التحصيلية هي اختبارات تحاول قياس ما حصله الطلاب فعلاً بعد أن درسوا برنامجاً معيناً. ويشير كل من "كمال دسوقي" (1988: 47)، و"جروسر وسبافورد" ***(Grosser & Spafford, 1995: 12)*** إلى أن التحصيل الدراسي هو مستوى محدد من الإنجاز أو الكفاءة أو الأداء في العمل المدرسي، أو في سلسلة اختبارات تربوية مقننة.

كما يشير "عبدالمجيد نشواتي" (1991: 601) إلي أن الاختبار التحصيلي نمط معين من أدوات القياس ووسائله، وينطوى علي مجموعة من الأسئلة أو العبارات أو المهام التعليمية، التي اختيرت وصيغت بطريقة منهجية معينة بحيث توفر –لدى إجابة الطالب عنها– قيمة رقمية لإحدى خصائصه المعرفية كالتحصيل والذكاء والابتكار، أو غير المعرفية كالخلقية الاجتماعية والاتجاهات والقيم والميول. بينما يركز "جابر عبدالحميد" (1994: 357) في تعريفه للاختبار علي طريقة الإجراء أو التطبيق نفسها فيوضح أن الاختبار إجراء نسقي لقياس عينه من سلوك شخص، لكي نقوَّم ونقدَّر ذلك السلوك في ضوء مستويات ومعايير . وهي نوعان اختبارات مرجعية المعيار ***Norm Referenced Tests***، واختبارات مرجعية المحك ***Criterion Referenced Tests***.

والاختبار التحصيلي هو الأداة التي تستخدم في قياس المعرفة والفهم والمهارة في مادة دراسية أو تدريبية معينة أو مجموعة من المواد، وهي أكثر الأدوات شيوعاً في التقويم التربوي وهي الوسيلة الوحيدة التي تستخدم في مصر في توجيه التلاميذ وانتقائهم، (محمد السيد حسونة، 1995: 18). وتشير "أمينة كاظم" (1996: 438) إلى أن التحصيل الدراسي يمكن تعريفه بواسطة بنود مناسبة متدرجة الصعوبة على تدرج متصل تكون فيما بينها اختباراً مناسباً لقياس أهداف مقرر ما، ويمكن التعرف على تقدير مستوى قدرة الأفراد على متغير التحصيل الدراسي عن طريق استجاباتهم على بنود اختبار متدرجة في الصعوبة حتى تكون مناسبة لجميع المستويات وتساعد في الكشف عن الفروق الفردية بين الطلاب.

كما يشير "سعد عبدالرحمن" (1998: 9) إلى أن الاختبار التحصيلي هو مجموعة من الأسئلة أو البنود لكل منها إجابة واحدة صحيحة فقط، وتختلف عن الاستفتاءات في أن الأخيرة ليس لها إجابات صحيحة أو إجابات خاطئة، إذ أن المطلوب هو معرفة رأي الفرد أو نوعية استجابته في موقف من المواقف التي يمثلها ذلك البند أو السؤال. كما تعرفه "عفاف محمد عبدالمنعم" (1999: 54) بأنه

القدرة على الاستدعاء للاستيعاب والمقدر كمياً بالدرجة الكلية لجميع المواد التي تم دراستها، ويتضح من خلال الاختبارات المدرسية المقننة أو غير ذلك من وسائل التقويم.

والاختبارات التحصيلية تقيس مدى استيعاب الطلاب لبعض المعارف والمفاهيم والمهارات المتعلقة بالمادة الدراسية، كما يدل التحصيل الدراسي على الوضع الراهن لأداء الفرد أو تعلمه أو ما اكتسبه بالفعل في برنامج تعليمى، (صلاح علام، 2000: 305– 306). ويوضح "فتحي الزيات" (2001: 475)، و"محسن محمد أحمد عبدالنبي" (2001) أن التحصيل الدراسي هو مقدار المعارف والمهارات التي يحصلها الطالب من خلال الموضوعات الدراسية أثناء العام الدراسي، والتي يمكن قياسها عن طريق الاختبارات المدرسية التي تجرى خلال العام الدراسي ونهايته، ويعبر عنها بالدرجات التي يحصل عليها الطالب في هذه الاختبارات.

كما تقيس الاختبارات التحصيلية المدرسية مدى تحقيق الأهداف التعليمية، وهى إما اختبارات عادية تعد بواسطة المعلمين، أو اختبارات عامة تعدها الوزارة في نهاية كل مرحلة دراسية، وبخاصة الثانوية العامة، (عبدالوارث الرازحى، 2001 : 110). كما ترى "عبلة محمد الجابر" (2002: 16) و"عاطف عبدالعزيز عبدالمقصود" (2003)، أن التحصيل الدراسي مؤشر لإتقان المعارف والمعلومات ويتم اكتسابه عن طريق التدريب على مهارة معينة أو تحصيل لمقرر ما، ويقدر الأداء تقديراً كمياً.

ومن خلال التعريفات السابقة يعرف الباحث الاختبار التحصيلي على أنه طريقة منظمة لقياس الجوانب المعرفية للطلاب بعد دراستهم لمجموعة من المواد، وذلك عن طريق التقدير الكمي لأدائهم علي مجموعة من الأسئلة تمثل الجوانب المعرفية التي يجب أن يكتسبها الطالب بعد دراسته لتلك المواد الدراسية. وفي ضوء ذلك يمكن النظر إلى التحصيل الدراسي على أنه مؤشر للمستوى الذي يصل إليه الطالب تحت ظروف الممارسة والتدريب التي تتيحها المدرسة، ويُعبر عنه بالدرجات التي يحصل عليها الطالب في الاختبارات التحصيلية العادية التي تعدها المدرسة أو الوزارة.

الصدق:

لا يكون الاختبار صادقًا إلا إذا كان قادرًا علي قياس ما وضع لقياسه، ولا يقيس شيئاً غيره، ***(Angoff, 1988: 19; Cronbach, 1988: 5; Gronlund & Linn, 1990: 79; Goldstein, 1996: 88; Taylor, 1999: 84)*** وكذلك أن يكون قادرًا علي التمييز بين طرفي القدرة التي يقيسها، أي التمييز بين الأداء القوى والأداء الضعيف، (سعد عبدالرحمن، 1998: 183). ويعرف "إيبل وفرسبي" ***(Ebel & Frisbie, 1986: 37)*** الصدق بأنه إلى أي مدى يمكن أن تفيد الدرجة الناتجة من الاختبار في اتخاذ قرارات أو الوصول لاستنتاجات ودلائل صحيحة حول الطالب. ***the appropriateness of making specific inferences and of making certain decisions on the basis of scores from a test (Messick, 1989)***، ويؤكد ذلك "ميسك"

بقوله "إن الاختبار الصادق هو الذي نستطيع التوصل من خلاله إلى استنتاجات أو دلائل صحيحة حول الطالب وليس حول الاختبار نفسه". فالاختبار الصادق هو ذلك الاختبار الذي تدعم نتيجته ما يمكن قوله أو اتخاذه بشأن طالب ما ***to what extent the result of the test supports what we are trying to say about the person*** ***(Wiliam, 1996)***.

وإذا أجريت عمليات الصدق واشتقت من مجموعة أو مجموعات معينة، فإن الصدق مرهون بخصائص هذه المجموعات والمجتمع الذي تمثله، ولذلك فالاختبار الذي ثبت صدقه في مجتمع أو ثقافة معينة، قد لا يصدق في مجتمع أو ثقافة أخرى، وبهذه الصورة فإن صدق الاختبار في قياس ما وضع من أجله يكون بالنسبة لخاصيتين: الأولى: قياس السمة المراد دراستها، أو الوظيفة التي يقيسها، والثانية: طبيعة العينة أو المجتمع المراد دراسة السمة كعينة مميزة لأفراده، (مجدى أحمد محمد، 1996: 93). والمقصود بأن الصدق مرهون بخصائص المجتمع أنه قد توجد بعض العبارات أو بعض الكلمات في الاختبارات، ويختلف معناها من مجتمع لآخر، وبالتالي تختلف إجابات الأفراد عليها، ومن ثم فهذه الاختبارات التي تحتوى مثل هذه العبارات أو الكلمات تكون صادقة لمجتمع، وغير صادقة في مجتمع آخر.

ويُعرف الصدق بالنسبة لنسبة التباين الحقيقي المرتبط بأهداف القياس، والذي يُعبّر عنه إحصائياً كما يلي: $ع^2$ س = $ع^2$ت + $ع^2$خ : $ع^2$ت = $ع^2$ص + $ع^2$ط حيث: $ع^2$س تباين الدرجات الفعلية، $ع^2$ت تباين الدرجات الحقيقية، $ع^2$خ تباين الخطأ $ع^2$ص التباين المرتبط بالصدق، $ع^2$ط التباين غير المرتبط بالصدق، وبالتالي فإن: $ع^2$ س = $ع^2$ص + $ع^2$ط + $ع^2$خ، (عبدالهادى السيد عبده، وفاروق عثمان، 1995: 237). أي أن تباين الدرجات الفعلية هو مجموع تباينات ثلاث هي: تباين الجزء الذي يرجع إلي صدق الاختبار، وتباين الجزء الذي يرجع إلي عدم صدق الاختبار، والذي يمكن اعتباره نوعاً من أنواع الخطأ، وتباين الجزء الذي يرجع إلي خطأ القياس، وبالتالي فإن العلاقة السابقة يمكن أن تأخذ الصورة التالية: $ع^2$س = $ع^2$ص + $ع^2$صَ : $ع^2$صَ = $ع^2$خ + $ع^2$ط (محمد حسين سعيد، 2001).

وباستعراض مفاهيم الصدق السابقة يتضح أنها تؤكد علي أن الاختبار الصادق هو الذي يقيس ما وضع لقياسه، ولا يقيس شيء غيره، فاختبار القدرة المكانية يجب أن يقيس القدرة المكانية، ولا يقيس معها قدرات أخرى، كأن يتأثر المصحح بتنظيم الإجابة أو جودة الخط وغيرهما، كذلك فالاختبار الصادق يستطيع أن يميز بين الطلاب ذوي القدرات المختلفة، فاختبار التحصيل الصادق يستطيع أن يميز بوضوح بين مرتفعي ومنخفضي التحصيل. ويعرف الصدق في هذه الدراسة إجرائياً بالقيمة الناتجة من الطرق المستخدمة في تقديره، ومن هذه الطرق:

1- صدق المحك:

يسمى في بعض الأحيان بالصدق الأمبريقي، والمقصود به إلى أي حد ترتبط درجات الاختبار المراد قياس صدقه بمحك خارجي، ويوجد نوعان من صدق المحك وهما:

أ- الصدق التلازمي:

يعتمد على حساب الارتباط بين درجات الأفراد على الاختبار، ودرجاتهم على محك آخر، بحيث لا يكون هناك فاصل زمني بين تطبيق الاختبار وبين الحصول على درجات الأفراد الراهنة كما يقيسها المحك.

ب- الصدق التنبؤي:

كثيرًا ما يكون الاهتمام هو استخدام الاختبار لأغراض التنبؤ بنتائج تقع في المستقبل، وهذا النوع من الصدق يعتمد على حساب معامل الارتباط بين درجات الأفراد على الاختبار ودرجاتهم عل محك آخر، بحيث يكون هناك فاصل زمني بين التطبيق على الاختبار والتطبيق على الاختبار المحك. والإجراء العملي المتبع هو أن يعطي الاختبار إلى مجموعة من الأفراد عند التحاقهم بالعمل أو البرنامج، ثم متابعتهم في المستقبل، ثم الحصول لكل فرد منهم على قياس في محك معين للنجاح في العمل أو البرنامج، ثم نحسب الارتباط بين الدرجات على الاختبار والدرجات على المحك، وكلما كان الارتباط مرتفعاً كلما كان الاختبار أكثر فاعلية في التنبؤ.

وسوف تركز الدراسة الحالية على النوع الأول من صدق المحك وهو الصدق التلازمي؛ حيث تستخدم مستوى ذكاء الطلاب كمحك، بالإضافة إلى تقديرات معلميهم لمستوياتهم التحصيلية.

2- صدق المقارنة الطرفية:

تقوم هذه الطريقة في جوهرها على مقارنة متوسط درجات الأقوياء في الميزان بمتوسط درجات الضعاف في نفس ذلك الميزان بالنسبة لتوزيع درجات الاختبار، ولذا سُميت بالمقارنة الطرفية لاعتمادها على الطرف الممتاز والطرف الضعيف للميزان. وعندما تدل نتائج الاختبار على أن الأقوياء في الميزان أقوياء في الاختبار وأن الضعاف في الميزان ضعاف في الاختبار

يصبح الاختبار صادقاً، ويزداد الصدق تبعاً لزيادة هذا الاقتران، ولذا تتضح الأهمية الطرفية لمستويات الميزان في هذه المقارنة.

ومن أبسط الطرق التي تستخدم لتحقيق هذه الفكرة مقارنة متوسطات درجات الأقوياء بمتوسطات درجات الضعاف، ثم حساب دلالة الفروق بين هذه المتوسطات، وعندما تصبح لتلك الفروق دلالة إحصائية واضحة يمكن تقرير أن الاختبار يميز بين الأقوياء والضعاف في الميزان، وبذلك يمكن الاطمئنان إلى صدقه، وعندما لا تصبح لتلك الفروق دلالة إحصائية واضحة، فإنه لا يمكن الاطمئنان إلى صدق مثل هذا الاختبار، (فؤاد البهي السيد، 1978: 405-409).

وسوف تستخدم الدراسة الحالية هذا النوع من الصدق؛ حيث يتم المقارنة بين مرتفعي ومنخفضي الذكاء، ومرتفعي ومنخفضي الأداء من وجهة نظر معلميهم في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي.

الدراسات السابقة:

1- دراسة "هاينز" *(Hines, 1981)*:

بحثت هذه الدراسة الصدق البنائي لست اختبارات تحصيلية مهنية تم استخدامها في برنامج الاختبارات التحصيلية المهنية بأوهايو. واشتملت مجالات هذا الاختبار على: الميكانيكا الزراعية، والتجارة، وميكانيكا الديزل، والخدمات الصناعية، واقتصاديات المنزل وخدمة الطعام. وللتحقق من صدق هذه الاختبارات، تم تطبيقهم على 4627 طالباً وطالبة في الصفوف العاشر والحادي عشر والثاني عشر بمقاطعة أوهايو. وقام الباحثون بتقييم الصدق البنائي لكل اختبار من خلال التحليل المقارن لمستويات أداء المجموعات على الست اختبارات. وباستخدام تصميم تحيل التباين ثنائي العامل لبحث الاختلافات الجماعية في مستويات الأداء على كل اختبار، أظهرت نتائج الدراسة أن هذه الاختبارات تعد بمثابة أدوات تقييمية يمكن استخدامها من قبل المعلمين والمديرين كجزء من تحليل المنهج والنظام.

2- دراسة "هوبكنز وآخرون" *(Hopkins & Others, 1985)*:

هدفت هذه الدراسة إلى تقييم الصدق التلازمي لاختبارات تحصيلية مقننة باستخدام تقديرات المعلمين للتحصيل الأكاديمي للطلاب كمحك، حيث قام 42 معلماً بتقييم كل طالب من طلابهم (ن = 1032) في القراءة، والحساب، وفنون اللغة، والعلوم، والدراسات الاجتماعية بالمرحلة الإعدادية. وتم توجيه المعلمين إلى تقدير كفاءة كل طالب بغض النظر عن الحضور أو الاتجاه أو غيرها من العوامل. وأكدت نتائج الدراسة أن الاختبارات التحصيلية المقننة تحظى بمستوى صدق مناسب في كل من: القراءة، والحساب، وفنون اللغة، والعلوم، والدراسات الاجتماعية.

3- دراسة "نولت وتيندال" *(Nolet & Tindal, 1990)*:

هدفت هذه الدراسة إلى تحليل البيانات المقدمة من قبل ناشري الاختبارات التحصيلية ***Achievement Test Publishers*** لبحث وجود مجموعة من الأدلة التي تؤيد الاستدلالات المختلفة حول معنى الدرجات التي يتم الحصول عليها من هذه الاختبارات. وركزت الدراسة بشكل أساسي على بحث الجوانب المتعددة للصدق البنائي لبطاريات الاختبارات التحصيلية مرجعية المحك المنشورة. وتم مراجعة اختبارات تحصيلية مطبقة فردياً وأخرى جماعياً. واشتملت المواد المفحوصة لكل اختبار كل من استجابات التلاميذ، والبروتوكولات المدرجة، وقوائم التطبيق، وقوائم الأهداف. وأظهرت نتائج الدراسة أن بطاريات الاختبارات التحصيلية تعد بمثابة مقاييس ملائمة للتحصيل بشكل عام في القراءة والرياضيات والتعبير اللغوي، على الرغم من أن أداء التلاميذ في الجوانب المهارية المقدم بواسطة العديد من الاختبارات الفرعية المشتملة في معظم البطاريات التحصيلية غير صادق، كما أكدت نتائج الدراسة أن بطاريات الاختبارات التحصيلية المنشورة موضع الدراسة كان لها صدق تنبؤي، ولكن ليس لها صدق تمييزي، كما أن هذه الاختبارات قد فشلت في تمثيل عدد كبير من السلوكيات الصفية.

4- دراسة "سكوت" *(Scott, 2000)*:

هدفت هذه الدراسة إلى بحث ما إذا كانت درجات التلاميذ على الاختبارات التحصيلية المقننة تمثل مقاييس معتمدة للجودة التدريسية. وباستخدام بروتوكول البحث المصمم بواسطة بوبهام ***Popham*** وموجهي الدراسة المحلية تم تقدير مفردات الاختبارات الفردية من اختبار تحصيلي مقنن متداول محليا، وتم هذا التقدير بواسطة المعلمين والوالدين لتحديد الدرجة التي إليها يشعر المقدرون أن المفردات تعكس محتوى هام تم تدريسه بالفعل في المدارس. وأظهرت نتائج الدراسة أن في المتوسط شعر القائمون على التقدير أن المحتوى المنعكس في أسئلة الاختبار يقيس مادة تكون هامة بالنسبة للتلاميذ أن يتعلموها.

5- دراسة "محمد حسين سعيد" (2001):

هدفت هذه الدراسة إلى بحث الخصائص السيكومترية (الصدق-الثبات-التحيز-ثبات درجة القطع) لنتائج امتحانات الثانوية العامة في عامي 1991، و1999، وتم الحصول على نتائج الطلاب في هذين العامين وتحليلها، والتعرف على مستوى الصدق والثبات لها، بالإضافة إلى مدى تحيزها ***Biased***، وأظهرت نتائج الدراسة إلى انخفاض مستوى الصدق والثبات لهذه الامتحانات في هذين العامين، بالإضافة إلى التحيز الواضح لها، ومن ثم فإنها امتحانات غير عادلة.

6- دراسة "بيل وآخرون" *(Bell & Others, 2002)*:

هدفت هذه الدراسة إلى بحث الصدق التلازمي لاختبار ذكاء ***Solsson Full-Range***

Intelligence Test (S-FRIT) من خلال مقارنة درجات هذا الاختبار بدرجات كل من اختبار وكسلر للذكاء الطبعة الثالثة ***Wechsler Intelligence Scale For Children Third Edition (WISC-III)***، واختبارات ووكودك جونسون التحصيلية المعدلة ***Woodcock-Johnson Tests of Achievement – Revised (WJ-R)*** . وأظهرت نتائج الدراسة أن درجات اختبار الذكاء ***(S-FRIT)*** كانت أكثر ارتباطا بالذكاء العام، والقدرات اللفظية والرياضية من الذكاء غير اللفظي، والقراءة، أو قدرات اللغة التحريرية (المكتوبة).

7- دراسة "أبيلا وآخرون" *(Abella & Others, 2005)*:

هدفت هذه الدراسة إلى اختبار 1700 متعلم للغة الإنجليزية وتلاميذ متعلمين سابقاً للغة الإنجليزية في الصفوف 4 ، 10 باستخدام كل من الاختبارات التحصيلية للغة الإنجليزية والاختبارات التحصيلية للغة الأسبانية، وتم مقارنة أدائهم على الاختبارين. وأظهرت نتائج الدراسة أن التلاميذ المتعلمين للغة الإنجليزية استطاعوا الإجابة على معظم المفردات على اختبار رياضيات باللغة الأم–***Home Language Mathematics Test*** مقارنة باختبار رياضيات باللغة الإنجليزية، بغض النظر عن مستواهم في قراءة اللغة الأم. هذا بالإضافة إلى أن التلاميذ متعلمي اللغة الإنجليزية السابقين كانوا غير قادرين على إظهار معرفتهم بجوانب المحتوى في الاختبارات التحصيلية للغة الإنجليزية، وأظهرت نتائج الدراسة أن نتائج الاختبارات التحصيلية للطلاب ليست دائما مقاييس صادقة لمعرفتهم بجوانب المحتوى.

تعقيب على الدراسات السابقة:

يتضح مما سبق أن الاهتمام بصدق الامتحانات التحصيلية كان على مستوى الدراسات الأجنبية فيما عدا الدراسة العربية التي قام بها الباحث (2001)، كما يلاحظ أن هذه الدراسات في تقديرها للصدق اعتمدت على اختبارات الذكاء، كما في دراسة "بيل وآخرون" ***(Bell & Others, 2002)***، أو تقديرات المعلمين لمستويات طلابهم التحصيلية، كما في دراسة "هوبكنز وآخرون" ***(Hopkins & Others, 1985)***، ويلاحظ أيضاً أن القليل من هذه الدراسات أظهرت تمتع هذه الامتحانات بمستوى صدق مناسب، كما في دراسة "بيل وآخرون" ***(Bell & Others, 2002)***، في حين أن العديد منها توصل لعكس ذلك كما في دراسة "تولت وتيندال" ***(Nolet & Tindal, 1990)*** ودراسة "محمد حسين سعيد" (2001)، ودراسة "أبيلا وآخرون" ***(Abella & Others, 2005)***.

فروض الدراسة:

الفرض الأول:

توجد علاقة دالة وموجبة عند مستوى دلالة 0.05 بين درجات الطلاب في الاختبارات التحصيلية المدرسية الرسمية في مرحلة التعليم قبل الجامعي ودرجاتهم في مقياس الذكاء.

الفرض الثاني:

توجد علاقة دالة وموجبة عند مستوى دلالة 0.05 بين درجات الطلاب في الاختبارات التحصيلية المدرسية الرسمية في مرحلة التعليم قبل الجامعي وتقديرات المعلمين لمستوياتهم التحصيلية.

الفرض الثالث:

توجد فروق ذات دلالة إحصائية عند مستوى دلالة 0.05 بين متوسطي مرتفعي (الإرباعي الأعلى) ومنخفضي (الإرباعي الأدنى) الذكاء في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي.

الفرض الرابع:

توجد فروق ذات دلالة إحصائية عند مستوى دلالة 0.05 في نتائج الطلاب في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي تبعا لتقديرات معلميهم (مرتفع، منخفض).

إجراءات الدراسة:

أولاً: منهج الدراسة:

تم استخدام المنهج الوصفي، حيث إنه المناسب لطبيعة وأهداف الدراسة، والتي تحاول التعرف على مستوى صدق الاختبارات التحصيلية المدرسية الرسمية في مرحلة التعليم قبل الجامعي.

ثانياً: عينة الدراسة:

اشتملت عينة الدراسة على (450) طالباً وطالبةً، ونتائجهم بمرحلة التعليم قبل الجامعي بمحافظة بني سويف في عام 2008/2007م على النحو التالي:

- نتائج (183) طالباً وطالبةً بالصف السادس الابتدائي،
- ونتائج (130) طالباً وطالبةً بالصف الثالث الإعدادي،
- ونتائج (137) طالباً وطالبةً بالصف الثاني الثانوي (المرحلة الأولى).

بالإضافة إلى معلمي هؤلاء الطلاب والذين بلغ عددهم (15) معلماً ومعلمةً بواقع (5) معلمين لكل مرحلة دراسية.

ثالثاً: أدوات الدراسة: اعتمدت الدراسة في تحقيق أهدافها على الأدوات التالية:

1- مقياس رافن للذكاء:

أعد هذا المقياس في الأصل رافن وقام "فؤاد أبوحطب" (1977) بتقنينه على البيئة السعودية، والمقياس يتكون من (5) مجموعات هي (أ ، ب ، ج ، د ، هـ) وكل مجموعة من المجموعات السابقة تتكون من (12) مفردة، ومن ثم فإن عدد مفردات المقياس الكلية هي (60) مفردة. وتتابع المجموعات الخمس حسب درجة الصعوبة، وكل مفردة عبارة عن رسم أو تصميم هندسي أو نمط شكلي، حذف منه جزء وعلى المفحوص أن يختار الجزء الناقص من بين (6) أو (8) بدائل معطاه.

صدق المقياس:

قام مُعد المقياس بتقدير الصدق المرتبط بالمحك، حيث بلغت قيمة معامل الارتباط بين هذا المقياس ومقياس ذكاء الشباب اللفظي لحامد زهران 0.73، وذكاء الشباب المصور لحامد زهران، 0.78 وجميعها دال عن مستوى دلالة 0.01، (فؤاد أبوحطب، 1977: 235)، وتشابهت قيم معاملات الصدق التي توصل إليها "أحمد عثمان" (1988) في البيئة المصرية (ن=7500 طالباً وطالبةً بجميع المراحل التعليمية)، مع ما توصل إليه "فؤاد أبوحطب" (1977) على البيئة السعودية.

وفي الدراسة الحالية تم استخدام طريقة صدق المحك (التلازمي) حيث بلغت قيمة معامل الارتباط بين هذا المقياس ومقياس الذكاء المصور لأحمد زكي صالح (0.87)، و(0.86)، و(0.83) بالنسبة لتلاميذ المرحلة الابتدائية (ن=73)، والمرحلة الإعدادية (ن=65)، والمرحلة الثانوية (ن=45)، على الترتيب، وجميع هذه القيم دالة عند مستوى دلالة إحصائية 0.05.

ثبات المقياس:

قام مُعد المقياس بتقدير ثبات المقياس بطريقة كيودر ريتشاردسون (20) على عينات في أعمار مختلفة (8–30 +)، وقد تراوحت قيم معاملات الثبات بهذه الطريقة ما بين 0.87 إلى 0.95، كما تراوحت قيم معاملات الثبات بطريقة إعادة التطبيق ما بين 0.46 إلى 0.86 (فؤاد أبوحطب، 1977: 236–237). وتشابهت قيم معاملات الثبات التي توصل إليها "أحمد عثمان" (1988) في البيئة المصرية (ن=7500 طالباً وطالبةً بجميع المراحل التعليمية)، مع ما توصل إليه "فؤاد أبوحطب" (1977) على البيئة السعودية.

وفي الدراسة الحالية تم استخدام طريقة ألفا لكرونباخ، حيث بلغت قيم معاملات الثبات بهذه الطريقة (0.79)، (0.83)، (0.81)، (0.80)، (0.81) لتلاميذ المرحلة الابتدائية، و(0.81)، (0.81)، (0.84)، (0.80)، (0.80) لطلاب المرحلة الإعدادية، و(0.81)، (0.80)، (0.80)، (0.80)، (0.80)، (0.83) لطلاب المرحلة الثانوية وجميعها قيم مناسبة.

2– استمارة تقدير مستوى التلاميذ التحصيلية، إعداد الباحث (ملحق1). وهي عبارة عن استمارة تقدير يطلب من المعلم فيها أن يقدر مستوى كل طالب من طلابه في ضوء تقديرات ست هي: ممتاز، جيد جداً، جيد، متوسط، ضعيف، ضعيف جداً، وهذه الاختيارات تأخذ التقديرات: (6)، و(5)،

و(4)، و(3)، و(2)، و(1) على الترتيب.

3- نتائج الطلاب في مرحلة التعليم قبل الجامعي (الصف السادس الابتدائي، الصف الثالث الإعدادي، الصف الثاني الثانوي) في العام الدراسي 2008/2007م.

المعالجة الإحصائية لبيانات الدراسة:

للإجابة عن أسئلة الدراسة، واختبار صحة الفروض، تم استخدام الأساليب الإحصائية التالية:

- معامل ارتباط بيرسون،
- معامل ارتباط الرتب لسبيرمان،
- اختبار "ت" لدلالة الفروق بين المتوسطات غير المرتبطة.

وقد تمت جميع المعالجات الإحصائية باستخدام البرنامج الإحصائي ***SPSS (12)***.

حدود الدراسة:

تتحدد نتائج الدراسة الحالية بعينة الدراسة والتي اشتملت على 450 طالباً وطالبةً، ونتائجهم، ومعلميهم (ن=15) بمرحلة التعليم قبل الجامعي بمحافظة بني سويف (نتائج 183 طالباً وطالبةً بالمرحلة الابتدائية، ونتائج 130 طالباً وطالبةً بالمرحلة الإعدادية، ونتائج 137 طالباً وطالبةً بالمرحلة الثانوية)، كما تتحدد بالأدوات المستخدمة فيها، والتي تمثلت في اختبار رافن للذكاء، واستمارة تقدير مستوى الطلاب، وكذلك بالأساليب الإحصائية المستخدمة في معالجة بيانات الدراسة.

نتائج الدراسة:

1- للتحقق من صحة الفرض الأول للدراسة والذي ينص على: توجد علاقة دالة وموجبة عند مستوى دلالة 0.05 بين درجات الطلاب في الاختبارات التحصيلية المدرسية الرسمية في مرحلة التعليم قبل الجامعي، ودرجاتهم في مقياس الذكاء. والفرض الثاني للدراسة والذي ينص على: توجد علاقة دالة وموجبة عند مستوى دلالة 0.05 بين درجات الطلاب في الاختبارات التحصيلية المدرسية الرسمية في مرحلة التعليم قبل الجامعي، وتقديرات المعلمين لمستوياتهم التحصيلية. تم استخدام معامل ارتباط بيرسون، ومعامل ارتباط الرتب لسبيرمان، وذلك على النحو التالي:

أولاً: بالنسبة للمرحلة الابتدائية:

جدول (1) قيم معاملات الارتباط بين درجات تلاميذ المرحلة الابتدائية في الاختبارات التحصيلية المدرسية ودرجاتهم في مقياس الذكاء وتقديرات معلميهم لهم

المادة	معامل الارتباط بالذكاء	معامل الارتباط بتقديرات المعلمين*
اللغة العربية	0.443**	0.709**
العلوم	0.381**	0.713**
الحساب	0.455**	0.607**
الدراسات الاجتماعية	0.477**	0.540**
اللغة الإنجليزية	0.171	0.189
المجموع الكلي	0.446**	0.713**a

**** دالة عند مستوى دلالة 0.01.** ***** دالة عند مستوى دلالة 0.05.**

يتضح من الجدول السابق ما يلي:

- وجود علاقة دالة وموجبة عند مستوى دلالة (0.05) بين درجات تلاميذ المرحلة الابتدائية في نتائج الاختبارات التحصيلية المدرسية لمواد: اللغة العربية، والعلوم، والحساب، والدراسات الاجتماعية، والمجموع الكلي، وبين درجاتهم في مقياس رافن للذكاء.
- عدم وجود علاقة دالة بين درجات تلاميذ المرحلة الابتدائية في مادة اللغة الإنجليزية ودرجاتهم في مقياس رافن للذكاء.
- وجود علاقة دالة وموجبة عند مستوى دلالة (0.05) بين درجات تلاميذ المرحلة الابتدائية في نتائج الاختبارات التحصيلية المدرسية لمواد: اللغة العربية، والعلوم، والحساب، والدراسات الاجتماعية، والمجموع الكلي، وبين تقديرات معلميهم.
- عدم وجود علاقة دالة بين درجات تلاميذ المرحلة الابتدائية في مادة اللغة الإنجليزية وتقديرات معلميهم لهم.

ثانياً: بالنسبة للمرحلة الإعدادية:

جدول (2) قيم معاملات الارتباط بين درجات تلاميذ المرحلة الإعدادية في الاختبارات التحصيلية المدرسية ودرجاتهم في مقياس الذكاء وتقديرات معلميهم لهم

* تم حساب معامل الارتباط بين متوسط تقديرات جميع المعلمين لكل طالب والدرجة الكلية له.

المادة	معامل الارتباط بالذكاء	معامل الارتباط بتقديرات المعلمين
اللغة العربية	0.642**	0.695**
العلوم	0.517**	0.415**
الحساب	0.571**	0.514**
الدراسات الاجتماعية	0.519**	0.471**
اللغة الإنجليزية	0.510**	0.637**
الدرجة الكلية	0.679**	0.634**

** دالة عند مستوى دلالة 0.01. * دالة عند مستوى دلالة 0.05.

يتضح من الجدول السابق ما يلي:

- وجود علاقة دالة وموجبة عند مستوى دلالة (0.05) بين درجات تلاميذ المرحلة الإعدادية في نتائج الاختبارات التحصيلية المدرسية لمواد: اللغة العربية، والعلوم، والحساب، والدراسات الاجتماعية، واللغة الإنجليزية، والمجموع الكلي، وبين درجاتهم في مقياس رافن للذكاء.
- وجود علاقة دالة وموجبة عند مستوى دلالة (0.05) بين درجات تلاميذ المرحلة الإعدادية في نتائج الاختبارات التحصيلية المدرسية لمواد: اللغة العربية، والعلوم، والحساب، والدراسات الاجتماعية، واللغة الإنجليزية، والمجموع الكلي، وبين تقديرات معلميهم لهم.

ثالثاً: بالنسبة للمرحلة الثانوية:

جدول (3) قيم معاملات الارتباط بين درجات طلاب المرحلة الثانوية في الاختبارات التحصيلية المدرسية ودرجاتهم في مقياس الذكاء وتقديرات معلميهم لهم

المادة	معامل الارتباط بالذكاء	معامل الارتباط بتقديرات المعلمين
اللغة العربية	0.217	0.155
اللغة الإنجليزية	0.314*	0.499**
اللغة الفرنسية	0.213	0.121
الرياضيات	0.311*	0.399*
علم النفس	0.021	0.113
الكيمياء	0.069	0.189
المجموع الكلي	0.138	0.180

** دالة عند مستوى دلالة 0.01. * دالة عند مستوى دلالة 0.05.

يتضح من الجدول السابق ما يلي:

- وجود علاقة دالة وموجبة عند مستوى دلالة (0.05) بين درجات طلاب المرحلة الثانوية في نتائج الاختبارات التحصيلية المدرسية لمادتي: اللغة الإنجليزية، والرياضيات ودرجاتهم في مقياس رافن

للذكاء.

- عدم وجود علاقة دالة بين درجات طلاب المرحلة الثانوية في نتائج الاختبارات التحصيلية لمواد: اللغة العربية، واللغة الفرنسية، وعلم النفس والكيمياء والمجموع الكلي، ودرجاتهم في مقياس رافن للذكاء.
- وجود علاقة دالة وموجبة عند مستوى دلالة (0.05) بين درجات طلاب المرحلة الثانوية في نتائج الاختبارات التحصيلية المدرسية لمادتي: اللغة الإنجليزية، والرياضيات وتقديرات معلميهم لهم.
- عدم وجود علاقة دالة بين درجات طلاب المرحلة الثانوية في نتائج الاختبارات التحصيلية لمواد: اللغة العربية، واللغة الفرنسية، وعلم النفس والكيمياء، والمجموع الكلي، وتقديرات معلميهم لهم.

2- للتحقق من صحة الفرض الثالث للدراسة والذي ينص على: توجد فروق ذات دلالة إحصائية عند مستوى دلالة 0.05 بين متوسطي مرتفعي (الإرباعي الأعلى) ومنخفضي (الإرباعي الأدنى) الذكاء في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي. تم استخدام اختبار "ت" لدلالة الفروق بين المتوسطات غير المرتبط، وذلك على النحو التالي:

أولاً: بالنسبة للمرحلة الابتدائية:

يوضح الجدول التالي نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الذكاء في الاختبارات التحصيلية بالمرحلة الابتدائية، والذي يتضح منه أن جميع هذه الفروق ذات دلالة إحصائية عند مستوى دلالة (0.05) لمواد اللغة العربية، والعلوم والحساب والدراسات، والمجموع الكلي، بينما كانت الفروق في اللغة الإنجليزية غير دالة إحصائياً.

جدول (4) نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الذكاء في الاختبارات التحصيلية بالمرحلة الابتدائية

المادة	المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة "ت"
اللغة العربية	منخفضي الذكاء	62	25.2419	8.99396	5.35**
	مرتفعي الذكاء	43	34.8605	8.45177	
العلوم	منخفضي الذكاء	62	16.7984	4.69693	5.32**
	مرتفعي الذكاء	43	21.2326	3.27208	
الحساب	منخفضي الذكاء	62	27.7339	7.36057	6.26**
	مرتفعي الذكاء	43	36.9651	7.52804	
الدراسات	منخفضي الذكاء	62	17.2823	5.14353	5.20**
	مرتفعي الذكاء	43	21.9186	3.33083	
اللغة الإنجليزية	منخفضي الذكاء	62	18.1048	3.55765	1.46
	مرتفعي الذكاء	43	19.1744	3.84813	
المجموع الكلي	منخفضي الذكاء	62	105.1613	24.79899	6.32**

المادة	المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة "ت"
	مرتفعي الذكاء	43	134.1512	20.38894	

** دالة عند مستوى دلالة 0.01. * دالة عند مستوى دلالة 0.05.

ثانياً: بالنسبة للمرحلة الإعدادية:

يوضح الجدول التالي نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الذكاء في الاختبارات التحصيلية بالمرحلة الإعدادية، والذي يتضح منه أن جميع هذه الفروق ذات دلالة إحصائية عند مستوى دلالة (0.05). لجميع المواد الدراسية والمجموع الكلي في هذه المرحلة.

جدول (5) نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الذكاء في الاختبارات التحصيلية بالمرحلة الإعدادية

المادة	المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة "ت"
اللغة العربية	منخفضي الذكاء	37	66.2162	10.79837	**5.79
	مرتفعي الذكاء	26	85.4615	15.62429	
العلوم	منخفضي الذكاء	37	33.0405	5.35591	**5.92
	مرتفعي الذكاء	26	41.7692	6.30116	
الرياضيات	منخفضي الذكاء	37	63.2973	8.58896	**6.40
	مرتفعي الذكاء	26	79.6538	11.72755	
الدراسات	منخفضي الذكاء	37	33.0270	4.99854	**6.28
	مرتفعي الذكاء	26	41.4615	5.58377	
اللغة الإنجليزية	منخفضي الذكاء	37	36.0270	3.07306	**5.98
	مرتفعي الذكاء	26	42.1538	5.04533	
المجموع الكلي	منخفضي الذكاء	37	231.6081	21.37793	**7.32
	مرتفعي الذكاء	26	290.5000	41.85857	

** دالة عند مستوى دلالة 0.01. * دالة عند مستوى دلالة 0.05.

ثالثاً: بالنسبة للمرحلة الثانوية:

يوضح الجدول التالي نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الذكاء في الاختبارات التحصيلية بالمرحلة الثانوية، والذي يتضح منه أن الفروق الدالة إحصائياً عند مستوى دلالة 0.05 هي في مادتي اللغة الإنجليزية، والرياضيات، أما بقية المواد الدراسية وهي: اللغة العربية، واللغة الفرنسية وعلم النفس والكيمياء، والمجموع الكلي، فلم تكن دالة إحصائياً.

جدول (6) نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الذكاء في الاختبارات التحصيلية بالمرحلة الثانوية

المادة	المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة "ت"
اللغة العربية	منخفضي الذكاء	40	21.5500	4.97661	1.77
	مرتفعي الذكاء	55	23.1545	3.86206	
اللغة الإنجليزية	منخفضي الذكاء	40	15.9250	4.96855	3.24**
	مرتفعي الذكاء	55	18.8455	3.55235	
اللغة الفرنسية	منخفضي الذكاء	40	37.8375	7.73386	0.873
	مرتفعي الذكاء	55	39.1364	6.70990	
الرياضيات	منخفضي الذكاء	40	37.8125	10.04712	2.461*
	مرتفعي الذكاء	55	42.1455	7.12434	
علم النفس	منخفضي الذكاء	20	46.5811	4.21227	1.084
	مرتفعي الذكاء	31	43.8871	10.35109	
الكيمياء	منخفضي الذكاء	20	37.0000	7.64532	1.680
	مرتفعي الذكاء	24	40.6458	6.42596	
المجموع الكلي	منخفضي الذكاء	40	158.3875	28.10842	1.40
	مرتفعي الذكاء	55	165.7545	22.96291	

** دالة عند مستوى دلالة 0.01. * دالة عند مستوى دلالة 0.05.

3- للتحقق من صحة الفرض الرابع للدراسة والذي ينص على: توجد فروق ذات دلالة إحصائية عند مستوى دلالة 0.05 بين مرتفعي ومنخفضي الأداء وفقاً لتقديرات معلميهم في الامتحانات المدرسية الرسمية في مرحلة التعليم قبل الجامعي. تم استخدام اختبار "ت" لدلالة الفروق بين المتوسطات غير المرتبطة، وذلك على النحو التالي:

أولاً: بالنسبة للمرحلة الابتدائية:

يوضح الجدول التالي نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الأداء وفقا لتقديرات معلميهم في الامتحانات المدرسية الرسمية بالمرحلة الابتدائية، والذي يتضح منه أن جميع هذه الفروق ذات دلالة إحصائية عند مستوى دلالة (0.05)، فيما عدا مادة اللغة الإنجليزية.

جدول (7) نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الأداء وفقا لتقديرات معلميهم في الامتحانات المدرسية الرسمية بالمرحلة الابتدائية

المادة	المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة "ت"
اللغة العربية	منخفضي الأداء	31	19.6613	6.88644	**13.68
	مرتفعي الأداء	22	42.5000	4.39968	
العلوم	منخفضي الأداء	17	14.6176	4.29945	**8.82
	مرتفعي الأداء	17	24.0294	94324.	
الحساب	منخفضي الأداء	32	27.0469	5.74366	**5.31
	مرتفعي الأداء	36	35.8750	7.69543	
الدراسات الاجتماعية	منخفضي الأداء	36	16.6806	4.56407	**6.74
	مرتفعي الأداء	26	23.0577	1.81839	
اللغة الإنجليزية	منخفضي الأداء	41	18.0732	2.64235	0.425
	مرتفعي الأداء	48	18.3854	4.01492	
المجموع الكلي	منخفضي الأداء	31	88.8065	18.95946	**10.97
	مرتفعي الأداء	39	135.3462	16.50644	

** دالة عند مستوى دلالة 0.01. * دالة عند مستوى دلالة 0.05.

ثانياً: بالنسبة للمرحلة الإعدادية:

يوضح الجدول التالي نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الأداء وفقا لتقديرات معلميهم في الامتحانات المدرسية الرسمية بالمرحلة الإعدادية، والذي يتضح منه أن جميع هذه الفروق ذات دلالة إحصائية عند مستوى دلالة (0.05).

جدول (8) نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الأداء وفقا لتقديرات معلميهم في الامتحانات المدرسية الرسمية بالمرحلة الإعدادية

المادة	المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة "ت"
اللغة العربية	منخفضي الأداء	40	76.5375	12.15413	*2.384
	مرتفعي الأداء	35	83.1286	11.70492	
العلوم	منخفضي الأداء	37	36.9730	6.13273	**4.71
	مرتفعي الأداء	18	44.0000	2.12824	
الحساب	منخفضي الأداء	36	65.5556	10.30380	**7.82
	مرتفعي الأداء	26	83.0577	5.72595	
الدراسات الاجتماعية	منخفضي الأداء	48	36.3854	6.28341	**5.26
	مرتفعي الأداء	32	42.9688	3.98169	
اللغة الإنجليزية	منخفضي الأداء	27	35.4259	2.57093	**10.67
	مرتفعي الأداء	29	43.1379	2.81861	
المجموع الكلي	منخفضي الأداء	20	231.5833	23.04806	**7.426
	مرتفعي الأداء	21	286.8421	23.31073	

** دالة عند مستوى دلالة 0.01. * دالة عند مستوى دلالة 0.05.

ثالثاً: بالنسبة للمرحلة الثانوية:

يوضح الجدول التالي نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الأداء وفقا لتقديرات معلميهم في الامتحانات المدرسية الرسمية بالمرحلة الثانوية، والذي يتضح منه أن هذه الفروق ذات دلالة إحصائية عند مستوى دلالة 0.05 لمادتي اللغة الإنجليزية والرياضيات، أما المواد الأخرى وهي: اللغة العربية واللغة الفرنسية والكيمياء وعلم النفس والمجموع الكلي، فقد كانت الفروق غير دالة إحصائياً.

جدول (9) نتائج اختبار "ت" لدلالة الفروق بين مرتفعي ومنخفضي الأداء وفقا لتقديرات معلميهم في الامتحانات المدرسية الرسمية بالمرحلة الثانوية

المادة	المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة "ت"
اللغة العربية	منخفضي الأداء	12	22.7500	2.49089	1.79
	مرتفعي الأداء	25	24.5400	2.99277	
اللغة الإنجليزية	منخفضي الأداء	25	17.4255	4.41937	*2.106
	مرتفعي الأداء	47	19.5200	3.79212	
اللغة الفرنسية	منخفضي الأداء	21	39.7075	8.50431	0.826
	مرتفعي الأداء	33	41.3600	5.91481	
الرياضيات	منخفضي الأداء	29	32.8333	9.97795	**5.186
	مرتفعي الأداء	43	43.5079	7.23036	
علم النفس	منخفضي الأداء	10	46.9500	5.62954	0.073
	مرتفعي الأداء	25	47.0600	3.25423	
الكيمياء	منخفضي الأداء	15	37.7333	8.27618	1.69
	مرتفعي الأداء	28	41.2500	5.34114	
المجموع الكلي	منخفضي الأداء	25	165.3585	20.77336	0.705
	مرتفعي الأداء	44	169.7400	26.33720	

** دالة عند مستوى دلالة 0.01. * دالة عند مستوى دلالة 0.05.

مناقشة وتفسير نتائج الدراسة:

هدفت الدراسة الحالية إلى التعرف على مدى صدق نتائج الامتحانات المدرسية الرسمية بمرحلة التعليم قبل الجامعي، وقد أظهرت نتائج الدراسة أن امتحانات المرحلة الابتدائية، بما في ذلك المجموع الكلي، فيما عدا مادة اللغة الإنجليزية، يمكن الوثوق فيها واستخدامها في البحوث التربوية، فلها درجة من الصدق مناسبة، وكذلك الوضع في جميع امتحانات المرحلة الإعدادية والمجموع الكلي، بينما كانت امتحانات الثانوية العامة النتائج فيها مختلفة؛ فلم تكن هناك إلا مادتي اللغة الإنجليزية والرياضيات التي يمكن الوثوق في نتائجهما والاعتماد عليهما في تصنيف الطلاب وانتقائهم، بينما كانت المواد الأخرى والمجموع الكلي نتائجها لا يمكن الاعتماد عليها، كما أشارت إلى ذلك قيم معاملات الارتباط بين درجات الطلاب في هذه المواد، وتقديرات معلميهم من ناحية، وبين درجات الطلاب في هذه المواد من ناحية أخرى. وبهذا تتفق نتائج هذه الدراسة مع نتائج العديد من الدراسات

والتي بينت مستويات متباينة من الصدق للاختبارات التحصيلية موضع الدراسة، كما في دراسة "بيل وآخرون" ***(Bell & Others, 2002)***، ودراسة "تولت وتيندال" ***(Nolet & Tindal, 1990)***، ودراسة "محمد حسين سعيد" (2001)، ودراسة "أبيلا وآخرون" ***(Abella & Others, 2005)***.

وتشير نتائج الدراسة الحالية إلى أن نتائج امتحانات المرحلتين الابتدائية والإعدادية –فيما عدا اللغة الإنجليزية في المرحلة الابتدائية– يمكن للباحثين وصانعي القرار الاعتماد على نتائجها في اتخاذ القرارات المرتبطة بالطلاب؛ كتصنيفهم إلى متفوقين أو موهوبين أو متأخرين دراسيا أو ذوي صعوبات تعلم أو بطيئي تعلم، وكذلك نقلهم إلى صفوف أعلى أو أي قرارات تربوية أخرى، وبالتالي فلا داعي للباحثين والتربويين القيام بإعداد اختبارات تحصيلية في هاتين المرحلتين –فيما عدا اللغة الإنجليزية في المرحلة الابتدائية– إلا إذا دعت الضرورة أو طبيعة البحث التربوي إلى ذلك.

وقد يرجع اختلاف درجة صدق نتائج المواد الدراسية في مرحلة التعليم قبل الجامعي (جودة امتحانات المرحلة الابتدائية –فيما عدا اللغة الإنجليزية– والإعدادية وسوء نتائج المرحلة الثانوية) إلى عوامل متعددة منها ثبات هذه الامتحانات؛ فالعلاقة وثيقة بين ثبات الاختبار ودرجة صدقه، حيث يتأثر الصدق بالقيمة العددية لمعامل ثبات الاختبار تأثراً مباشراً مطرداً، فيزداد الصدق تبعاً لزيادة الثبات، ويقل تبعا لنقصان الثبات (فؤاد البهي السيد، 1978). كما قد ترجع إلى طبيعة الأسئلة في هذه الامتحانات فالأسئلة الغامضة وغير الواضحة أو الأسئلة السهلة والصعبة؛ يمكن أن تؤثر على صدق هذه الاختبارات، كما أن طول الامتحان (عدد الأسئلة في الامتحان) أو قصره ممكن أن يؤثر في درجة صدق هذه الاختبارات، حيث يزداد صدق الاختبار تبعاً لزيادة عدد أسئلته لأن ذلك الطول يضعف أثر الشوائب أو أخطاء القياس لكبر حجم عينة الأسئلة، (سعد عبدالرحمن، 1998).

كما قد يرجع الاختلاف في جودة هذه الامتحانات الرسمية قبل الجامعية إلى الاختلاف في مستوى تباين درجات الاختبار، فالتباين الضعيف يقلل من الصدق، والتباين القوي يزيد من الصدق، (سعد عبدالرحمن، 1998). كما أن لمشتتات الإجابة ***Distracters*** (قدرة البديل في الاختبارات من نوع الاختيار من متعدد على أن تبدو مقبولة بالنسبة لمن لا يعرف الإجابة الصحيحة) أثر على الصدق الفارقي، حيث يزداد معامل الصدق بزيادة قوة تشتيت البدائل، (فتحي مصطفي الزيات، 1989). كما يرى "إيزاك وإدوين" ***(Isac & Edwin, 1981)*** أن تضخم درجات ***Grade Inflation*** الاختبار يمكن أن تؤثر على صدق الاختبارات التحصيلية إذا كانت

هذه الزيادة غير منتظمة، أما إذا كان التضخم عبارة عن إضافة ثابتة لدرجات الاختبار فإنه لا يؤثر علي صدقها.

كما أن الحالة النفسية للطالب، وظروف تطبيق الاختبار، لها دور كبير في التأثير في مستوى

صدق هذه الاختبارات؛ فالأسرة والأقران كلها من العوامل التي تؤدي إلى اضطرابات نفسية للطالب أثناء الامتحان، قد يدفع البعض منهم إلى التفكير في الانتحار (عبدالرقيب البحيري، محفوظ عبدالستار، 2008)، فقد أصبح الامتحان وسيلة تحكم به الجماعة على أداء الفرد بالنجاح أو الفشل، ولم يكن الهدف منه معرفة الفرد لمقدار تقدمه في عمله، أو معرفة مواطن فشله أو ضعفه ليتغلب عليها، حيث تبنى الامتحانات أو الاختبارات التحصيلية بوضعها الراهن، علي أساس فلسفة تربوية تؤكد بل وتشجع إبراز الفروق الفردية، وتحث على التعلم التنافسي من أجل حصول الفرد علي مركز نسبي متفوق بين أقرانه، دون محاولة لتحديد ما يمتلكه الفرد من مهارات وظيفية وأخلاقيات وسلوكيات إيجابية بناءة، لذلك فقد أدت في كثير من الأحيان إلى العديد من المشكلات الاجتماعية والتربوية التي تعوق تقدم المجتمعات مثل إثارة البغض بين الأفراد، والتحيز ضد الطلاب محدودي القدرة أو بطيئ التعلم، وعدم تكافؤ الفرص التعليمية لهم، والاختيار غير الصائب للطلاب للالتحاق بالجامعة، كما أنها لا تقيس بدقة الأهداف التعليمية الجوهرية التي ينبغي أن يسعى المعلم إلى تحقيقها داخل الصف الدراسي،(صلاح الدين محمود علام، 1995).

التوصيات:

في ضوء ما أسفرت عنه نتائج الدراسة الحالية فإنه يمكن تقديم مجموعة التوصيات التالية:

1. الاهتمام بامتحانات المرحلة الثانوية ومحاولة تطويرها وبنائها في ضوء أسس علمية وموضوعية سليمة.
2. الاهتمام بمادة اللغة الإنجليزية في المرحلة الابتدائية ومحاولة تطويرها وبنائها في ضوء أسس علمية وموضوعية سليمة.
3. يمكن الاعتماد على المجموع الكلي وكذلك درجات امتحانات المواد المختلفة في المرحلتين الابتدائية والإعدادية، فيما عدا مادة اللغة الإنجليزية في المرحلة الابتدائية، وبالتالي يمكن الاعتماد عليها في البحوث النفسية والتربوية دون الحاجة إلى إعداد اختبارات تحصيلية فيها.
4. عند دراسة ذوي الاحتياجات الخاصة (المتفوقين–الموهوبين–ذوي صعوبات التعلم–بطيئي التعلم–المتأخرين دراسيا) في المرحلة الثانوية، فإنه لا يمكن الاعتماد على درجات الطلاب في نتائج الامتحانات في هذه المرحلة، وإنما يجب إعداد اختبارات تحصيلية تتميز بالصدق.
5. الاهتمام بالحالة النفسية للطلاب أثناء الامتحانات، وبصفة خاصة طلاب المرحلة الثانوية، فهم في حاجة إلى برامج تربوية وإعلامية للحد من القلق الذي يعتريهم أثناء الامتحانات.
6. يمكن الاعتماد على تقديرات المعلمين لمستويات طلابهم التحصيلية، فقد كان هناك اتفاق بين تقديرات المعلمين ونتائج اختبارات الذكاء في علاقة كل منها بدرجات الطلاب في الامتحانات المدرسية الرسمية.

البحوث المقترحة:

في ضوء نتائج الدراسة الحالية فإنه يمكن إجراء العديد من الدراسات والتي منها:

1. دراسة تحليلية لأسئلة امتحانات المرحلة الابتدائية وخاصة اللغة الإنجليزية للتعرف على جوانب القصور والقوة فيها.
2. دراسة تحليلية لأسئلة امتحانات المرحلة الإعدادية للتعرف على جوانب القصور والقوة فيها.
3. دراسة تحليلية لأسئلة امتحانات المرحلة الثانوية للتعرف على جوانب القصور والقوة فيها.
4. دراسة صدق الاختبارات التحصيلية الرسمية في مرحلة التعليم قبل الجامعي في محافظات أخرى.
5. دراسة ثبات الاختبارات التحصيلية الرسمية في مرحلة التعليم قبل الجامعي.
6. دراسة المستويات المعرفية التي تقيسها الامتحانات في مرحلة التعليم قبل الجامعي.
7. دراسة صدق وثبات الامتحانات في مرحلة التعليم الجامعي (الحكومي والخاص) في كليات وتخصصات مختلفة.

المراجـــع

أولاً: المراجع العربية:

1- أحمد زكي صالح (1978). اختبار الذكاء المصور للأطفال. القاهرة: مكتبة الأنجلو المصرية.

2- أحمد عثمان صالح (1988). أثر عامل الثقافة في الاختبارات المتحررة من أثر الثقافة في ضوء تقنين اختبار المصفوفات المتتابعة على البيئة المصرية، مجلة البحث في التربية وعلم النفس، يناير، ص ص211-243.

3- أمينة محمد كاظم (1996). اتجاهات معاصرة في القياس والتقويم النفسي والتربوي. القاهرة: مكتبة الأنجلو المصرية.

4- أمينة محمد كاظم، صلاح مراد، إسحاق بطرس (1989). تطوير نظم الامتحانات بمراحل التعليم العام وبناء بنوك الأسئلة فى بعض الدول العربية. حقيبة دراسية إقليمية حول تطوير الامتحانات وبناء بنوك الأسئلة فى الدول العربية، برنامج التجديد التربوى من أجل التنمية فى الدول العربية (إبيداس) القاهرة 25-29 نوفمبر، ص ص 1-43.

5- توما جورج خورى (1991). الاختبارات المدرسية ومرتكزات تقويمها. بيروت: المؤسسة الجامعية للدراسات والنشر والتوزيع.

6- جابر عبدالحميد (1994). علم النفس التربوي، ط 3. القاهرة: دار النهضة العربية.

7- حسن حسينى جامع (2001). تقويم الأوراق الامتحانية للثانوية العامة فى الفترة من 1992-1999م، قسم البحوث، المركز القومى للامتحانات والتقويم التربوى بالقاهرة.

8- حمدى عبدالعظيم البنا (2001). دراسة تحليلية للمستويات المعرفية فى امتحانات الكيمياء للشهادة الثانوية العامة، مجلة التربية العلمية، الجمعية المصرية للتربية العلمية، المجلد الرابع، العدد الأول، يناير ص ص 9-35.

9- رضا عبدالقادر درويش (1995). تقويم "دليل تقويم التلميذ فى مادة العلوم العامة للصف الثالث الإعدادى" فى ضوء أهداف تدريس العلوم. مجلة كلية التربية ببنها، الجزء الثانى، يناير ص ص 227-255.

10- رمزية الغريب (1996). التقويم والقياس النفسي والتربوي. القاهرة: الأنجلو المصرية.

11- زين بن حسن ردادي (2002). المعتقدات الدافعية واستراتيجيات التنظيم الذاتي للتعلم في علاقتهما بالتحصيل الدراسي لدى تلاميذ مدارس منارات المدينة المنورة، مجلة كلية التربية، جامعة الزقازيق، العدد 41، ص ص 171- 234.

12- سعد جلال (1985). القياس النفسي "المقاييس والاختبارات". القاهرة: دار الفكر العربي.
13- سعد عبدالرحمن (1998). القياس النفسي "النظرية والتطبيق"، ط3. القاهرة: دار الفكر العربي.
14- سيدة حافظ محمد شاهين (2005). فاعلية برنامج تعلم بعض المهارات ما بعد المعرفية في تنمية النضج الاجتماعي والاتجاه نحو التعلم الذاتي والتحصيل الدراسي، رسالة ماجستير، معهد الدراسات التربوية، جامعة القاهرة.
15- صفوت فرج (1989). القياس النفسي، ط2. القاهرة: الأنجلو المصرية.
16- صلاح الدين محمود علام (1995). التوجهات المستقبلية لتقويم تحصيل الطلاب في ضوء متطلبات القرن الحادي والعشرين، مجلة كلية التربية، جامعة الأزهر، العدد 49، إبريل، ص ص 15 – 53.
17- صلاح الدين محمود علام (2000). القياس والتقويم التربوى والنفسى، أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة: دار الفكر العربى.
18- عادل محمد محمود العدل (1996). التنبؤ بالتحصيل الدراسي من بعض المتغيرات غير العلمية، مجلة دراسات نفسية، رابطة الأخصائيين النفسيين، المجلد 6، العدد 1، ص ص 82-117.
19- عاطف عبدالعزيز عبدالمقصود (2003). فعالية استخدام استراتيجية التعلم التعاوني على التحصيل الدراسي وبقاء أثر التعلم لدى طلاب المدرسة الثانوية الزراعية، مجلة البحوث النفسية والتربوية، كلية التربية، جامعة المنوفية، العدد 3، السنة 18، ص ص 157 – 219.
20- عبد الهادى السيد عبده، فاروق عثمان (1995). الإحصاء التربوي والقياس النفسي. القاهرة: دار المعارف.
21- عبدالرقيب البحيري، محفوظ عبد الستار (2008). بعض الاضطرابات النفسية المرتبطة بالتفكير الانتحاري لدى عينة من طلاب المرحلة الثانوية. المجلة المصرية للدراسات النفسية، العدد 60، المجلد الثامن عشر، يوليو، ص ص 1-55.
22- عبدالمجيد نشواتي (1991). علم النفس التربوي، ط5. عمان: دار الفرقان.
23- عبدالوارث عبده الرازحى (2001). تطوير نموذج معيارى لتقويم كفاءة نظام إعداد الاختبارات العامة. المؤتمر العربى الأول للامتحانات والتقويم التربوى: رؤية مستقبلية، المركز القومى للامتحانات والتقويم التربوى، القاهرة: 22-24 ديسمبر، ص ص 107-120.

24- عبلة محمد الجابر مرتضى (2002). التحصيل الدراسي وعلاقته بأسلوب التروي/الاندفاع عند مستويات مختلفة من القلق والذكاء لدى طالبات المرحلة الثانوية، رسالة ماجستير، كلية البنات للآداب والعلوم والتربية، جامعة عين شمس.

25- عفاف محمد عبدالمنعم محمد (1999). تأكيد الذات وعلاقته بالتحصيل الدراسي لدى طلاب المرحلة الإعدادية، رسالة ماجستير، معهد الدراسات العليا للطفولة، جامعة عين شمس.

26- علي حامد الشبيتى (1996). الصدق التنبؤي لمعايير القبول في كلية المعلمين بالطائف، رسالة الخليج العربي، مكتب التربية العربي لدول الخليج، العدد 57، السنة 16، ص ص 57 – 83.

27- فؤاد أبوحطب (1977). بحوث في تقنين الاختبارات النفسية، المجلد الأول. القاهرة: مكتبة الأنجلو المصرية.

28- فؤاد أبوحطب (1983). القدرات العقلية، ط3. القاهرة: الأنجلو المصرية.

29- فؤاد البهي السيد (1978). علم النفس الإحصائي وقياس العقل البشري. القاهرة: دار الفكر العربي.

30- فتحي مصطفي الزيات (1989). أثر قوة تشتيت البدائل في أسئلة الاختيار من متعدد علي صدق الاختبار وثباته، دراسة تحليلية، مجلة كلية التربية، جامعة المنصورة، العدد الحادي عشر، سبتمبر، ص ص 87-108.

31- فتحي مصطفى الزيات (2001). علم النفس المعرفي، الجزء الأول، دراسات وبحوث. القاهرة: دار النشر للجامعات.

32- فكرى حسن ريان (1994). تقويم نماذج أسئلة المواد الدراسية بالتعليم الثانوى لعام 1991/1990 قسم البحوث، المركز القومى للامتحانات والتقويم التربوى بالقاهرة.

33- كمال محمد دسوقي (1988). ذخيرة علم النفس، الجزء الأول. القاهرة: الدار الدولية للنشر والتوزيع.

34- اللجنة الوطنية القطرية للتربية والثقافة والعلوم (1984). الامتحانات ما لها وما عليها، مجلة التربية، العدد الرابع والستون، رجب 1404 هـ- إبريل، ص ص 16 – 38.

35- مجدي أحمد محمد (1996). علم النفس التجريبي بين النظرية والتطبيق. الإسكندرية: دار المعرفة الجامعية.

36- محسن محمد أحمد عبدالنبي (2001). العلاقات التفاعلية بين الذكاء الانفعالي والتفكير

الابتكاري والتحصيل الدراسي للطالبات الجامعيات السعوديات، مجلة البحوث النفسية والتربوية، كلية التربية، جامعة المنوفية، العدد 3، السنة 16، ص ص 129 – 166.

37- محمد السيد حسونة (1995). التقويم التربوي، مفهومه، أهدافه، أسسه وأساليبه، صحيفة التربية، العدد الثاني، السنة السادسة والأربعون، يناير، ص ص 8 – 22.

38- محمد حسين سعيد (2001). درجات امتحان الثانوية العامة "دراسة سيكومترية". رسالة ماجستير غير منشورة، كلية التربية ببني سويف، جامعة القاهرة.

39- محمد حسين سعيد (2005). تطوير أساليب التقويم ضرورة حتمية لضمان جودة المؤسسات التعليمية. ورقة عمل مقدمة إلى المؤتمر العلمي السنوي الثالث عشر للجمعية المصرية للتربية المقارنة والإدارة التعليمية بعنوان "الاعتماد وضمان جودة المؤسسات التعليمية"، في الفترة 29–30 يناير، كلية التربية ببني سويف، جامعة القاهرة.

40- محمد خيرى محمود (1994). دراسة تقويمية تحليلية للأسئلة المتضمنة فى كتب العلوم للمرحلة الاعدادية. المجلة المصرية للتقويم التربوى، المجلد الثانى، العدد الأول، ص ص 55–105.

41- محمد رضا البغدادي (1998). الأهداف والاختبارات في المناهج وطرق التدريس بين النظرية والتطبيق. القاهرة: دار الفكر العربي.

42- محمود عبدالحليم منسي، حسن أحمد عيسى، فوزي عزت علي، مصطفي السعيد جبر (1993). تحليل نتائج امتحانات الثانوية العامة في جمهورية مصر العربية في الفترة من 1991 إلي 1993، المركز القومي للامتحانات والتقويم التربوي، قسم البحوث.

43- نادية على سعود (1999). دراسة تتبعية تقويمية لأسئلة امتحانات إتمام الشهادة الإعدادية فى جمهورية مصر العربية من عام 1991 إلى 1994 فى مادة اللغة العربية. المجلة المصرية للتقويم التربوى، المجلد السادس، العدد الأول، ص ص 63–104.

ثانياً: المراجع الأجنبية:

44- Abella, R. & Urrutia, J. & Shneyderman, A. (2005). An Examination of the Validity of English-Language Achievement Test Scores in an English Language Learner Population. Bilingual Research Journal, v29, n1, p127-144 Spr.

45- Angoff, W. (1988). Validity: An Evoloving Concept. In Wainer, H & Braun, H. (Eds), Test Valllidity. Hillsdale, NJ: Lawrence Erlbaum.

46- Bell, L. & Rucker, M. & Finch, A. & Alexander, J. (2002). Concurrent Validity of the Slosson Full-Range Intelligence Test: Comparison with the Wechsler Intelligence Scale for Children--Third Edition and the Woodcock Johnson Tests of Achievement-Revised. Psychology in the Schools, v39, n1, p31-38 Jan.

47- Cronbach, L. (1988). Five Perspectives on Validity Argument. . In Wainer, H & Braun, H. (Eds), Test Valllidity. Hillsdale, NJ: Lawrence Erlbaum.

48- Ebel, R. & Frisbie, D. (1986). Essentails of Educational Measurement. Fourth Edition. Prentice-Hall Inc. Englewood Cliffs.

49- Goldstein, H. (1996) Assessment: Problems, Developments and Statistical Issues. Wiley Series in Probability and Statistics. Chichester: Wiley.

50- Gronlund, N. & Linn, R. (1990). Measurement and Evaluation in Teaching, (6th ed.) New York: Macmillan.

51- Grosser, G. & Spafford, C. (1995). Psychology Dictionary, New York, McGraw – Hill.

52- Hines, V. (1981). A Study of the Construct Validity of Six Vocational Achievement Tests in the Ohio Vocational Education Achievement Test Program. Journal of Industrial Teacher Education, v19, n1, p3-13 Fall.

53- Hopkins, K. & George, C. & Williams, D. (1985). The Concurrent Validity of Standerdised Achievement Tests by Content Area using Teachers' Ratings as Critera. Journal of Educational Measurement, v22 n3, 177–182.

54- Issac, I. & Edwin, O. (1981). Grade Inflation And The Validity Of The Scholastic Aptitude Test, American Educational Research Journal, v18, n2, p143-156.

55- Messick, S. (1989). Validity. In Linn, R.L. (Ed) Educatioanl Measurement. Thrd Ed. American Council on Education. Macmillan Publishing Co. N.Y.

56- Nolet, V. & Tindal, G. (1990). Evidence of construct validity in published acheivement tests. Paper presented at the annual meeting of the american educational research association, Boston, KA, April 16-20.

57- Scott, C. (2000). Should Achievement Tests be Used to Judge School Quality?. Education Policy Analysis Archives. v8, n46, September 17, 20-27.

58- Taylor, D. (1999). National Testing: Past, Present and Future. Issues in Assessment and Testing. The British Psychological Society.

59- Wiliam, D. (1996). National Curriculum Assessment and Programmes of Study: valisity and impact. British Educational Research, v22, n1, p15-19.

The Validity of Achievement Tests in the pre- university Stage."An Evaluative Study"

Mohamed Hussein Said Hussein
Lecturer at the Faculty of Education in Beni- Suef
Educational Psychology Department

Abstract:

The present study aims at detecting the validity of formal school tests in the pre-university stage through recognizing the relation between the results of students on one hand and the levels of their intelligence and the estimations of their teachers to their achievement levels on the other hand. The sample of the study consists of 450 students (males and females) in the pre-university stage in Beni-Suef governorate (the results of 183 students in the primary stage, the results of 130 students in the preparatory stage and the results of 137 stdnts in the secondary stage). The descriptive methodology is used as it suits the aims of the study. The study depends on the following tools: Ravn Scale for intelligence, the translated version of Foaad Abou Hatab (1977), a form of estimation of students levels from the viewpoint of their teachers , prepared by the researcher and the results of students. Person and Spearman Brawn Coefficients of Correlation are used in this study. T test is also used to determine the differences between means that are not correlated. The results of the study show the following: 1-The results of tests in the primary stage (Arabic, Math, Science and Social studies) and the total summation are valid according to criterion and peripheral comparison. The problem was in English subject as the marks of students were not connected to the standards of intelligence or the estimations of their teachers to their levels. 2- The results of tests in the preparatory stage (Arabic, Math, Science, Social Studies and English) and the total summation were valid according to criterion and peripheral comparison. 3- The results of tests in the secondary stage were valid in English and math only according to criterion and peripheral comparison, but they were not correlated in the other subjects: Arabic, French, Chemistry, Psychology and the total summation with their standards of intelligence or the estimations of their teachers to their levels.