

العنوان:	التقييم الحقيقي "البورتفوليو" كمدخل تكاملي للاختبارات التحصيلية
المصدر:	المجلة المصرية للدراسات النفسية
الناشر:	الجمعية المصرية للدراسات النفسية
المؤلف الرئيسي:	عامر، عبدالناصر السيد
المجلد/العدد:	مج14، ع44
محكمة:	نعم
التاريخ الميلادي:	2004
الشهر:	يوليه
الصفحات:	171 - 219
رقم MD:	1008664
نوع المحتوى:	بحوث ومقالات
اللغة:	Arabic
قواعد المعلومات:	EduSearch
مواضيع:	الاختبارات والقياسات التعليمية، التحصيل العلمي، علم النفس التربوي
رابط:	http://search.mandumah.com/Record/1008664

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

إعداد

د. عبد الناصر السيد عامر

مدرس القياس والتقويم بكلية التربية
بالإسماعيلية - جامعة قناة السويس

ملخص الدراسة

ركزت الدراسة على أسلوب التقييم الحقيقي كمدخل تكاملي للاختبارات المعيارية. ولعينة مكونة من عنتين فرعيتين هما عينة الدراسة (١٢٦) طالب من طلاب الدراسات العليا (دبلوم مهني في التربية) ٢٠٠٣ بكلية التربية بالإسماعيلية - جامعة قناة السويس. وعينة صدق النتائج وتكونت من (٦٤) طالب. والأدوات شملت استمارة التقييم الذاتي ومقياس الاتجاه نحو خصائص التعلم والمتعلم والتدريس واختبار تحصيلي مكون من ٣٠ سؤال واستمارة لتقييم الأنشطة التقييمية بعض البيانات الارشيفية.

تبين أن اتجاهات العينة كانت إيجابية بدرجة كبيرة نحو خصائص التعلم والمتعلم والتدريس في حالة استخدام التقييم الحقيقي وتميز استجابات العينة بالسلبية نحو نفس الخصائص في حالة استخدام الاختبارات المعيارية. وتبين وجود علاقة ارتباطية موجبة بين الأداء في ضوء التقييم الحقيقي والأداء في ضوء الاختبارات المعيارية بلغت ٠,٦٤ لعينة الدراسة و ٠,٥٢ لعينة صدق النتائج، ووجود فروق ذات دلالة إحصائية بين الأداء على التقييم الحقيقي والأداء في ضوء الاختبارات المعيارية. وجود علاقة ارتباطية سالبة بلغت ٠,٢٣ - بين التقييم الذاتي والأداء في ضوء أنشطة الحقيبة التقييمية. وتوصلت الدراسة إلى أن أفضل نموذج للتنبؤ بالتحصيل في نهاية الترم هو المعتمد على درجة الاختبارات التقليدية ودرجة التقييم الحقيقي معاً ذلك لعينة الدراسة وعينة صدق النتائج وهذه النتيجة تؤكد على أهمية التكامل بين التقييم الحقيقي والاختبارات المعيارية لصناعة قرارات حول التلاميذ.

التقييم الحقيقي (البورتفوليو) كمدخل

تكاملي للاختبارات التحصيلية

إعداد

د. عبد الناصر السيد عامر

مدرس القياس والتقويم بكلية التربية

بالإسماعيلية - جامعة قناة السويس

مقدمة:

ركز أصحاب التجديد التربوي اهتماماتهم على تحسين المناهج وطرق التدريس ولم يعطوا نفس الاهتمام لتحسين الممارسات التقييمية الحالية والخبرة الواقعية تقول انه يمكن بسهولة تغيير الممارسات في مجال المناهج ولكن هذا التغيير يكون أكثر صعوبة إذا كان الأمر يتعلق بالممارسات التقييمية وذلك لعدم تحمس القائمين على العملية التعليمية على تغيير الوضع التقييمي السائد (Firestone et al., 1998, 2000 Shepard). والتقييم Assessment عملية متعددة الأوجه تتضمن أكثر من إدارة الاختبار وتهدف إلى جمع البيانات بغرض صناعة القرار على مستوى الطالب أو على مستوى عملية التدريس (McLoughlim & Lewis, 1990). وتتضمن عملية التقييم الأنشطة المصاحبة للحصول على التقديرات أو الدرجات وتتم في ضوء دلائل متعددة.

والامتحانات الحالية تجعل المناهج جامدة وتجعل المدرسة مكاناً غير محبوباً وتبث الرعب والقلق في التلاميذ وتجعل عملية التدريس من أجل الامتحان (Heck & Crislip, 2001) وتعد في نهاية عملية التعلم لذلك فهي منفصلة عن عملية التدريس، ولا تهدف إلى تطوير المنهج (Camp & Levine, 1991) والممارسات التقييمية المتبعة في مدارسنا يوجه إليها كثير من الانتقادات ورغم ذلك لم تتغير منذ زمن بعيد ولم تواكب التغيرات المذهلة في مجال التربية ونظرية التعلم.

وظهرت الحاجة بصورة متزايدة لتحسين الممارسات التقييمية في النظام التعليمي لتواكب التغيرات المتلاحقة في مجال علم نفس التعلم Learning Psychology
المجلة المصرية للدراسات النفسية - العدد ٤٤ - المجلد الرابع عشر - يوليو ٢٠٠٤ - (١٧٢) =

ومبادئ النظرية البنوية في رؤيتها لعملية التعلم، وأيضاً لتغير فلسفة التقييم منذ العقد الأخير من القرن الماضي وبحيث أصبحت فلسفة التقييم قائمة على أساس Large-Scale Assessment ، حيث يتم تقييم الطالب في ضوء أدلة متعددة مثل الأنشطة أو التكاليف بدلاً من التركيز على الأداء في ضوء اختبار وحيد (Camp, 1993)، والتقييم في ضوء حزمة الأداء Performance Package حتى يظهر الطالب مدى الفهم العميق لما يتعلمه (Newell, 2002) وأهمية تضمين ومشاركة المتعلم في عملية التقييم (Cross, 1998).

ونتيجة لتغير فلسفة التقييم وتغير مفهوم التعلم في ضوء آراء مدرسة علم النفس المعرفي ووجهة النظرية البنوية لعملية التعلم Constructivist Learning Theory تم استحداث طرق أو مداخل تقييمية عديدة أطلق عليها مفهوم التقييم البديل Alternative assessment أو التقييم في ضوء الأداء Performance based assessment أو التقييم الأصيل (الحقيقي) Authentic assessment.

والتقييم البديل هو شكل من أشكال التقييم بالمقارنة بالتقييم التقليدي (الاختبار من متعدد مثلاً) ويتضمن أشكال عديدة مثل الحقيقية ومجموعات التعلم التعاوني والملاحظة والعرض الشفوي، والشيء الجوهرى لهذا التقييم البديل هو تركيزه على العمليات تماماً مثل النواتج (Davies & Michael, 1999).

وفي ضوء الحكم الهائل من الدراسات والمقالات المنشورة، ظهرت مداخل تقييمية جديدة منها التقييم المرتكز على المنهج Curriculum based assessment والتقييم الدينامي Dynamic assessment والتقييم الحقيقي (البورتفوليو) Portfolio assessment ويطلق عليه البعض البورتفوليو، أو تقويم ملفات الأعمال*.

العوامل التي أدت إلى إعادة النظر في الممارسات التقييمية الحالية
ظهرت مداخل التقييم البديل نتيجة لعدة عوامل أهمها:

* هذا المصطلح مأخوذ عن علام، صلاح الدين (تحت النشر). التقويم التربوي البديل: أسسه النظرية والمنهجية وتطبيقاته الميدانية، دار الفكر العربي.

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

أولاً: تغير مفهوم التعلم في ضوء المدرسة السلوكية عن مفهومه في ضوء توجهات النظرية المعرفية وأصحاب اتجاه علم النفس البنوي

ظهرت النظرية الارتباطية Association Theory على يد Thondike (1921) وما تبعه من بزوغ المدرس السلوكية Behaviorism على يد Hull (1943) و (1938, 1954) Skinner و (1965) Gagne ورؤية هذه النظرية السلوكية لعملية التعلم باعتباره تراكم لأجزاء المعرفة وهذا التراكم ينشأ نتيجة مجموعة كبيرة من الارتباطات بين المثيرات والاستجابات (النظرية الارتباطية) وما يتبعها من عملية تعزيز (السلوكية). ونظرت السلوكية إلى عملية التعلم باعتباره ذات طبيعة هرمية أو خطية كما وضع ذلك Bloom (1956) و Gagne (1965). وفي هذه النظرية لعبت الاختبارات دوراً هاماً وجوهرياً في عملية التعلم وذلك لمعرفة درجة التمكن من المهارات والأهداف السلوكية، أى أن الأهداف موجه لتحقيق أو للوصول إلى التمكن للمعلومات أو المعارف وهذا ما حدث على يد Skinner (1954) في نموذج التعليم المبرمج حيث لا ينتقل المتعلم إلى الهدف التالي إلا إذا أنجز الأهداف السابقة ويعود الفضل إلى السلوكيين في ابتداعهم للاختبارات الموضوعية خاصة على يد Thondike والسلوكيين يضعون الاختبارات في نهاية عملية التعلم وتتم بصورة مباشرة حيث يحدث تماثل بين أسئلة الامتحانات والأهداف السلوكية التي درسها المتعلم ويتسم دور المتعلم بالسلبية سواء في عملية التعلم أو التقييم. ولا تشجع عملية التقييم الطالب على الإبداع أو الابتكار إنما هي محددة بإطار ومحتوى وأسئلة يجيب عليها المتعلم في نهاية عملية التعلم.

ومع ظهور تطورات في مجال علم نفس التعلم منها توجهات المدرسة المعرفية ومبادئ النظرية البنوية (علم النفس البنوي)، وإبراز مفهوم التعلم الذاتي، بدأ ينظر إلى مفهوم التعلم باعتباره عملية نشطة يشارك فيها المتعلم بإيجابية وأن دورة ليس مستقبلاً للمعارف إنما دوره أن يبنى المعرفة بنفسه، وترى Shepard (2000) أن مبادئ عملية التعلم في ضوء هذه التطورات هي:

١- تتمى القدرات العقلية تتمى في ضوء إطارها الاجتماعي الثقافي: ويلاحظ أن

مفهوم القدرات العقلية تم إحلاله محل مفهوم الذكاء وأن القدرات العقلية يتم تتميتها في إطار البيئة الاجتماعية الثقافية.

٢- المتعلم صانع وباني للمعارف: وهذا ما أكدت عليه المبادئ البنائية وهي أن كل المعارف الإنسانية بنائية وليست مكتشفة والمتعلم يبذل جهداً كبيراً للحصول على المعارف. وهذا التحول من الموقف السلبي للمتلم (كمتلقي للمعرفة) في ضوء المبادئ السلوكية إلى الدور النشط الإيجابي للمتلم وإسهامه الفعال في بناء المعرفة كان له أثر في إعادة النظر لعملية التقييم.

٣- أن تتضمن عملية التعلم وعي المتعلم بما يتعلمه: وهذا يتطلب أن تشمل عملية التعلم تنمية مهارات ما وراء المعرفة والتعلم المضبوط ذاتياً Self-monitoring learning.

٤- الفهم العميق يدعم انتقال أثر التعلم: حيث أن انتقال أثر التعلم في ضوء المبادئ السلوكية يحدث نتيجة لدرجة التشابه بين الموقفين وليس للفهم أي دور في انتقال أثر التعلم.

وقد أشار Herman (1992) أنه وفقاً للباحثين المعرفيين فإن التعلم القائم على معنى يجب أن يكون تأملي وبنائي ومنظم ذاتياً (في: Davies & Michael, 1999). ولابد من استحداث مداخل جديدة للتقييم لكي تنمي هذه المهارات وغالباً ما تفشل الاختبارات التقليدية في تحقيق هذه الأهداف (O'Neil, 1992).

في ضوء ذلك فإن دور المتعلم أصبح أكثر إيجابية، حيث يسهم بدور فعال في بناء المادة المتعلمة وهذا ما دعى إلى إعادة النظر في الأشكال التقييمية التقليدية التي هدفت إلى قياس ما استوعبه المتعلم من معارف جاهزة وتأتي في نهاية عملية التعلم، ولكن في ضوء بناء المتعلم للمعارف لم يعد التقييم في ضوء النواتج النهائية فعالاً ولكن لابد من التركيز على تقييم العمليات Processes assessment.

التقييم الحقيقي (البورنوليوس) كمدخل تكاملي للاختبارات التحصيلية

ثانياً: الانتقادات الموجهة إلى الممارسات التقييمية الحالية (الاختبارات التقليدية^(*)) وجهت إلى الممارسات التقييمية الحالية كثيراً من الانتقادات أهمها إنها غير قادرة على توجيه العمل التدريسي اليومي (Kim et al., 1996, Salvia & Yesseldyke, 1995) ولا تمدنا بصورة دقيقة عن مدى قدرة الطالب على ممارسة العمليات العقلية العليا (Murphy, 1994) ولا تقيس سوى مستوى الحفظ والاستظهار (Newell, 2000) وغير قادر على إظهار قدرات الطالب العقلية مثل القدرة على التكامل بين المعلومات وتطبيقها في المواقف الحياتية (Shepard, 2000) وغير قادرة على تقييم العمليات والاستراتيجيات التي يستخدمها المتعلم في الوصول إلى الحل (Camp, 1993) وأن تدريس المعلم لمحتوى بأسئلة محددة يعطي تضخم Inflation عن مستوى تحصيل الطالب (Newell, 2002, Shepard, 2000) ويصبح موقف التلميذ سلبي في الممارسة التقييمية حيث لا يمتلك صوت فيها (Pairs & Ayers, 1994) ولا يمكن الاستفادة منها في تقويم المنهج (Pairs & Ayers, 1994, Salvia & Yesseldyke, 1995) وتسبب القلق والرعب وتفش عملية الغش ومشاهدة الأحلام اليومية المخيفة (Heck & Crislip, 2001, Pairs & Ayers, 1999) وتركز على التنافس وليس النمو الشخصي لكل تلميذ (Crooks, 1988) وأهملت التفاعل بين المكونات الثلاثة للمنظومة وهي التعلم والتدريس والتقييم (Crooks, 1988, Kim et al., 1996). وترى (Shepard, 2000) إلى أن التدريس والتقييم عمليتان منفصلتان وأرجعت ذلك إلى عدم الاتفاق بين وجه النظرية التقليدية لعملية التقييم والمتمثلة في الاختبارات التقليدية ووجهة النظر الحديثة لعملية التعلم، وأعزت (Graue, 1993) هذا الفصل إلى طبيعة أساليب التقييم السائدة في المدارس (في: Shepard, 2000)، بينما يرى (Camp & Levine, 1991) سبب هذا الفصل إلى أن الامتحانات يتم

(*) يقصد بالاختبارات التقليدية، الاختبارات التحصيلية المبنية في إطار نظريات القياس التقليدية (جماعية /

محكية - المرجع)، تمييزاً لها عن تلك التي تبنى وترج في إطار نظرية القياس الحديثة (نظرية

الاستجابة المفردة IRT Item Response Theory) ومن أهمها نموذج (راش) Rasch Model.

عقدما في نهاية عملية التدريس، ويرى Kim et al., (1996) ويعود هذا الفصل إلى أن نمو عملية التعلم والتدريس والتقييم نمت بصورة مستقلة عن الأخرى في التراث. كما أن الممارسات الحالية تشجع على التعلم عديم المعنى مما يسهل عملية النسيان للمعلومات ولا تتيح الفرصة للمعلمين لمراجعة أسئلة الامتحانات ولا يتناقشون حولها بفاعلية (Crooks, 1988).

وفي ضوء هذه الانتقادات الموجهة للممارسات التقييمية الحالية، نادى كثيراً من المحدثين وعلماء القياس والتقييم بضرورة تغيير هذه الممارسات لكي نتجنب كثيراً من الانتقادات السابقة.

ثالثاً: فلسفة التقييم في القرن الحالى

يمكن تلخيص فلسفة التقييم في القرن الحالى في أنها تركز على تقييم الجودة Quality assessment وليس فقط الحصول على شهادة (Cross, 1998) وأن يشارك الطالب في تقييم أعماله لأنها شيء ضرورى حتى يتحمل مسئولية تعلمه (Cross, 1998, Flachikov & Goldfinsh, 2000) وتصبح عملية التقييم عملية تفاوضية وليست من جانب المعلم، ويتم التقييم في ضوء دلائل أو تكاليف متنوعة بدلاً من الاعتماد على الاختبارات فقط (Camp, 1993) ويجب أن تتسع لتشمل مقابلات وعرض شفوى وعينات من أعمال الطالب وهنا يزيد من خصوبة ودينامية عملية التقييم (Shepard, 2000)، وحتى يظهر الطالب مدى فهمه العميق لما يتعلمه (Newell, 2000) وأن يتم التركيز على تقييم العمليات تماماً مثل النواتج. ويتم عملية التقييم أثناء عملية التدريس وتسعى إلى تحسينها وهذا بدوره ينمى ويوطد العلاقة بين التدريس والتقييم (Camp, 1993) ويزيد التواصل بين المعلم والمتعلم من جهة وبين الطلاب ومجتمعهم من جهة أخرى (Cole & Struyk, 1997). وفي ضوء مشاركة الطالب في تقييم أعماله تم تنمية مفهوم التقييم الذاتى Self assessment (Falchikov & Goldfinch, 2000) وأن يتم التركيز على المدخل المحكى في عملية التقييم وبحيث يتم وضع معايير للأداء حتى يتمكن الطالب منها، أى يجب التركيز على الطالب في حد ذاته وليس الاهتمام

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

بالتنافس. ويجب على التقييم أن يسعى إلى تنمية مهارات التفكير العليا والفهم العميق والقدرة على حل المشكلات (Shepard, 2000) والتركيز على التعامل والتفاعل مع المشكلات الواقعية الحياتية حتى يكون التعلم قائم على معنى (Camp, 1993).

في ضوء العوامل الثلاثة السابقة، ظهرت الحاجة الملحة إلى ضرورة استحداث مداخل تقييمية بديلة تلبي التغير الحادث في مفهوم عملية التعلم وتغير فلسفة التقييم التي تهدف إلى الوصول إلى الإتقان والجودة والتميز Excellence في الأداء، ويبدو أن العامل الأكثر إلحاحاً على ضرورة إعادة النظر في الممارسات التقييمية الحالية هي جعل عملية التقييم كجزء متفاعل مع عملية التعلم والتدريس. ويبدو أن التطور في طرق التعلم وتبنى مدخل "التعلم الذاتي" و"التعلم عن بعد" جعلت الممارسات التقييمية الحالية لا تناسب هذه المداخل في عملية التعلم.

كل هذا أدى علماء القياس والتقويم إلى الحديث عن مداخل تقييمية تعتمد على قياس الأداء بصورة مباشرة وعملية وتعتمد على أن يكون للطلاب دوراً هاماً في تقييمهم وتصبح عملية التقييم عملية تفاعلية بين المعلم والمتعلم ومن أهم هذه المداخل هو "مدخل التقييم الحقيقي" (البورتفوليو) وقد يطلق عليه البعض الحقيقية التقييمية أو يطلق عليه البعض البورتفوليو. وفي عام ١٩٩٨ تبنت "٢١" ولاية من الولايات الأمريكية مدخل التقييم المعتمد على عينات الأداء (الحقائب) وأزاد هذا العدد إلى "٣٤" ولاية عام ١٩٩٩ (Heck & Crislip, 2001) وفي الجزء الأتى سيتم تناول هذا المدخل التقييمي بالتفصيل.

مدخل التقييم الحقيقي (البورتفوليو) Portfolio Assessment Approach

الحديث حول هذا المدخل يتم في ضوء المحاور الآتية:

أولاً: مفهوم التقييم الحقيقي

كما سبق ذكره سالفاً، أن الاتجاه الحديث في التقييم يؤكد على عدم الاعتماد ولو بصورة كلية على الاختبارات التقليدية ونتيجة للتطورات في مجال علم نفس التعلم وتغير فلسفة التقييم في القرن الحالى تم استحداث مداخل عديدة أهمها مدخل الحقيقة

فى التقييم. وهذا الشكل التقييمى ينتمى إلى تصنيف أوسع وهو التقييم المبنى على الأداء الأصيل أو الحقيقى Authentic Performance – based assessment (David et al., 2001, Paul & Dylan, 1998). وقد برزت فكرة استخدام المدخل الحقائقى فى عملية التقييم فى مؤتمر عقد فى ضاحية واشنطن (١٩٩١) وقال White (1994, pp 25-26) "وجدت نفسى مشاركاً فى مؤتمر طموح وفعال وناقش أفكار تربوية جديدة فى مجال التقييم القومى لتعلم طلاب الجامعة National Assessment of College Student Learning ومن الأفكار التربوية التى طرحت هو استخدام البورتفوليو (الحقيبة) فى عملية التقييم". والمدخل الحقيقى ليس فكرة جديدة بل انه مفهوم قديم، ففى منتصف الثمانينات من القرن الماضى استخدمه رواد التجديد التربوى فى عملية التقييم وهو ليس بطريقة منهجية بل إطار عام للاستجابة (Sommer, 1991) وليس جديد فى مجال التقييم، فاستخدمه الفنانون فى المدارس الفنية لإظهار قدراتهم وفى دقة وتناسق أعمالهم (Camp & Levine, 1991) ودمج بين الدليل الكمى والدليل الكيفى كمؤشرات لعملية التقييم (Nancy & Marker – Sehroer, 1992)، ويراه Wolf (1989) بأنه تقويم كلى وإنسانى وخالى من التحيز الثقافى (فى: Nancy & Marker – Sehroer, 1992). والبورتفوليو أو الحقيبة هو تجميع منظم وهادف لأعمال الطالب بحيث يعرض جهده وإنجازاته نحو تحقيق الأهداف الموضوعية وهذه الأعمال قد تكون أنشطة أو تكليفات يؤديها الطالب (Cole & Struyk 1997, Ina) Duffy et al., (1999) يعرفه (Dylan, 1994, Smith & Tillema, 2002) بأنه تجميع متعدد الأبعاد للأدلة بحيث تسمح للمعلم والمتعلم أن يبنوا صورة مفصلة ومنظمة عن عملية التعلم. ويراه Murphy (1994) بأنه مجموعة مختارة من أعمال التلميذ تتعلق بالمقرر كجزء مكمل له ويتم إنجاز هذه الأعمال فى الفصل أو خارجه لاستخدامه لأغراض التدريس والتقييم وتبرز Salvia & Yesseldyke (1995) دوره فى إصدار أحكام حول التلاميذ.

وفى ضوء التعريفات السابقة يتبين أن التقييم الحقيقى (البورتفوليو) هو مجموعة

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

من التكاليفات أو الأنشطة يقوم بها الطالب ويعتمد على نفسه بالقدر الأكبر وقد يكون إنجازها داخل الفصل أو خارجه. وهنا يبرز الهدف الأساسي من عملية التقييم الحقيقي وهو التأكيد على إيجابية الطالب ونشاطه في ضوء ما يقوم به تحت إشراف وتوجيه المعلم وفي ضوء ذلك يسعى إلى التفاعل مع عملية التدريس من خلال التوجيهات والمناقشات واللقاءات بين المعلمين والمتعلمين. وفي المدخل الحقيقي (البورتفوليو) يتم إعطاء التكاليفات على فترات زمنية مختلفة منذ بداية العام الدراسي حتى نهايته وهذا يسهل جمع بيانات عن الطالب حتى تسهل صناعة القرار. وتري Nacncy & Mark Schroer (1992) انه في ضوء المدخل الحقيقي بدلاً من أن نقول "نحن نريد أن نرى ماذا يفعل الطالب في امتحان الساعتين أو الثلاث ساعات" يمكن أن تتغير هذه المقولة إلى "ماذا يفعل الطالب على طول العام الدراسي".

ثانياً: مميزات المدخل الحقيقي في عملية التقييم:

يتميز استخدام المدخل الحقيقي في عملية التقييم بعدة مميزات أهمها انه يعطى صورة أكثر شمولية وموثوقية عن أداء الطالب، حيث يتميز بأنه متعدد المصادر Multisourced وفيه يتم الحصول على بيانات ومعلومات من خلال التكاليفات والأنشطة والمناقشات (Camp, 1993, Murphy, 1994). ويمدنا بأدلة عن وعي الطالب بالعمليات والاستراتيجيات المستخدمة في الوصول إلى الحلول وبذلك يوصف بأنه وثوقى (حقيقى) Authentic، وينمى مهارات التفكير العليا والقدرة على حل المشكلات من خلال الأنشطة والواجبات التى يتناولها المتعلم (Cole & Struyk, 1997, Davies & Michael 1999)، ويتطلب التكامل بين المعارف والقدرة على تطبيقها في مواقف الحياة لذلك يتميز بالتكاملية Integrated، ويزيد الاتصال بين المدرس والمتعلم مما يساعد على توجيه عملية التدريس (Notle, 1992) واستخدامه يزيد من كفاءة الذات (في: Cole & Struyk, 1997)، ويجعل الطالب مشاركاً بإيجابية في عملية التعلم (Duffy et al., 1999; Ina & Da Leen, 1999) ويجعل المتعلمين يتحملون مسئوليتهم نحو إنجاز أعمالهم لذلك يتميز بالوضوح Explicit، ويركز على التغذية الراجعة بصورة مستمرة (Smith &

(Tillema, 2001, White, 1994)، ويحدد نواحي القوة والضعف في أداء المتعلمين بصورة مستمرة حتى يسعى المتعلم إلى تحسين أدائه (Smith & Tillema, 2001)، ويتغلب على مشكلة التحيز للاختبارات Test Bias التي تعاني منها الاختبارات التقليدية من خلال الحكم الكيفي ومشاركة المتعلمين في عملية التقييم (Salvia & Yesseldyke, 1995, Wolf, 1989) ويكامل بين الممارسات التدريسية والتقييمية (Duffy et al., 1999, Ina & Daleen, 1999)، ومن خلال التقييم الحقايبى يصبح المتعلم مقدراً ذاتياً Self assessor لأعماله ويصبح مفهوم التقييم الذاتى واقع في عملية التعلم (White, 1994)، ويشجع التلاميذ على ممارسة مهارات ما وراء المعرفة للوعى بأدائهم (Duffy et al., 1999, Smith, 2000)، وينمى مهارات التعلم مدى الحياة Life Long Learning Skills وينمى مفهوم التعلم المتمركز حول المتعلم Student-centered Learning (Paul & Dylan, 1998) ويقوى التعلم المنظم ذاتياً Self-Regulated Learning (Pairs & Ayers, 1994)، ويصبح المتعلم موحهاً ذاتياً لتعلمه وتزيد من الاستقلالية في عملية التعلم مما ينمى الابتكار (Ina & Daleen, 1999).

وفي ضوء المميزات العديدة السابقة جراء استخدام مدخل الحقيقة في عملية التقييم، يتضح أن هذا المدخل يلبي النظرة الحديثة لعملية التعلم من إيجابية المتعلم ودوره الأساسى في بناء المادة المتعلمة، ويتضح أن أهمية استخدام هذا المدخل هو إتاحة الفرصة للتلاميذ للمشاركة في تقييم أعمالهم بصورة منطقية بعيدة عن الذاتية وذلك من خلال عملية تفاوضية بين المتعلم والمعلم.

ثالثاً: الانتقادات الموجهة إلى مدخل التقييم الحقايبى (البورتفوليو)

على الرغم من المميزات العديدة التي يتمتع بها استخدام الحقايب في عملية التقييم إلا انه يواجه انتقادات عديدة أهمها معارضة قوية من قبل المدرسين على استخدامه بدرجة كبيرة على أساس توسيع استخدامه يفقداهم ذاتهم وسيطرتهم في الفصل (White, 1994)، وعدم قدرة المدرسين على تصميم وتنفيذ النشاط التدريسي التقييمى القائم على مدخل الحقايب (Cole & Struyk, 1997)، وفي هذا

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التقييمية

الشأن توصل (1993) Calfee & Perfumo من خلال مسح قومي للمدارس التي تطبق نظام التقييم الحقايبى فى الولايات المتحدة إلى وجود فجوة بين ما يعرفه المدرسين عن التقييم الحقايبى وبين ممارستهم الفعلية (فى: Salvia & Yesseldyke, 1995). ومن أهم الانتقادات الموجه إلى هذا المدخل بأنه يأخذ وقتاً كبيراً من قبل المدرسين لتنظيم وتصحيح الأنشطة التقييمية خاصة إذا استخدم بجانب الاختبارات التقليدية. وهذا المدخل تنقصه المعيارية والخصائص السيكمترية خاصة الثبات والصدق، وذلك للأنشطة التقييمية خاصة إذا كانت هذه الأنشطة تختلف من تلميذ إلى آخر (Nancy & Mark Schroer, 1992)، وقد توصل (1993) Calfee & Perfumo إلى أن هذا المدخل يؤخذ عليه غياب إجراءات تقدير الثبات والصدق. ولكن من أهم المحددات لهذا المدخل هو كآى أداه تمدنا بالبيانات الكيفية فإنه توجد صعوبة فى تحليلها وعرضها وتنظيمها.

رابعاً: دينامية تصميم وتنفيذ المدخل الحقايبى (البورتفوليو) فى عملية التقييم:

لتحقيق أكبر فائدة من تطبيق المدخل الحقايبى فى عملية التقييم يجب أن تكون عملية تصميم وتنفيذ هذا المدخل على درجة عالية من الدقة والوضوح.

وفيما يلى تصور لتصميم وتنفيذ المدخل الحقايبى وقد استعان الباحث بأراء (Barton & Collins, 1997, Salvia & Yesseldyke, 1995, Shepard,) (2000) وفى هذا التصور يتضمن ثلاثة مداخل هى كالاتى:

١- مرحلة التخطيط Planning Stage

فى هذه المرحلة يجب تحديد الهدف من استخدام التقييم الحقايبى، فهل يستخدم للتكامل بين التدريس والتقييم، أى يتم الاعتماد عليه كوسيلة مكملة فى المقرر الدراسى ويستخدم بجانب الاختبارات التقليدية لتوثيق أداء التلاميذ. واستخدام مدخل الحقية فى عملية التقييم والتدريس يتطلب جهداً أكبر من كونه يستخدم بصورة مدعمة للاختبارات التقليدية. فى ضوء ذلك فإن مرحلة التخطيط تتطلب معرفة الآتى:

أ- تحديد الهدف من استخدام التقييم الحقائبي (البورتفوليو):

ويتنوع الهدف من استخدامه كالآتي:

- الاعتماد على البورتفوليو كعملية تدريسية وتقييمه في نفس الوقت.
- الاعتماد عليه لتوثيق أداء الطلاب بالمشكلات الواقعية.
- الاعتماد عليه بصورة تكلفات إضافية مرتبطة بأهداف المقرر.

ب- تحديد المشاريع أو الأنشطة المراد إنجازها.

قد تكون هذه التكاليفات أو الأنشطة عامة لكل المتعلمين مع إعطاء أنشطة أو تكلفات خاصة لكل طالب على حده، أى إنها تكون أنشطة عامة (غالباً قليلة) وبعضها خاصة (كثيرة)، فى تحديد الأنشطة أو التكاليفات لابد من إتاحة الحرية للطلاب فى اختيار الواجبات أو الأنشطة التى تتفق مع اهتماماته وذلك بالتشاور مع المعلم ويمكن أن يكون إنجاز بعض هذه الأنشطة بصورة تعاونية، حيث يمكن إعطاء تكليف يشترك فيه عدد من الطلاب وذلك بهدف تنمية القدرة على التعلم التعاونى ولابد من ارتباط هذه التكاليفات بأهداف المقرر فى حالة استخدامه كناحية تدريسية بقدر الإمكان.

ج- تحديد معايير التقييم:

بعد أن تتحدد التكاليفات والمشاريع لكل الطلاب، يقوم المعلم بتحديد معايير التقييم لهذه الأعمال ويتم الإعلان عن هذه المعايير لكل التلاميذ وتتنوع معايير التقييم وتشمل:

- الاستراتيجيات المستخدمة لإنجاز التكاليفات
- مدى توظيف الطريقة العلمية (خاصة لمراحل دراسية متقدمة، التعليم الجامعي مثلاً)
- مدى جودة العمل وارتباطه بالأهداف الموضوعية.
- حسن التنظيم فى كتابة المشروعات أو التقارير
- مدى جدية الباحث فى أداء الأنشطة
- مدى توظيف المراجع الذى استعان بها فى كتابة تقارير أنشطته.

٢- مرحلة التنفيذ Implementation

بعد أن يتم تحديد التكاليفات والأنشطة لكل متعلم، يبدأ المتعلم فى إنجاز الأنشطة المكلف بها ويتم ذلك فى المكتبة أو فى الفصل أو فى المنزل وخلال إنجاز الأنشطة لابد من إطلاع المعلم بصورة مستمرة على ما أنجزه الطالب ويطرح المعلم ملاحظاته أو تعديلاته وآرائه فى ضوء مناقشات مع طلابه ولابد أن يعبر التلاميذ عن رؤيتهم ووجه نظرهم وقد تختلف أو تتفق مع المعلم وفى هذه الخطوة ينمو لدى التلاميذ القدرة على التأمل الذاتى Self-Reflection وكما ذكرنا فلابد فى مرحلة التنفيذ أن يمارس التلاميذ مهارات التعلم التعاونى من خلال تكليف كل مجموعة بعمل نشاط مشترك. ويمكن أن يقوم كل طالب بعرض ما تم إنجازه وذلك من خلال عرض شفوى على زملائه فى مناقشات يديرها المعلم. وفى ضوء هذه المناقشات يقوم أقرانه بتقييم أعماله peer assessment

٣- مرحلة التقييم assessment stage

تنقسم عملية التقييم فى استخدام المدخل الحقيقى إلى مرحلتين:

أ- تقييم العمليات Processes Assessment

ويتم هذا خلال مرحلة تنفيذ أنشطة الحقيبة وذلك من خلال التغذية الراجعة إلى يتلقاها المتعلم من معلميه وزملائه خلال أداء أو إنجاز الأنشطة المنوط بها ولذلك يوصف التقييم الحقابى بأنه تقييم تكوينى Formative Assessment وهذا أهم ما يميز مدخل التقييم الحقابى عن الاختبارات التقليدية المعيارية وهذا يعمق التفاعل بين العملية التدريسية والعملية التقييمية حيث أن آراء المعلمين وتعليقاتهم لابد أن تكتب فى الصورة المبدئية للنشاط.

ب- تقييم الناتج النهائى Product Assessment

فى نهاية السنة الدراسية يقوم الطالب بتسليم الحقيبة المتضمنة على التكاليفات أو الأنشطة المكلف بها كل طالب ويتم التقييم النهائى فى ضوء المعايير السابق تحديدها فى مرحلة التخطيط. وفى هذه المرحلة تظهر قضية تحديد الحد الأدنى للأداء، إذ كيف تتم الحدود الدنيا المقبولة للأداء ليطلق على الطالب بأنه متمكن فى

أداء التكاليفات وهذه قضية محل مناقشة وجدل فى التراث السيكمترى ولكن يجب أن يكون تحديد الحدود الدنيا للأداء مسئولية مشتركة بين المعلمين والطلاب والآباء ويجب أن لا يكون للسياسيين دوراً فى وضع هذه المعايير وخاصة إذا كنا نهدف من التقييم ليس الحصول على شهادة بل التقييم من أجل التميز أو الجودة. وعند التقييم النهائى لابد أن يكون الطالب مشاركاً فى تقييم أعماله من خلال جلسات تفاوضية بينه وبين المعلم وهذا ما يميز المدخل الحقيقى عن الاختبارات التقليدية.

ويتم التقييم للأنشطة فى ضوء مقياس تدريجى من ثلاثة أو أربعة أو خمسة استجابات عن مدى مناسبة ودقة إنجاز الأعمال ويقوم المعلم بتصحيح الأنشطة فى ضوء هذا المقياس فى ضوء المعايير التى سبق تحديدها فى مرحلة التخطيط وقد يتم تصحيح هذه الاستمارة من قبل مصححين عديدين ثم يؤخذ متوسط تقدير المصححين، ويتم حساب ثبات بيانات هذه الاستمارة من خلال ثبات المقدرين.

Inter Rater Reliability

خامساً: قضايا سيكومترية مرتبطة بالتقييم الحقيقى:

يرى (Camp 1993) أن التكاليفات المنوط القيام بها من قبل المتعلمين بعضها مشتركة وبعضها تكاليفات خاصة بكل طالب على حدة، من ثم فإن المنهجية السيكمترية التقليدية فى حساب الخصائص السيكمترية غير مناسبة لاختلاف عينات الأعمال المكلف بها الطلاب خاصة لمؤشرى الثبات والصدق. وهذا ما دعا (Snow 1989) إلى المناشدة بتطوير نظرية قياسية جديدة. ولا يوجد اتفاق على كيفية حساب صدق وثبات البيانات المتحصل عليها من التقييم الحقيقى فى التراث (Ina & Daleen, 1999) ويمكن التغلب على مشكلة نقص الثبات للبيانات الناتجة من التقييم الحقايبى من خلال تعدد المصححين للتكاليفات التى يقوم بها الطلاب ثم تقدر ثبات المقدرين (Salvia & Yesseldyke, 1995, Ina & Dalean, 1999). فى حين يمكن الإدعاء بدون شك أن التقييم الحقيقى أكثر صدقاً من الاختبارات التقليدية لأنه يظهر أداء الطالب على عينة من الأنشطة وليس على قياس مفرد (White, 1994) ولزيادة صدق التقييم الحقيقى يجب أن تتعدد الأنشطة

والتكاليف بدلاً من الاعتماد على نشاط واحد (Murphy, 1994). وترى Ina & Daleen (1999) انه لتحسين صدق وثبات التقييم الحقيقي فلا بد من:

١- عند تسليم التلاميذ التكاليف لابد من تقدير مدى إنجاز هذه التكاليف بصورة مستمرة على مقياس تدريجي.

٢- تعدد المصححين في تصحيح التكاليف وحساب الارتباط بين تقدير المصححين حتى يمدنا بمؤشر صادق عن الثبات.

٣- المقارنة أو العلاقة بين الأداء في ضوء التقييم الحقيقي والأداء في ضوء طرق أخرى مثل الاختبارات التقليدية أو المقابلات مع الطلاب.

وكما تم ذكره أن التقييم الحقيقي (البورتفوليو) تنقصه المعيارية فالأنشطة تختلف من طالب إلى آخر بالتالي فمحتواه ليس موحد لكل الطلاب، بالتالي فالمقارنة بين التلاميذ تبدو عملية صعبة والقضية الأساسية في تصحيح الأنشطة والتكاليف التي تتضمنها الحقيقية هي الوقت المستغرق في تصحيح هذه التكاليف حيث يتطلب تصحيحها وقتاً كبيراً خاصة إذا كانت الحقيقة تتضمن أنشطة متعددة وتستخدم بجانب الاختبارات التقليدية.

وكما سبق ذكره في مرحلة التخطيط فإن التصحيح يتم في ضوء معايير معينة يتضمنها مقياس تدريجي ثم تعطى درجة لكل نشاط وتؤخذ الدرجة الكلية للأنشطة التي تتضمنها الحقيقة وبما أن كل نشاط يتضمن محتوى ومهارات واستراتيجيات مختلفة عن النشاط الآخر فإن البعض ينتقد هذا الإجراء ويعتبرون الدرجة الكلية في هذه الحالة ما هي إلا محصلة لأنشطة مختلفة ويشبهون هذا الإجراء بالجمع بين التفاح والبرتقال.

المشكلة وتحديها :

يبدو أن عملية التجديد في مجال التقييم عملية صعبة، كما أن المجتمع الف عملية التقييم في شكل الاختبارات الحالية وأي عملية تجديد يقابلها صدود وعدم تقبل من قبل المؤسسات التربوية والمجتمع معاً وذلك نتيجة للتراث الثقافي الذي يعتبر الامتحانات والدرجات هي عملية يجب أن لا يشملها أي تعديلات. وبدأ في نهاية

العقدين الأخيرين من القرن الماضي الحديث عن الممارسات التجديدية فى مجال التقييم وجاء ذلك كرد فعل لتغيير فلسفة عملية التعلم فى ضوء آراء أصحاب علم النفس البنويى الذين تبنا أن المعرفة هى عملية بنائية وليست مكتسبة، وأيضاً نتيجة لتغيير أو تطور فلسفة القياس والتقييم فى القرن الحالى، وانتشار طرق جديدة للتعلم مثل التعلم الذاتى والتعلم عن بعد، وأيضاً الانتقادات الموجهة إلى العلاقة بين مكونات العملية التعليمية ومنها عملية التدريس والتقييم، حيث فى ضوء الوضع التعليمى الحالى يتم عقد اختبارات بصفة دورية أو بصورة نهائية، ولكن الأداء فى هذه الاختبارات لا يعكس نظرة تقييمية لعملية التدريس، وهذا ما أدى بالمحدثين والمجدين التربويين وأصحاب القياس النفسى إلى إعادة النظرة فى العملية التقييمية الحالية واستحداث مداخل جديدة للتقييم قائمة على الأداء المباشر، وتعتمد هذه المداخل على إتاحة الفرصة لتضمين التلميذ فى تحديات ومشكلات المجتمع الحقيقية وتزيد من فعالية الطالب ليسهم بنشاطه فى المجتمع (Davis & Michael, 1999) وتعتبر أحد هذه المداخل التى لاقت قبولاً فى عملية التقييم هو التقييم الحقيقى (البورتفوليو)، والتراث البحثى الحالى لا يمدنا بدلائل أمبريقية عن فعالية هذا النوع من التقييم بصورة كافية وأن كانت معظم الدراسات والمقالات التى تناولت هذا المدخل التقييمى تناولته من ناحية التأصيل النظرى والآراء المبنية على الخبرة والتجربة. ورغم أن استخدام المدخل الحقيقى تميز بالحدثة نسبياً فى عملية التقييم وتؤكد White (1994) على أن التقييم القائم على تنوع الأدلة (البورتفوليو) بأنه حركة للمستقبل سواء بنا أو بدوننا وأن الاقتناع هو سبب للقبول بالمدخل فى عملية التقييم وليس التدعيم الأمبريقى وأن التراث الحالى لا يمدنا بأسس أمبريقية مبنية على التجريبية. ونادت بأهمية الحصول على دلائل تشير على أن استخدام التقييم الحقيقى (البورتفوليو) ينمى مهارات التعلم الذاتى.

وفى ضوء تفوق التقييم الحقيقى بالعديد من المميزات عن الاختبارات التقليدية (Cole & Struyk, 1997, Camp, 1993, Duffy et al., 1999, Ina & Daleen, 1999, Salvia & Yesseldyke, 1995)

وفى ضوء دراسة العلاقة بين الأداء على الاختبارات التقليدية والأداء فى ضوء

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

البورتفوليو توصل (Supovitz & Bernnan 1997) إلى أن العلاقة بين الأداء في الاختبارات التقليدية والأداء في ضوء المدخل الحقيقي تراوحت بين ٠,٥ إلى ٠,٦ (في: Heck & Crislip, 2001)، وتوصل (Heck & Crislip 2001) إلى أن العلاقة بين الأداء على الاختبارات التقليدية والأداء في ضوء المدخل الحقيقي لطلاب المرحلة الابتدائية في الصف الأول بلغت ٠,٥٦.

يمكن طرح التساؤلات الآتية:

١- ما اتجاهات طلاب الدراسات العليا (دبلوم مهني) نحو خصائص عملية التعلم

والمتعلم وعملية التدريس في ضوء استخدام المدخل الحقيقي في عملية التقييم؟

٢- ما الاتجاهات نحو خصائص عملية التعلم والمتعلم وعملية التدريس في ضوء

الممارسات التقييمية الحالية (الاختبارات التقليدية)؟

٣- هل تختلف الاتجاهات نحو خصائص عملية التعلم وعملية التدريس والمتعلم

باختلاف مدخل التقييم المستخدم (حقيقي - تقليدي)؟

٤- هل توجد علاقة بين التقييم الذاتي للأنشطة التقييمية والأداء على هذه الأنشطة

التي تتضمنها الحقيقية التقييمية؟

٥- ما مقدار العلاقة الارتباطية بين الأداء (الدرجات) على الأنشطة التي تتضمنها

الحقيقية والأداء في الاختبارات التقليدية؟

وقد أراد الباحث معرفة مدى موثوقية ومصداقية الدرجة المتحصل عليها من

الأداء في ضوء التقييم الحقيقي، والدرجة المتحصل عليها من الاختبارات التقليدية،

والدرجتان معاً في التنبؤ بتحصيل عينة الدراسة (دبلوم مهني) في المقررات

الدراسية الأربعة التي يدرسونها في الترم الثاني، حيث ترى (Valencia 1992)

أنه لا بد أن يتكامل التقييم الحقيقي مع التقييم الرسمي لتوثيق أداء الطلاب أفضل من

الاعتماد على معيار واحد في حد ذاته (في: Nancy & Marker - Schroer,

1992)، وفي ضوء ذلك تم طرح ثلاثة نماذج تنبؤية هم:

- الدرجة المتحصل عليها من تصحيح الأنشطة التي تتضمنها الحقيقية تصلح

للتنبؤ بتحصيل العينة في المقررات الأربعة في الترم الثاني.

- درجة الاختبارات التقليدية تصلح للتنبؤ بتحصيل الترم الثاني.
- الدرجة المتحصل عليها من تصحيح الأنشطة التي تتضمنها الحقيبة ودرجة الاختبارات التقليدية معاً تصلح للتنبؤ بالتحصيل في الترم الثاني. والباحث يطرح التساؤل الآتي:

٦- أى من هذه النماذج أكثر مطابقة واثقاً مع بيانات العينة؟ أو بكلمات أخرى أى من النماذج الثلاثة السابقة أكثر قدرة تنبؤية بالتحصيل النهائي في الترم الثاني؟
أهمية الدراسة:

- تتبع أهمية الدراسة الحالية في إطار منهجها الوصفي الارتباطي في الآتي:
- ١- توجيه النظر إلى أهمية مدخل التقييم الحقيقي (البورتفوليو) Portfolio assessment في عملية التقييم في مؤسستا التعليمية.
 - ٢- إبراز أهم العوامل التي ساعدت على إعادة النظر في الممارسات التقييمية الحالية.
 - ٣- معرفة جدوى استخدام وتطبيق المدخل التقييمي الحقيقي (البورتفوليو) خاصة لطلاب الدراسات العليا مقارنة بالاختبارات التقليدية.
- هدف الدراسة:**

- تهدف الدراسة إلى:
- ١- التعرف على الاتجاهات نحو خصائص عملية التعلم وعملية التدريس والمتعلم في ضوء مدخل التقييم الحقيقي وفي ضوء الاختبارات التقليدية الحالية.
 - ٢- مقارنة لاتجاهات نحو خصائص عملية التعلم وعملية التدريس والمتعلم في ضوء مدخل التقييم الحقيقي (البورتفوليو) في مقابل الاختبارات التقليدية.
 - ٣- التعرف على العلاقة بين الأداء في الأنشطة التي تتضمنها الحقيبة والتقييم الذاتي لأنشطة الحقيبة التقييمية.
 - ٤- معرفة العلاقة بين الأداء في ضوء الأنشطة التي تتضمنها الحقيبة والأداء في ضوء الاختبارات التقليدية.

٥ معرفة أى من نوعى التقييم سواء الحقيقى أو التقليدى أكثر إسهاما فى التنويع
بالتحصيل فى المقررات الدراسية الأخرى فى نفس الفترة الدراسية (الترم)

الطريقة والإجراءات

العينة:

انقسمت العينة إلى:

١- عينة الدراسة

تكونت عينة الدراسة من طلبة الدراسات العليا (دبلوم مهني في التربية) ٢٠٠٣
بكلية التربية بالإسماعيلية . جامعة قناة السويس وبلغ عددهم ١٢٦ طالب وطالبة
بمتوسط عمر ٢٢.٩٢ بانحراف معياري ٢.٣٦

٢- عينة صدق النتائج: Cross-validity sample

وتكونت هذه العينة من (٦٤) طالب من إجمالي "١٢٦" المكونة لعينة الدراسة
وقد تم اختيار هذه العينة عشوائيا من "١٢٦" طالب.

كيفية تطبيق مدخل الحقيقة أثناء عملية التدريس

هدف الباحث من استخدام الأنشطة التقييمية أن تكون مساعدة لعملية التقييم
التقليدية وهي استخدام الاختبارات الجماعية. ولم يطبق الباحث المدخل الحقيقي
بصورة تامة في عملية التدريس والتقييم. وقد عرض الباحث خلال تطبيق هذا
المدخل عدة تكاليف على طلبة الدبلوم المهني في بداية العام الدراسي وتم التنبيه
عليهم بأهمية عمل هذه التكاليف للأخذ بها في عملية التقييم وتكونت هذه الأنشطة
من:

١- إعداد اختبار تحصيلي في مجال تخصصه وتطبيقه على عينة من طلابه وجمع
بعض البيانات الأساسية عن هؤلاء الطلاب مثل النوع ومكان المعيشة وعمل
الأم، ثم إدخال درجات الاختبار التحصيلي والبيانات الوصفية في الكمبيوتر من
خلال البرنامج الإحصائي SPSS، ثم صياغة أكبر عدد من الأسئلة البحثية
يوظف الدارس فيها أكبر عدد من الاختبارات الإحصائية الوصفية والاستدلالية
وأيضا الاختبارات الإحصائية البارامترية واللابارامترية، ثم إجراء الحسابات

اليدوية لهذه الأسئلة البحثية من خلال خطوات اختبارات الفروض، ثم إجراء هذه الاختبارات من خلال البرنامج الإحصائي SPSS.

٢- البحث في عدة قواميس عن بعض الكلمات والمفاهيم في مجال القياس والإحصاء وتم التنبيه على الدارسين بأهمية البحث في أكثر من قاموس قد يكون في مجال علم النفس أو في أى مجال يرغبه ومن أهم هذه المفاهيم (الإحصاء - الصدق العاملي - الثبات - التقويم - المتغير - العينة - فرض البحث -)

٣- أن يقوم الدارس بتوظيف مفاهيم ومبادئ الإحصاء الوصفي والاستدلالي مثل فرض البحث والنموذج الإحصائي والدلالة العملية وقوة الاختبار على دراسة منشورة في أحد المجلات أو المؤتمرات أو رسائل الماجستير والدكتوراه.

٤- تم إعطاء بحث يقوم به عدد من الدارسين ولايزيد عددهم عن ستة دارسين ضمن عمل جماعي تعاوني ويقوم هؤلاء الدارسين بالبحث عن الموضوع المكلفين به على الإنترنت من خلال مواقع عديدة، ثم يقوموا بالبحث في المكتبة عن نفس الموضوع ويقوموا بأعداد تقرير عن الموضوع المكلفين به ومن أهم الموضوعات المكلف بها الطلبة (حجم التأثير - قوة الاختبار - الصدق العاملي - تحيز أسئلة الاختبار.....).

في ضوء ذلك فإن الحقيقة أو ملف الطالب يتكون من أربعة تكاليفات أو أنشطة. وقد تم طرح هذه الأنشطة أو التكاليفات في بداية العام الدراسي وكان يقوم الدارس باستشارة أستاذ المقرر من حين لآخر من خلال مناقشة في المحاضرة أو في الساعات المكتبية المخصصة لهذا الغرض وكان أستاذ المقرر يعطي بعض الآراء والمقترحات لأداء هذه التكاليفات لكل طالب على حدة، ثم يقوم الدارس بالأخذ بهذه المقترحات أو الآراء حتى يسعى إلى تحسين أدائه في هذه التكاليفات. وتم التركيز على تقييم هذه التكاليفات في نهاية العام الدراسي وأن تمت عملية المراجعة المستمرة خلال الترم وخاصة في البحث المكلف به مجموعة من الطلاب وحيث يقوم كل مجموعة بعرض ما أنجزته في هذا العمل التعاوني في بداية كل محاضرة ويستغرق

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

عرضهم ما يقرب من نصف ساعة. وركز الباحث على إعطاء تقييمات للنواتج النهائية وذلك في ضوء استمارة تقييم لهذه الأنشطة. وفي نهاية الترم تم إعطاء الطالب استمارة للتقييم الذاتي للمقرر والأنشطة المصاحبة له.

أدوات الدراسة:

تم تطبيق عدة أدوات لإتمام هذه الدراسة هي كالاتى:

أولاً: استمارة التقييم الذاتى لموضوعات مقرر الإحصاء والأنشطة المصاحبة له:

تكونت الاستمارة فى بدايتها من ٢٢ عنصر تقيس تقييم الدارسين لأنفسهم فى المهام والتكاليفات الأربعة السابق توضيحهم فى بند (كيفية تطبيق مدخل الحقيبة أثناء عملية التدريس) وهى التكاليف الخاص ببناء اختبار تحصيلي وتكليف البحث فى الإنترنت وتكليف البحث فى عدة قواميس وتكليف تقويم دراسة ما بالإضافة إلى عناصر تقيس تقييم الدارسين لأنفسهم فى محتوى المقرر ككل ولكن فى ضوء حذف "١١" عنصر أصبحت الاستمارة مكونة من "١١" عنصر يتعلق بعضها بالتقييم الذاتى بالنشاط المتعلق بالبحث فى الإنترنت وبعضها يتعلق بالاختبار التحصيلي وبعضها يتعلق بتقييم المقرر ككل.

وقد تم بناء هذه الاستمارة فى ضوء مقياس ليكرت وذلك بإعطاء العبارة وأمامها خمسة بدائل (بدرجة كبيرة جداً ^(٥) - بدرجة كبيرة ^(٤) - بدرجة متوسطة ^(٣) - بدرجة قليلة ^(٢) - بدرجة قليلة جداً ^(١)). تكونت هذه الاستمارة فى بدايتها من (٢٢) عنصر وبحساب معامل الثبات باستخدام المعامل الفاء، اتضح أن ثباتها (٠,٥٣) وتم حذف عدد من العناصر حتى أصبحت الاستمارة مكونة من (١١) عنصر وحقيقة يبدو ان انخفاض معامل الثبات يعود إلى الأبعاد العديدة التى تكونت منها الاستمارة فتم الاعتماد على حساب الخصائص السيكمترية (الثبات والصدق) للأبعاد على حدة. وتكونت الاستمارة من ثلاثة أبعاد هى كالاتى:

١- التقييم الذاتى لنشاط البحث المتعلق بالإنترنت:

وتكون هذا البعد من ثلاثة عناصر حيث بلغ معامل الثبات الفأ (٠,٧٨) لعينة مكونة من ١٢٢ دارس وفيما يلى ثبات البعد بعد حذف كل عنصر ومعامل الارتباط المصحح بين العنصر والمجموع الكلى للبعد محذوف منه درجة العنصر.

جدول (١): معامل الثبات للبعد بعد حذف العنصر ومعامل الارتباط المصحح

معامل الارتباط المصحح	معامل الفا بعد حذف السؤال	العنصر
0.69	0.63	Z11*
0.46	0.78	Z12
0.74	0.58	Z13

* Z مسمى عناصر الاستمارة فى البرنامج

ويتضح من الجدول (١) أن معامل الثبات بعد حذف العنصر انخفضت قيمته لكل العناصر عن قيمة ثبات البعد ككل وهى (٠,٧٨)، مما يدل على أهمية هذه العناصر فى البعد، كما تبين العلاقة الموجبة بين درجة العنصر والمجموع الكلى للبعد محذوف منه درجة العنصر على الرغم من حذف درجة العنصر من المجموع الكلى ولم يعتمد الباحث على حساب معامل الارتباط بين المجموع الكلى ودرجة العنصر لأن هذا المعامل يحدث له تضخم Inflation لأن درجة السؤال تكون ضمن المجموع الكلى.

٢- التقويم الذاتى للنشاط المتعلق بالبحث الميدانى (إعداد اختبار تحصيلي)

وتوظيف نواتج برنامج SPSS

تكون هذا البعد من أربعة عناصر وبلغ معامل الثبات الفا (٠,٧٦) لعينة ١٢٣ وفيما يلى ثبات البعد بعد حذف العنصر ومعامل الارتباط المصحح.

جدول (٢): معامل الثبات للبعد بعد حذف

العنصر ومعامل الارتباط المصحح للبعد الثانى

معامل الارتباط المصحح	معامل الفا بعد حذف العنصر	العنصر
0.60	0.69	Z10
0.56	0.67	Z9
0.47	0.71	Z8
0.64	0.75	Z7

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التصيلية

تبين من جدول (٢) أن معامل الثبات للبعد بعد حذف العنصر انخفض قيمته عن معامل الثبات للبعد ككل (٠,٧٦)، ويدل هذا على أهمية هذه العناصر في البعد، كما أن ارتباط العناصر بالمجموع الكلي للبعد بعد حذف العنصر من البعد ارتباطاً موجباً مما يؤكد على أهمية هذه العناصر في البعد.

٣- التقييم الذاتي للمقرر بصفة عامة:

وتكون هذا البعد من أربعة عناصر وبلغ معامل الثبات $\alpha = ٠,٥٩$ لعينة مكونة من ١٢٣ طالب.

جدول (٣): معامل الثبات للبعد بعد حذف

العنصر ومعامل الارتباط المصحح للعنصر

العنصر	معامل الفا بعد حذف العنصر	معامل الارتباط المصحح
Z14	0.47	0.42
Z15	0.50	0.40
Z20	0.54	0.34
Z19	0.56	0.32

تبين من جدول (٣) أن قيمة معامل الثبات للبعد بعد حذف العنصر انخفض عن قيمة معامل الثبات للبعد ككل وأيضاً تبين أن قيمة معامل الارتباط المصحح موجبة جزئية إلى حد ما مما يدل على أهمية هذه العناصر في البعد.

- صدق الاستمارة: الصدق العاملي التوكيدي لاستمارة التقييم الذاتي المكونة من ثلاثة أبعاد:

قام الباحث بحساب الصدق العاملي التوكيدي باستخدام برنامج 8 Lisrel من إعداد (Joreskog & Sorbon 1993) للأداة المكونة من "١١" عنصر موزعة على ثلاثة أبعاد وتم الحصول على قيم التشبع (٨) للمتغيرات على الأبعاد الخاصة بها وذلك بحساب التحليل العاملي التوكيدي.

جدول (٤): قيم تشبعات العناصر على أبعادها وقيمة (T) المناظرة

العنصر	العامل الأول		العامل الثاني		العامل الثالث	
	التشبع λ	قيمة T	λ	T	λ	T
Z11	0.88	8.92				
Z12	0.47	6.92				
Z13	0.91	8.96				
Z10			0.77	10.56		
Z9			0.82	10.78		
Z8			0.59	9.05		
Z7			0.52	8.23		
Z14					0.52	5.94
Z15					0.62	6.53
Z20					0.35	4.74
Z19					0.53	6.26

واتضح أن كل العناصر تشبعت على الأبعاد الخاصة بها، حيث زادت قيمة T عن ١,٩٦ وهى الحد الأدنى لقبول تشبع العنصر على البعد. وأتضح أن نموذج التحليل العاملى للاستمارة بأبعادها الثلاثة وعناصرها "١١" أثبتت حسن مطابقة للبيانات المجمعة من العينة وظهر ذلك من خلال مؤشرات حسن المطابقة حيث بلغت قيمة χ^2 لحسن المطابقة ٤٠,٠٠ ($P = 0.49$) وهى غير دالة إحصائياً عند ٠,٠٥ مما يدل أن النموذج متطابق مع البيانات وظهر ذلك التطابق من خلال مؤشرات حسن المطابقة الأخرى، وهى $RMSEA = 0.0$ حيث انخفضت قيمتها عن ٠,٠٥ ومؤشر GFI ومؤشر AGFI ومؤشر NNFI ومؤشر CFI حيث بلغت قيمهم على التوالي ٠,٩٧ و ٠,٩٤ و ١,٠٠ و ١,٠٠ أى أن قيمة هذه المؤشرات زادت عن نقطة القطع ٠,٩ وهى الحد الأدنى لقبول مطابق النموذج

* يمكن الرجوع إلى دراسة للباحث بعنوان "أداء مؤشرات حسن المطابقة لتقويم نموذج المعادلة البنائية" مقدمة إلى المؤتمر السلوى العشرون لعلم النفس فى مصر والثانى عشر العربى فى الفترة ١٩-٢١ يناير ٢٠٠٤.

التقييم الحقيقي (البورتفوليو) كمدخل تكاملى للاختبارات التحصيلية

وهذا يؤكد على أن الاستمارة تتمتع بخصائص سيكومترية عالية من حيث ثباتها وصدقها.

ثانياً: مقياس الاتجاه نحو خصائص عملية التعلم وعملية التدريس والمتعلم

تم بناء المقياس فى ضوء طريقة ليكرت وهى أن توضع أمام كل عبارة خمسة بدائل للاستجابة وهى بدرجة كبيرة جداً^(٥) - بدرجة كبيرة^(٤) - بدرجة متوسطة^(٣) - بدرجة قليلة^(٢) - بدرجة قليلة جداً^(١). وتكونت هذه الأداة من (١٧) عنصر لقياس الاتجاه نحو خصائص التعلم وخصائص عملية التدريس وخصائص المتعلم وتم تقديمها للدارس لقياس الاتجاه نحو هذه الخصائص فى ضوء مدخل التقييم الحقيقى وقياس الاتجاه نحو نفس الخصائص فى ضوء التقييم التقليدى الحالى.

١- الخصائص السيكومترية لاستجابات العينة على الأداة فى ضوء مدخل التقييم الحقيقى (البورتفوليو):

بلغ معامل الثبات الفا للأداة المكونة من "١٧" عنصر (٠,٨٦) لعدد أفراد العينة ١١٨ وتكونت الأداة من ثلاثة أبعاد كالاتى:

البعد الأول: خصائص عملية التعلم

وهذا البعد تكون من عنصرين وبلغ معامل الثبات ٠,٦٦ لعينة ١٢٥ وفيما يلى ثبات البعد بعد حذف العنصر ومعامل الارتباط المصحح.

جدول (٥): ثبات البعد بعد حذف العنصر ومعامل الارتباط المصحح

العنصر	معامل الفا بعد حذف العنصر	معامل الارتباط المصحح
P14*	0.50	0.66
P15	0.50	0.54

* P مسمى العنصر فى برنامج الكمبيوتر

يتضح من جدول (٥) أن قيمة معامل الثبات للبعد بعد حذف العنصر الأول وبعد حذف العنصر الثانى انخفضت عن قيمة معامل الثبات للبعد ككل (٠,٦٦) وأن قيمة معامل الارتباط المصحح عالية، مما يدل على أهمية العنصرين كمكونين للبعد.

البعد الثاني: خصائص عملية التدريس

تكون هذا البعد من (خمس) عناصر وبلغ معامل الثبات الفا (٠,٦٧) لعدد أفراد العينة ١٢٣

جدول (٦): ثبات البعد بعد حذف العنصر ومعامل الارتباط المصحح

العنصر	معامل الفا بعد حذف العنصر	معامل الارتباط المصحح
P5	0.64	0.37
P6	0.61	0.44
P7	0.56	0.54
P8	0.63	0.37
P17	0.63	0.38

يتضح من جدول (٦) أن ثبات البعد بعد حذف العنصر انخفض عن ثبات الأداة ككل (٠,٦٧) كما أن قيمة معامل الارتباط المصحح يعتبر مقبولة وهذا يؤكد أهمية العناصر الخمسة المكونة لبعد خصائص عملية التدريس.

البعد الثالث: خصائص المتعلم

يتكون هذا البعد من (تسعة) عناصر وبلغ معامل الثبات الفا (٠,٧٨) لعدد أفراد العينة ١١٩.

جدول (٧): ثبات بعد خصائص المتعلم بعد

حذف العنصر ومعامل الارتباط المصحح

العنصر	الفا بعد حذف العنصر	معامل الارتباط المصحح
P1	0.76	0.42
P2	0.77	0.40
P3	0.77	0.38
P4	0.75	0.54
P9	0.77	0.40
P10	0.77	0.39
P11	0.74	0.58
P12	0.75	0.49
P13	0.76	0.43
P16	0.76	0.49

التقييم الحقيقي (البورنوليوي) كمدخل تكاملي للاختبارات التحصيلية

الصدق العاملي التوكيدي لمقياس الاتجاه نحو خصائص عملية التعلم والتدريس والمتعلم في ضوء التقييم الحقيقي:

أجرى الباحث التحليل العاملي التوكيدي باستخدام برنامج (8) Lisrel وذلك للتأكد من البنية العاملية للمقياس.

جدول (٨): تشبعات عناصر مقياس الاتجاه نحو خصائص

التعلم وخصائص التدريس وخصائص المتعلم في ضوء التقييم الحقيقي

العنصر	البعد الأول		البعد الثاني		البعد الثالث	
	λ	T	λ	T	λ	T
P1	0.52	12.10				
P2	0.42	9.85				
P3	0.30	7.16				
P4	0.60	13.74				
P9	0.69	15.45				
P10	0.47	11.07				
P11	0.63	19.28				
P12	0.57	13.40				
P13	0.48	11.32				
P16	0.59	13.40				
P5			0.59	11.66		
P6			0.70	13.31		
P7			0.43	8.84		
P8			0.41	8.37		
P17			0.76	9.79		
P14					0.71	9.57
P15					0.74	9.68

تبين من الجدول السابق أن عناصر البعد الأول تشبعت على العامل الأول وأيضاً عناصر البعد الثاني تشبعت على العامل الثاني وعناصر البعد الثالث تشبعت على العامل الثالث وأتضح أن قيمة التشبع كانت عالية وذلك من خلال كبر قيمة T حيث زادت عن ١,٩٦ زيادة كبيرة مما يؤكد على أن البنية العاملية لهذا المقياس على درجة عالية من الثقة. وأنعكس ذلك في قيم مؤشرات حسن المطابقة حيث بلغت قيمة RMSEA = ٠,٠٥٥ وقيمة GFI = ٠,٩٤ ومؤشر AGFI = ٠,٩٢ ومؤشر NNFI = ٠,٩٧ ومؤشر CFI = ٠,٩٧ أي أن قيمة المؤشرات زادت عن قيمة ٠,٩ وهي الحد الأدنى لقبول معقولة ومقبولية ومطابقة النموذج.

الخصائص السيكمترية لاستجابات العينة على المقياس في ضوء التقييم التقليدي ونفس الأداة السابقة (١٧) عنصر، بلغ معامل الثبات ألفا (٠,٩٠) لحجم عينة ١١٤ وتبين أن معامل الثبات للبعد الأول خصائص التعلم المكون من عنصرين ٠,٧٢ (N=125) ومعامل الثبات ألفا للبعد الثاني خصائص عملية التدريس ٠,٧٤ (N=121) ومعامل الثبات للبعد الثالث خصائص المتعلم ٠,٨٣ (N=117). التحليل العاملي التوكيدي لاستجابات أفراد العينة على المقياس في ضوء التقييم التقليدي

جدول (٩): تشبعات عناصر مقياس الاتجاه نحو خصائص

التعلم والتدريس والمتعلم في ضوء التقييم التقليدي

العنصر	البعد الأول		البعد الثاني		البعد الثالث	
	λ	T	λ	T	λ	T
T1	0.57	13.71				
T2	0.43	10.53				
T3	0.27	6.73				
T4	0.70	16.32				
T9	0.64	15.62				
T10	0.70	16.86				
T11	0.55	13.40				
T12	0.60	14.46				
T13	0.46	11.43				
T16	0.69	16.29				
T5			0.64	13.35		
T6			0.63	13.11		
T7			0.57	12.46		
T8			0.31	7.38		
T17			0.63	13.15		
T14					0.79	12.24
T15					0.86	12.34

تبين من الجدول السابق أن عناصر البعد الأول تشبعت على العامل الأول وهكذا عناصر البعد الثاني وعناصر البعد الثالث وتميزت قيم التشبع بأنها عالية جداً وأنعكس ذلك في قيمة T. وتبين أن مؤشرات حسن المطابقة لهذا النموذج تتقارب مع مؤشرات حسن المطابقة لاستجابات العينة على نفس الأداة في ضوء التقييم الحقيقي وبلغت مؤشرات حسن المطابقة للنموذج الحالي $RMSEA = 0.058$ و $GFI = 0.95$ و $AGFI = 0.93$ و $NNFI = 0.98$ و $CFI = 0.98$. وهذه مؤشرات مطابقة تعكس حسن مطابقة جيدة للنموذج مع البيانات.

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

ثالثاً: استمارة تقييم الأنشطة أو التكاليف المتضمنة في الحقيبة

تم إعداد استمارة لتقييم الأنشطة والتكاليف السابقة ذكرها سالفاً بالإضافة إلى تقييم مدى الجودة والتنظيم والجدية في عمل هذه الأنشطة وتكونت هذه الاستمارة من "٢٠" عنصر وبلغ معامل ثباتها ٠,٨٩ (N = 118) والبعد الأول الخاص بتقييم النشاط المتضمن إعداد اختبار تحصيلي وتوظيف الأساليب الإحصائية والمعالجة الإحصائية باستخدام برنامج SPSS تكون من (تسعة) عناصر وبلغ معامل ثباته ٠,٨٦ (N=118) والبعد الثاني الخاص بتقييم النشاط المتضمن أداء بحث مصادرة من الإنترنت والمكتبة شمل على (خمس) عناصر وبلغ معامل الثبات ألفا 0.73 (N = 120) والبعد الثالث الخاص بتقييم البحث عن بعض المفاهيم في مجال القياس والإحصاء في عدة قواميس تكون من عنصرين وبلغ معامل الثبات ٠,٧٧ (N = 120) والبعد الرابع الخاص بتقييم توظيف ما درسه من أساليب ومفاهيم إحصائية على دراسة ما شمل على عنصرين وبلغ معامل ثباته ٠,٨٠ (N = 120) والبعد الخامس الخاص بتقييم الجودة والتنظيم والجدية شمل على عنصرين وبلغ معامل ثباته ٠,٩١ (N = 120).

التحليل العاملي التوكيدي لاستمارة تقييم الأنشطة والتكاليف المتضمنة في الحقيبة

جدول (١٠): تشيعات عناصر استمارة تقييم الأنشطة على العوامل الخمسة

العنصر	العامل الأول		العامل الثاني		العنصر	العامل الثالث		العامل الرابع		العامل الخامس	
	λ	T	λ	T		λ	T	λ	T	λ	T
AC1*	0.63	16.22			AC15	0.76	10.27				
AC2	0.89	21.62			AC16	0.74	10.57				
AC3	0.76	19.25			AC17			0.93	12.63		
AC4	0.38	10.09			AC18			0.75	12.63		
AC5	0.92	22.18			AC19					0.88	15.19
AC6	0.57	14.89			AC20					0.91	14.83
AC7	0.78	19.69									
AC8	0.38	10.07									
AC9	0.70	10.44									
AC10			0.70	10.44							
AC11			0.71	10.55							
AC12			0.73	10.37							
AC13			0.71	10.55							
AC14			0.66	9.86							

* مسمى العنصر في البرنامج

تبين من الجدول السابق أن كل العناصر تشبعت على العامل الخامس بها وتبين

أن الصدق العاملي لهذا الاستمارة تميز بالدقة والمطابقة مع بيانات العينة حيث بلغت قيمة $RMSEA = 0.025$ أى إنها انخفضت عن ٠,٠٥ وبلغت قيمة $GFI = 0.95$ و $AGFI = 0.95$ و $NNFI = 0.99$ و $CFI = 1.00$ وهذه قيم عالية جداً لحسن مطابقة أبعاد الاستمارة الخمسة مع بيانات العينة.

رابعاً: الاختبار التحصيلي:

تم إعداد الاختبار التحصيلي في مقرر الإحصاء التربوي. تكون الاختبار التحصيلي من ٤٠ سؤال وبلغ معامل الثبات ألفا ٠,٧٦ وتكون الاختبار التحصيلي من ستة أبعاد حسب العمليات العقلية الستة وهي كالآتي:

- مستوى التذكر: تم تمثيل هذا البعد من عشرة أسئلة بمعامل ثبات ٠,٦٧ وتم حذف سؤال فأصبح عدد أسئلة البعد "تسعة" عناصر بمعامل ثبات ٠,٧٣.
 - مستوى الفهم: تم تمثيله بـ (ثمانية) أسئلة وبلغ معامل الثبات ٠,٥٤، ثم حذف سؤالين فأصبح عدد عناصر "ستة" بمعامل ثبات ٠,٥٨.
 - مستوى التطبيق: تكون هذا البعد من (ستة) عناصر وبلغ معامل الثبات ألفا ٠,٢٧، ثم حذف ثلاثة أسئلة فأصبح عدد أسئلته ثلاثة وبلغ معامل ثباته ٠,٧٦.
 - مستوى التحليل: تكون من (سبعة) أسئلة وبلغ ثباته ٠,٧٦ وتم حذف ثلاثة عناصر وأصبح ثباته ٠,٨٥.
 - مستوى التركيب: تكون من (خمسة) أسئلة بمعامل ثبات ٠,٧٤ وبعد حذف سؤال أصبح ثباته ٠,٨٠.
 - مستوى التقويم تكون من (٤) أسئلة وبلغ معامل ثباته ٠,٧٣.
- في ضوء ذلك فإن عدد الأسئلة المكونة للاختبار التحصيلي (٣٠) عنصر موزعة على العمليات العقلية الستة وبلغ معامل ثباته ٠,٨٨ لعينة مكونة من ٦٤ طالب وهذا الاختبار طبق في امتحان أعمال السنة وقام الباحث بإجراء التحليل العاملي التوكيدي للتأكد من الصدق العاملي لعناصر الاختبار موزعة على الأبعاد الستة.

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

جدول (١١): تشبعات أسئلة الاختبار التحصيلي على العوامل الستة

العنصر	التذكر	الفهم	التطبيق	التحليل	التركيب	التقويم
X1	0.46					
X3	0.46					
X4	0.48					
X8	0.43					
X12	0.63					
X13	0.46					
X14	0.54					
X15	0.49					
X17	0.42					
X2		0.54				
X6		0.43				
X9		0.42				
X18		0.41				
X19		0.56				
X21		0.34				
X29			0.30			
X32			0.33			
X31			0.30			
X33				0.54		
X35				0.47		
X36				0.43		
X34				0.37		
X25					0.93	
X27					0.34	
X28					0.57	
X26					0.72	
X37						0.60
X38						0.34
X39						0.71
X40						0.51

ويتضح من جدول (١١) أن قيم التشبع زادت عن ٠,٣ ولم يطبع البرنامج قيم T لمعرفة الدلالة الإحصائية للتشبعات وذلك لصغر حجم العينة. وكذلك لم يطبع البرنامج مؤشرات حسن المطابقة الخاصة بالنموذج.

خامساً: بيانات الأرشيف التى تمثل تحصيل الطلاب فى امتحانات نهاية الترم كما هى مدونة فى الكنترول شيت

قام الباحث بالحصول على درجات الطلاب فى المواد الخمسة التى يؤدى فيها الامتحان فى نهاية الترم الثانى وكانت المقررات الخمسة هى: الإحصاء التربوى بمتوسط تحصيل ٤٩ وانحراف معيارى ٧,٠٩، ومقرر التعليم العلاجى بمتوسط ٥٤,٤٨ وانحراف معيارى ٦,٠٢، ومقرر استراتيجيات تدريس بمتوسط ٤٠,٣٠ وانحراف معيارى ٨,٦٢ ومقرر تربية تعويضية بمتوسط ٥٩,٤١ وانحراف معيارى ٥,٦٦، ومقرر إدارة مدرسية بمتوسط ٣٦,٤٨ وانحراف معيارى ٥,٨٣، وبلغ متوسط تحصيل مجموع المقررات الأربعة (تعليم علاجى وإدارة مدرسية وتربية تعويضية واستراتيجيات تدريس) ١٩٠,٨٢ وانحراف معيارى ٢١,٠٤.

تطبيق أدوات الدراسة

تم تطبيق أدوات الدراسة وهى استمارة التقييم الذاتى والاتجاه نحو خصائص التعلم والمتعلم والتدريس فى ضوء التقييم الحقيقى والتقييم التقليدى قبل امتحان نهاية الترم بأسبوعين، كما تم تطبيق الاختبار التحصيلى المتضمن العمليات العقلية الستة فى امتحان أعمال السنة قبل امتحان نهاية الترم بأسبوع وذلك على عينة مكونة من "٦٤" (تمثل عينة صدق النتائج) بحيث تم تطبيق الامتحان النهائى فى نهاية الترم (العينة البحث ١٢٦)، وتم تصحيح الأنشطة أو التكاليفات لكل طالب فى ضوء استمارة التقييم للأنشطة بعد تسليمها فى نهاية الترم.

استراتيجية التحليل الإحصائى:

أعتمد الباحث فى الإجابة على الأسئلة من الأول حتى الرابع على العينة الكلية للدراسة المكونة من "١٢٦" طالب، واعتمد على الإجابة على الأسئلة من الخامس حتى السابع على عينة الدراسة وهى عينة الدراسة المكونة من ١٢٦ وعينة صدق النتائج (التقييم) المكونة من (٦٤). وكما سبق توضيحه أن عينة الدراسة طبق عليها الاختبار التحصيلى فى نهاية الترم ولكن عينة صدق النتائج طبق عليها امتحان أعمال السنة وطبق عليهم نفس الأدوات الأخرى وفى حالة السؤال السادس الخاص

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

بالتنبؤ بالتحصيل النهائي (نهاية الترم) في ضوء التقييم الحقيقي والتقييم التقليدي، واستخدم الباحث مؤشرات حسن المطابقة للمقارنة بين نماذج التنبؤ وهي مؤشر مربع معامل الارتباط R^2 ومؤشر مربع معامل الارتباط المصحح R^2 ومؤشر متوسط مربعات الخطأ MSE ومؤشر مالو (1970) Cp.

النتائج والمناقشة

التساؤل الأول:

للإجابة على التساؤل الأول الخاص بمعرفة اتجاهات طلاب العينة نحو خصائص عملية التعلم وخصائص عملية التدريس وخصائص المتعلم في ضوء التقييم الحقيقي قام الباحث بإجراء التوزيعات التكرارية لمعرفة تكرارات الاستجابات ونسبتها المئوية لكل عبارة وفيما يلي التكرارات والنسبة المئوية لكل عبارة.

جدول (١٢): التكرارات والنسب المئوية لاستجابات عبارات مقياس الاتجاه

نحو خصائص التعلم والتدريس والمتعلم في ضوء التقييم الحقيقي

٢	الخاصية	الاستجابة									
		درجة كبيرة جداً		درجة كبيرة		درجة متوسطة		درجة قليلة		درجة قليلة جداً	
		النسبة %	التكرار	النسبة %	التكرار	النسبة %	التكرار	النسبة %	التكرار	النسبة %	التكرار
١	يظهر كدراسي وأكاديمي بصورة جيدة	٣٩	٣١,٢	٥٧	٤٥,٦	٢٤	١٩,٢	٤	٣,٢	١	٠,٨
٢	أبذل جهداً كبيراً للحصول على المعلومات	٥٤	٤٣,٥	٤٤	٣٥,٥	١٩	١٥,٣	٧	٥,٦	٠	٠,٠
٣	أضد على نفسي للحصول على المادة المطلوبة	٤٩	٣٩,٥	٤٨	٣٨,٧	٢٢	١٧,٧	٤	٣,٢	١	٠,٨
٤	أبذل قصارى جهدي للحصول على المادة المطلوبة	٧٢	٥٨,٥	٤١	٣٣,٣	٨	٦,٥	٢	١,٦	٠	٠,٠
٥	أبذل قصارى جهدي للحصول على المادة المطلوبة	٥٢	٤١,٦	٥٤	٤٣,٢	١٤	١١,٢	٤	٣,٢	١	٠,٨
٦	أبذل قصارى جهدي للحصول على المادة المطلوبة	٤٤	٣٥,٢	٥٧	٤٥,٦	٢١	١٦,٨	٣	٢,٤	٠	٠,٠
٧	أبذل قصارى جهدي للحصول على المادة المطلوبة	٥٦	٤٤,٨	٥١	٤٠,٨	١٤	١١,٢	٣	٢,٤	٠	٠,٠
٨	أبذل قصارى جهدي للحصول على المادة المطلوبة	٥٥	٤٤	٥٢	٤١,٦	١٦	١٢,٨	٢	١,٦	٠	٠,٠
٩	أبذل قصارى جهدي للحصول على المادة المطلوبة	٤١	٣٣,٦	٤٨	٣٨,٣	٢٤	١٩,٢	٦	٤,٨	٣	٢,٤
١٠	أبذل قصارى جهدي للحصول على المادة المطلوبة	٤٧	٣٧,٦	٦٠	٤٨	١٦	١٢,٨	٢	١,٦	٠	٠,٠

تابع جدول (١٢)

م	الخاصية	الاستجابة									
		درجة كبيرة جداً		درجة كبيرة		درجة متوسطة		درجة قليلة		درجة قليلة جداً	
		المتوسط	المتوسط	المتوسط	المتوسط	المتوسط	المتوسط	المتوسط	المتوسط	المتوسط	المتوسط
	البعد الثاني: صلية التعلم										
١٤	يجعل صلية المتعلم متحركة حول المتعلم وليس المعلم	٧٤	٥٩,٢	٣٧	٢٩,٦	١٢	٩,٦	٢	١,٦	٠	٠,٠
١٥	يجعل صلية للتعلم أكثر إيقاعاً وجاذبية	٤٣	٣٤,٤	٥٦	٤٤,٨	٢١	١٦,٨	٥	٤	٠	٠,٠
	البعد الثالث: صلية التدريس										
٥	تمدت خلاله صلية التغذية الراجعة بصورة مستمرة	٤٨	٣٨,٤	٥٩	٤٧,٢	١٣	١٠,٤	٣	٢,٤	٢	١,٦
٦	يربط صلية التقييم بالتدريس	٤٧	٣٧,٩	٤٤	٣٥,٥	٢٥	٢٠,٢	٦	٤,٨	٢١,٦	٠,٠
٧	يسمح الفرصة للمعلم التعاوني في الفصل	٥١	٤١,١	٥٠	٤٠,٣	١٧	١٣,٧	٦	٤,٨	٠	٠,٠
٨	يسمح بالمشاركة مع زملائه في الأنشطة والمناقشات	٥٢	٤١,٦	٥٢	٤١,٦	١٨	١٤,٤	٣	٢,٤	٠	٠,٠
١٧	يزيد صلية الاتصال والتواصل بين المعلم والمتعلم	٥٨	٤٦,٨	٤٥	٣٦,٣	١٦	١٢,٩	٥	٤	٠	٠,٠

تبين من الجدول السابق أن متوسط كل العبارات زاد عن ٤,٠٠ (ماعدا العبارة ١٢) مما يدل على أن استجابات الطلاب نحو خصائص التعلم والتدريس والمتعلم تميزت بالإيجابية بدرجة كبيرة، فنلاحظ أن في ضوء استجابات العينة فإن استخدام مدخل التقييم الحقيقي ينمى القدرة على التفكير والتأمل ويجعل المتعلم أكثر قدرة على تطبيق ما تعلمه في المواقف المختلفة بدرجة كبيرة وهذه الاستجابات في العينة تتفق مع رأى (1999) Davies & Michael و (1997) Cole & Struyk و (1992) Notle، كما تبين من استجابات العينة أن التقييم الحقيقي يسعى إلى التغذية والمراجعة المستمرة للأداء بدرجة كبيرة ويجعل المتعلم أكثر إيجابية ونشاطاً في عملية التعلم بدرجة كبيرة وأيضاً وهذا يتفق مع آراء (1999) Ina & Dalean و (1999) Duffy et al., و (2001) Smith & Tillema. ونلاحظ أن استجابات العينة حول ارتباط عملية التقييم بالتدريس تميزت بالإيجابية حيث بلغ قيمة المتوسط

التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية

٤,١٨، ويبدو هذا من أهم المميزات لاستخدام المدخل الحقيقي في التقييم. وينمى مفهوم التعلم المتمركز حول المتعلم واتضح ذلك من استجابات العينة (٤,٤٦) وهذا يتفق مع (Paul & Dylan (1998 و Pairs & Ayers (1999. التساؤل الثاني:

للإجابة على التساؤل الثاني الخاص بالتعرف على اتجاهات طلاب العينة نحو خصائص عملية التعلم والتدريس والمتعلم في ضوء مدخل التقييم التقليدي المتمثل في الممارسات الحالية (استخدام الاختبارات في نهاية عملية التدريس)، قام الباحث بإجراء التوزيعات التكرارية للتعرف على تكرارات الاستجابات لكل خاصية ونسبتها المئوية.

جدول (١٣): التكرارات والنسب المئوية لاستجابات أفراد العينة نحو خصائص التعلم والتدريس والمتعلم في ضوء التقييم التقليدي

الخاصية	الاستجابات									
	درجة كبيرة جداً		درجة كبيرة		درجة متوسطة		درجة قليلة		قليلة جداً	
	تكرار	%	تكرار	%	تكرار	%	تكرار	%	تكرار	%
١. الحد الأول: الفهم										
١. تأثير الفهم والتفكير في مسودة	٨	٦,٤	١٥	١٢	٧٠	٥٦	٢٤	١٩,٢	٨	٦,٤
٢. أبل جيداً كثيراً تفهم على الشرائح	١٣	١٠,٦	٢٧	٢٢	٤١	٣٢,٨	٢٧	٢٢	٧	٥,٧
٣. أخذ على نفسي كمعول على قادة الفعلة	١٣	١٠,٤	٢١	١٧,٨	٣٦	٢٨,٨	٢٣	١٨,٤	١٢	٩,٦
٤. على الفهم على التفكير وفهم	٣	٢,٤	١٥	١٢,٢	٤٧	٣٨,٧	٣٧	٣٠,٦	٢١	١٧,٦
٥. دفع لي الفرصة للصحح لمفاتي	٦	٤,٨	١١	٨,٨	٣٠	٢٤,٠	٤٦	٣٦,٨	٢٢	١٧,٦
١٠. سمع في مخرج فكافى خلال الفعالتات	٦	٩,٧	١٢	٩,٧	٤٥	٣٦,٣	٤٢	٣٣,٦	١٤	١١,٣
١١. يمتطي لمثلث بالشرائح لمدة أطول	٢	١,٦	١٠	٨	٣٠	٢٤	٤٢	٣٣,٦	٤٠	٣٢,٠
١٢. يمتطي أكثر فترة على تفهم ما كتبه في موقت أفوى	٥	٤,٠	١٢	٩,٦	٢٨	٢٢,٤	٤٧	٣٧,٦	٢٣	١٨,٤
١٣. دفع لي فرصة أفوى لي تكون مستوى أفوى	٦	٤,٨	٨	٦,٦	٢٣	١٨,٦	٤٠	٣٢,٨	٥٠	٤٠
١٦. على أفوى ولتأوى الوصول إلى الإثنان	٣	٢,٤	١١	٨,٩	٤٦	٣٦,٦	٤٨	٣٨,٧	١٦	١٢,٩
الحد الثاني: صلية الفهم										
١٤. جعل صلية الفهم متركزة حول الفهم وليس الفهم	٤	٣,٢	٩	٧,٢	٢٧	٢٢,٦	٤٠	٣٢,٦	٢٥	٢٠
١٥. جعل صلية الفهم أكثر شاملاً وجالباً	٦	٤,٨	١٢	٩,٦	٤١	٣٢,٨	٤٥	٣٦,٨	٢١	١٦,٨
الحد الثالث: صلية الفهم										
٥. سمع خلال صلية الفعلة الفرصة صلية مسودة	٦	٤,٨	٩	٧,٢	٥٢	٤١,٣	٢١	١٦,٨	٢٢	١٧,٦
٦. ربط صلية الفهم بالتدريس	٨	٦,٤	١٨	١٤,٦	٤٦	٣٦,٤	٤١	٣٢,٣	١٠	٨,٦
٧. دفع الفرصة لعل التناوى لي الفهم	٢	١,٦	١٤	١١,٤	٢٩	٢٣,٦	٥٠	٤٠,٧	٢٨	٢٢,٨
٨. سمع بالشاركة مع ولتأوى لي الفعلة والفعالتات	٣	٢,٤	١١	٨,٨	٢٨	٢٢,٤	٤٦	٣٦,٨	٢٧	٢١,٦
١٧. دفع صلية الفهم والوصول إلى الفهم والفهم	٣	٢,٤	١٠	٨,٩	٤٨	٣٨,٧	٤٢	٣٣,٦	٢٠	١٦,٦

تبين من جدول (١٣) أن متوسط معظم العبارات (ماعدًا ٢ و ٣) انخفضت عن (٣,٠٠) مما يدل على أن استجابات الطلاب نحو خصائص المتعلم والتدريس والتعلم تميزت بالسلبية إلى حد ما.

تبين أن التقييم التقليدي يرتبط بعملية التدريس في ضوء استجابات العينة بدرجة قليلة (٢,٧٨) وأن جاز القول في ضوء التقريب لهذه القيمة بدرجة متوسطة وهذا يتفق مع رأى Kim et al., (1996) حيث أشار إلى أن الممارسات التقييمية الحالية غير قادرة على توجيه العمل التدريسي اليومي. كما أن الاختبارات التقليدية تظهر قدرات وإمكانات الطالب في ضوء استجابات العينة بدرجة متوسطة (٢,٩٢) وهذا يتفق مع (Shepard (2000). وأظهرت استجابات العينة أن التقييم في ضوء الاختبارات التقليدية لا يتيح الفرصة للمتعلّم على تطبيق ما تعلمه في المواقف الأخرى إلا بدرجة قليلة (١,٩٣) وهذا يتفق مع (Shepard (2000 و Newell (2000) ودور عملية التقييم التقليدي في جعل عملية التعلم أكثر إمتاعاً وجاذبية كانت بدرجة قليلة.

وعموماً فإن استجابات الطلاب نحو خصائص عملية التعلم والتدريس والمتعلم في ضوء التقييم الحقيقي كانت أكثر إيجابية عكس استجاباتهم نحو نفس الخصائص في ضوء التقييم التقليدي الذي تميزت بالسلبية.

التساؤل الثالث:

للإجابة على التساؤل الثالث الخاص بمعرفة وجود فروق جوهرية بين الاتجاهات نحو خصائص عملية التعلم والتدريس والمتعلم بين مدخل التقييم الحقيقي ومدخل التقييم التقليدي.

وللتحقق من هذا استخدم الباحث اختبار T-Test للعينات المرتبطة واتضح من المعالجة الإحصائية وجود فرق ذات دلالة جوهرية للاتجاهات نحو خصائص عملية التعلم بين مدخل التقييم الحقيقي ومدخل التقييم التقليدي حيث بلغت قيمة $T = 17.92$ ($P=0.00$) وهي دلالة إحصائية عند ٠,٠١ وكذلك وجود فرق بين الاتجاهات نحو خصائص عملية التدريس بين المدخلين وقيمة $T = 17.42$ ($P =$

==التقييم الحقيقي (البورتفوليو) كمدخل لتكامل الاختبارات التحصيلية==

0.00) وهى دالة أيضاً عند ٠,٠١ وتبين وجود فروق بين الاتجاهات نحو خصائص المتعلم بين مدخلى التقييم وأن قيمة $T = 19.71$ ($P = 0.00$) وهى دالة إحصائياً عند ٠,٠١ وتبين أن منشأ هذه الدلالة يعود إلى الاتجاهات الإيجابية العالية لخصائص عملية التعلم والمتعلم والتدريس فى ضوء التقييم الحقيقى فى مقابل الاتجاهات السلبية إلى حد ما لنفس الخصائص فى ضوء التقييم التقليدى، أى أن الدلالة لصالح التقييم الحقيقى. وهذا يتفق مع العديد من الآراء Davies & Miacheal (1999) و Cole & Struyk (1997) و Ina & Dalean (1999) و Duffy et al., (1999) و Salvia & Yesseldyke (1995).

التساؤل الرابع:

للإجابة على هذا التساؤل الخاص بمعرفة العلاقة بين التقييم الذاتى للأنشطة التقييمية التى تتضمنها الحقيبة والأداء على هذه الأنشطة وبحساب معامل ارتباط بيرسون يتضح أن قيمته $= ٠,٢٣$ - ($P = 0.013, N = 113$) وهذا يعنى وجود علاقة سالبة غير دالة إحصائية عند ٠,٠١ ويبدو أن هذه القيمة تخالف التوقع الذى يشير إلى أن استخدام التقييم الحقيقى ينمى مهارات التقييم الذاتى ونتيجة الدراسة تتعارض مع رأى White (1994) وأن كانت White ترى أن استخدام التقييم الحقيقى ينمى مهارات التقييم الذاتى ولكنها نادت بإيجاد دلائل أمبريقية على ادعاءها هذا. وقد يعود هذا الارتباط السالب إلى حداثة استخدام هذا المدخل لهؤلاء الطلاب حيث أنهم تربوا فى نظام تربى قائم على أساس الامتحانات فكان هذا التحول المفاجئ إلى نظام التقييم الحقيقى بمثابة تحول من نظام تقيمى مألوف إلى نظام آخر لم يألفه من قبل.

التساؤل الخامس:

للإجابة على التساؤلات من الخامس حتى السابع أعتمد الباحث أعتد على عينتين، عينة الدراسة (١٢٦) وهى العينة التى طبق عليها الاختبار التحصيلى فى نهاية الترم وتم الحصول على درجات الطلاب فى هذا الاختبار من الكنترول شيت، وعينة صدق النتائج وهى مكونة من (٦٤) طالب طبق عليهم الاختبار التحصيلى

لأعمال السنة وهى تقريباً نصف عينة الدراسة من ثم فإن عرض النتائج للأسئلة الآتية سيتم فى ضوء العينتين.

ولمعرفة وجود علاقة بين الأداء على الاختبارات التحصيلية والأداء فى ضوء التقييم الحقيقى تم حساب معامل ارتباط بيرسون، وفى عينة الدراسة بلغت قيمة معامل الارتباط $0,64$ ($P = 0.00$) لحجم عينة (١٢٠) وهى دالة إحصائياً عند مستوى دلالة إحصائية 0.01 ، أما فى عينة صدق النتائج بلغت قيمة معامل الارتباط 0.52 ($P = 0.00$) لحجم عينة (٦٠) وهى دالة إحصائياً عند مستوى دلالة $0,01$ ويتضح أن قيمة معامل ارتباط بيرسون فى عينة الدراسة تزيد عن قيمة معامل ارتباط بيرسون فى عينة صدق النتائج ولكنهم ينحصران فى المدى من $0,5$ إلى $0,65$ وهذا يتفق مع ما توصل إليه Supovitz & Brennan (1997) حيث توصل إلى أن العلاقة بين الأداء فى ضوء المدخل الحقيقى والأداء فى ضوء الاختبارات التقليدية يتراوح من $0,5$ إلى $0,6$ ويتفق أيضاً مع Heck & Crislip (2001) حيث وجد قيمة هذه العلاقة $0,56$ ، وإن كانت قيمة معامل الارتباط فى عينة صدق النتائج تقترب من نتائج الدراسات السابقة.

التساؤل السادس

لمعرفة مدى وجود فروق بين أداء الطلاب على الاختبارات التقليدية وأدائهم فى ضوء التقييم الحقيقى، أتضح من خلال عينة الدراسة (١١٧) وجود فروق ذات دلالة إحصائية بين الأداء على الاختبار التحصيلى فى نهاية الترم والأداء على الأنشطة التقييمية حيث بلغت قيمة $T = 4,85$ ($P = 0.00$) وهى دالة إحصائياً عند مستوى دلالة $0,01$ ، أما فى عينة صدق النتائج فإن $T = -17,97$ ($P = 0.00$) وهى دالة عند مستوى دلالة $0,01$ أيضاً وهذا يشير إلى وجود فروق ذات دلالة إحصائية بين الأداء على الاختبارات التحصيلية التقليدية والأداء فى ضوء التقييم الحقيقى وذلك لعينة الدراسة ولعينة صدق النتائج. ونتيجة الدراسة الحالية تتعارض مع ما توصل إليه Slater et al., (1997) بعدم وجود فروق بين أداء المجموعة الضابطة التى اعتمدت فى تقييمها على الاختبارات التقليدية وأداء المجموعة التجريبية التى اعتمدت فى تقييمها على أنشطة التقييم الحقيقى.

التساؤل السابع:

لمعرفة أى من النماذج الثلاثة المطروحة تصلح للتنبؤ بالأداء النهائى (تحصيل المقررات الدراسية الأربعة)، فالنموذج الأول قائم على أساس التنبؤ بتحصيل المقررات الأربعة من درجة الإحصاء فى امتحان نهاية الترم والنموذج الثانى قائم على أساس التنبؤ فى ضوء الأداء على الأنشطة التقييمية والنموذج الثالث قائم على أساس التنبؤ فى ضوء درجة (تحصيل) الإحصاء فى نهاية الترم والأداء فى ضوء الأنشطة التقييمية. ففى عينة الدراسة تم إجراء تحليل الانحدار المتعدد وللمفاضلة بين النماذج الثلاثة لمعرفة أى هذه النماذج أكثر مطابقة للعينة بيانات العينة تم الاعتماد وعلى مؤشرات حسن المطابقة* وهى R^2 (مربع معامل الارتباط المتعدد) و R_a^2 (مربع معامل الارتباط المصحح) و MSE (متوسط مربع الخطأ) و Cp (مؤشر Mallow 1970).

جدول (١٤): مؤشرات حسن المطابقة للنماذج التنبؤية الثلاثة لعينة الدراسة

الدالة الإحصائية	Cp	MSE	R_a^2	R^2	مؤشر حسن المطابقة / نموذج التنبؤ
P = 0.00	٢	٢٤٠,٧٩	٠,٤٢٠	٠,٤٢	درجة الإحصاء
P = 0.00	٢	٣٠١,٨٤	٠,٢٥٢	٠,٢٥	الأداء على الأنشطة التقييمية
P = 0.00	٣	٢١٥,٤٤	٠,٤٦	٠,٤٧	المؤشرين معاً

يتبين من جدول (١٤) أن أفضل قيمة لمؤشر R^2 للنماذج الثلاثة هى القيمة الخاصة بنموذج التنبؤ فى ضوء الاختبار التحصيلي والأداء على الأنشطة التقييمية معاً حيث بلغت قيمة R^2 (٠,٤٧)، يليها نموذج التنبؤ فى ضوء درجة الإحصاء فى اختبار نهاية الترم، ثم نموذج التنبؤ فى ضوء الأداء على الأنشطة التقييمية. وأن أفضل قيمة لمؤشر R_a^2 هى أيضاً لنموذج التنبؤ فى ضوء المؤشرين معاً حيث بلغت (٠,٤٦)، بينما قيمته لنموذج التنبؤ فى ضوء الاختبار التحصيلي (٠,٤٢)،

* للمزيد عن ذلك يمكن الرجوع إلى دراسة للباحث تحت النشر بعنوان "مقارنة امبريقية لنماذج تنبؤية باستخدام بعض مؤشرات حسن المطابقة".

وقيمته لنموذج التنبؤ في ضوء الأداء مع النشطة التقييمية (٠,٢٥)، وتتضح أن أقل قيمة لمؤشر MSE (متوسط مربعات الخطأ) هي لنموذج التنبؤ القائم على أساس المعيارين معاً ٢١٥,٤٤. أما في ضوء مؤشر (1970) Mallow (Cp) يكون أفضل نموذج مطابق للبيانات في ضوء هذا المؤشر إذا كانت حاصل طرح قيمته من قيمة (K + 1) (حيث K عدد المتغيرات المستقلة في النموذج) أقل ما يمكن.

وفي ضوء الجدول السابق يتضح أن كل النماذج متساوية من حيث مطابقتها للبيانات، حيث كانت حاصل طرح قيمته من (K+1) تساوى صفر.

في ضوء ماسبق يتبين أن أفضل نموذج يصلح للتنبؤ بالتحصيل في نهاية الترم في المقررات الأربعة هو النموذج المتضمن درجة اختبار التحصيل في الإحصاء في نهاية الترم ودرجة الأداء على الأنشطة التقييمية وكان أكثر مطابقة للبيانات من النماذج السابقة وظهر ذلك في تفوقه في مؤشرات R^2 ، R_a^2 ، MSE، بينما تساوت أفضليته مع النماذج الأخرى في ضوء مؤشر Cp وهذا يتفق مع رأى Valencia (1992) وحيث ترى أن الاعتماد على التقييم الحقيقي بجانب التقييم الرسمي لتوثيق أداء الطلاب أفضل من الاعتماد على معيار واحد.

أما بالنسبة لعينة صدق النتائج (عينة التعميم) فيتضح أن مؤشرات حسن المطابقة لها كالاتى:

جدول (١٥): مؤشرات حسن المطابقة للنماذج الثلاثة لعينة صدق النتائج

الدالة الإحصائية	Cp	MSE	R_a^2	R^2	مؤشر حسن المطابقة / نموذج التنبؤ
P = 0.00	٢	٢٢٠,١٦	٠,٥٠	٠,٥١	الاختبار التحصيلي لأعمال السنة
P = 0.00	٢	٢٩٧,٢٣	٠,٢٨	٠,٣٠	الأداء على الأنشطة التقييمية
P = 0.00	٣	١٨٧,٣١	٠,٥٥	٠,٥٦	المؤشران معاً

يتضح من جدول (١٥) ان نموذج التنبؤ القائم على أساس المؤشرين الاختبار التحصيلي والأداء في ضوء الأنشطة التقييمية يفوق النموذجين الآخرين حيث كانت

التقييم الحقيقي (البورفوليوس) كمدخل تكاملي للاختبارات التحصيلية

له أعلى قيمة لمؤشر R^2 (٠,٥٦) و مؤشر R^2 (٠,٥٥) وكانت له أقل قيمة لمؤشر MSE (١٨٧,٣١) وتساوى مع النموذجين الآخرين في مؤشر Cp وبليها في الأفضلية نموذج التنبؤ في ضوء الاختبار التحصيلي ثم نموذج التنبؤ في ضوء الأداء مع الأنشطة التقييمية وهذه النتيجة تتفق مع نتيجة عينة الدراسة.

التعليق

يبدو أن استخدام وتوظيف مدخل التقييم الحقيقي مازال في مرحلة الخداسة ولم ترسخ كيفية تطبيقه وكيفية ممارسة النواحي التقييمية للأنشطة المتضمنة فيه. ولكن هذا المدخل يعتبر من التطورات الهامة في مجال القياس والتقييم والتدريس، وأصبح ضرورة ملحة في ضوء التطورات والتغيرات في نظرية التعلم وتغير فلسفة التقييم في القرن الحالي من فلسفة قائمة على اختبارات تقليدية إلى فلسفة قائمة على تنوع الأدلة للأداء. حتى تزيد من درجة المصداقية للحكم على الأداء. والتقييم الحقيقي يعتبر أكثر ثراء من حيث محتواه وأساليبه من طرق التقييم التقليدية وذلك في ضوء آراء كثير من المحدثين والمجددين في مجال التربية والقياس وأصبح هذا المدخل أكثر تطبيقاً في كثير من الأنظمة التعليمية.

ومدخل التقييم الحقيقي القائم على مدى إنجاز الطالب لأنشطة وتكليفات يطبق في النظام التعليمي المصري - خاصة الجامعي - بدرجة خفيفة، أذ أحياناً يطلب الأستاذ من تلاميذه إعداد بحث أو نشاط ما لكي يكافئ الطالب عليه في أعمال السنة، ولكن تطبيقه كمدخل تقييمي تدريسي أشمل من ذلك بكثير.

والدراسة الحالية لم تعتمد على تطبيق المدخل التقييمي الحقيقي بصورة تامة في عملية التدريس والتقييم، إنما اعتمدت عليه بصورة جزئية كأسلوب مكمل لعملية التدريس وأيضاً مدعم لعملية التقييم التقليدية.

وهدف الباحث من هذه الدراسة إلى التعرف على أهم العوامل والتغيرات التي أدت إلى الاهتمام المتزايد بأهمية تطبيق التقييم الحقيقي ومميزاته والصعوبات التي تقابل تنفيذه وتوصلت الدراسة إلى نتائج هامة، حيث توصلت إلى أن الاتجاهات نحو خصائص المتعلم والتدريس وعملية التعلم تكون أكثر إيجابية ودينامية في

ضوء تطبيق هذا المدخل وذلك مقارنة بالاتجاهات نحو هذه الخصائص فى ضوء التقييم وذلك من وجهة نظر عينة الدراسة. وتوصلت الدراسة إلى أن معامل الارتباط يبين الأداء فى ضوء الاختبارات التقليدية والأداء فى ضوء التقييم الحقيقى يقع فى المدى من ٠,٥ إلى ٠,٦٥ وهى علاقة ليست قوية وأن كانت دالة إحصائياً. واتضح بأهمية التكامل بين الأداء فى ضوء التقييم الحقيقى والأداء فى ضوء الاختبارات التقليدية للتنبؤ بالأداء الأكاديمى العام خاصة فى مرحلة الدراسات العليا فى التعليم الجامعى وهذا يدعو إلى أهمية تضمين مدخل التقييم الحقيقى لصناعة القرارات الخاصة بالطلاب.

ولكن الدراسة توصلت إلى نتيجة تبدو مخالفة لكثير من التوقعات وهى أن استخدام مدخل التقييم الحقيقى لا ينمى مفهوم التقييم الذاتى للمتعلمين وقد يكون ذلك بسبب عدم ألفة الدارسين على التعامل مع هذا النمط، خاصة أن الدارسين كان يشكون من كثرة الأنشطة أو التكاليفات المنوطون بها وأدى ذلك إلى تكوين صورة تتميز بالسلبية لأدائهم ومدى تمكنهم من المهارات التى تتضمنها هذه الأنشطة التقييمية ولكن ماذا لو تم تطبيق استمارة التقييم الذاتى منذ بداية إعطائهم التكاليفات حتى انتهاءهم من أداء هذه التكاليفات؟ وكإجابة مرحلية متوقعة كان يمكن أن تتحسن صورة الطلاب للحكم على أنفسهم لدرجة التمكن من المهارات المتضمنة فى الأنشطة التقييمية.

وتميزت نتائج الدراسة بالاستقرار عبر عينة الدراسة وهى عينة الدراسة (١٢٦) وعينة صدق النتائج (٦٤) مما يزيد من درجة الثقة فى نتائج الدراسة وهذا يدعو إلى تكون النتائج قابلة للتعميم على مستوى المجتمع.

وأخيراً فإن مدخل التقييم الحقيقى ضرورة ملحة لتطبيقه فى نظامنا التعليمى خاصة مرحلة التعليم الجامعى ولكن لابد من إجراء الدراسات حول هذا المدخل، وذلك لمعرفة مدى فعاليته مقارنة بالمدخل التقليدى فى عملية التقييم والباحث الحالى يدعو الى دراسة هذه الفعالية من خلال تصميم تجريبى لمجموعة يتم التعامل معها من خلال نظام تقييمى معتمد على المدخل الحقيقى بدرجة تامة ومجموعة يتم التعامل

==التقييم الحقيقي (البورتفوليو) كمدخل تكاملي للاختبارات التحصيلية==

معها من خلال نظام تقييمي يعتمد على الاختبارات التقليدية بدرجة تامة ومجموعة يتم التعامل معها من خلال نظام تقييمي يشمل المدخلين معاً.

كما توجد العديد من التساؤلات حول هذا المدخل التقييمي وهى مدى قدرة استخدام هذا المدخل التقييمي على تنمية مهارات ما وراء المعرفة وتقوية التعلم المنظم ذاتياً وتنمية مهارات التفكير العليا وممارسة مهارات حل المشكلات وغيرها من الأسئلة لعل يوليهها الباحثين النفسيين والتربويين قدراً من الاهتمام والدراسة وذلك لإيضاح الرؤية حول مدخل التقييم الحقيقي (البورتفوليو) الذى يعتبر فى مرحلة التطبيق فى واقعنا التعليمي.

المراجع

- 1- Barton, J. & Collins, A. (1997). Portfolio assessment: A handbook for educators. Menlo Park: Adison – Wesley Publishing Co.
- 2- Calfee, R., & Perfumo, P. (1993). Student Portfolios Opportunities for a revolution in assessment. Journal of reading, 36, 532-537.
- 3- Camp, R. (1993). The place of portfolios in our changing views of writing assessment. In: R.E. Bennett & W.C. Ward (Eds.), construction versus choice in cognitive measurement: Issues in constructed response performance testing, and portfolio assessment (pp. 183-212). New Jersey: Lawrence. ERLBUM Associates.
- 4- Camp, R., & Levine, D. (1991). Portfolios Evolving: Background and variations in sixth-through twelfth – grade classroom. In: P. Belanoff & M. Dickson (Eds.), portfolios: Process and Product (pp. 194-205). Portsmouth: Boynton/Cook Publishers.
- 5- Cole, K.B., & Struyk, L.R. (1997). Portfolio assessment: Challenges in secondary education. High School Journal, 80, 261-272.
- 6- Condon, W. & Lyons, L.W. (1994). Maintaining a portfolio-based writing assessment: Research that informs program development. In: L. Black, D.A. Diaker, J. Sommers & G. Stygall (Eds.), New directions in portfolio assessment (pp. 277-285). Portsmouth: Boynton/Cook Publishers, Inc.
- 7- Crooks. T.J. (1988). The impact of classroom evaluation practices on student. Review of educational research, 58, 438-481.
- 8- Cross, K.P. (1998). Classroom Research: Implementing the scholarship of teaching. In: T. Angelo (Eds.),

- Classroom assessment and research. An update on uses, approaches, and research findings (pp. 5-12). San Francisco: Jossey – Bass Publishers.
- 9- David, F.B., Davis, M.H., Harden, R.M., Howie, P.W., Ker., & Pippard, M.J. (2001). AMEE medical Education Guide No.24: Portfolios as a method for student assessment. *Medical Teacher*, 23, 534- 545.
- 10-Davies, M.A., & Michael, W. (1999). Alternative assessment. *New Direction in teaching and learning. Contemporary Education*, 71, 39- 47./
- 11-Duffy, M.L., Jones, J., & Thomas, S.W., (1999). Using portfolios to foster Independent thinking. *Intervention in school & Clinic*, 35, 34- 37.
- 12-Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in Higher Education: A Meta-analysis comparing peer and teacher Marks. *Review of Educational Research*, 70, 287-322.
- 13-Firestone, W.A., Mayrowetz, D. & Fairman, J (1998). Performance based assessment and Instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and policy analysis*, 20, 95-113.
- 14-Graue, M.E. (1993). Integrating Theory and practice through instructional assessment. *Educational Assessment*, 1, 293-309.
- 15-Heck, R.H., & Crislip, M. (2001). Direct and Indirect writing assessments: Examining Issues of Equity and Utility *Educational Evaluation and Policy Analysis*, 23, 275-292.
- 16-Herman, J.L. (1992). What research tells us about good assessment. *Educational leadership*, 49, 74-78.
- 17-Ina, F., & Daleen, V.N. (1999). Using Portfolio assessment in a module in research information skills. *Education for information*, 17, 333-337.

- 18-Kim, J., McLean, J.E., & Iran-Nejad, A. (1996). The Integration of Teaching, Testing, and Learning. *Educational Research Quarterly*, 19, 29-50.
- 19-McLoughlim, J.A., & Lewis, R.B. (1990). *Assssing special students* (3rd Ed.). New York: Macmillan Publishing Company.
- 20-Murphy, S. (1994). Writing portfolios in K-12 schools: Implications for linguistically diverse students. In L. Black, D.A. Diaker, J. Sommers, & G. Stygall (Eds.), *New directions in portfolio assessment* (pp. 140-156). Portsmouth: Boynton /Cook Publishers, Inc.
- 21-Nancy, H., Marker-Schroer, M.F. (1992) Multidimensional assessment of the gifted minority student. *Roper Review*, 15, 73-77.
- 22-Newell, R.J. (2002). A different look at accountability. The Envisions approach. *Phi Delta Kappan*, 84, 208-211.
- 23-Nolet, V. (1992). Classroom-based measurement and portfolio assessment. *Diagnostique*, 18, 5-23.
- 24-O'Neil, J. (1992). Putting performance assessment to test, *Educational Leadership*, 49, 14-19.
- 25-Paris, S.G., & Ayres L.R. (1994). *Becoming Reflective Students and Teachers with Portfolios and Authentic Assessment*. Washington: American Psychological Association Books.
- 26-Paul, B. & Dylan, W. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5, 7-74.
- 27-Salvia, J., & Yesseldyke, J.E. (1995). *Assessment* (6th Ed.). Boston: Houghton Mifflin Company.
- 28-Shepard, L.A. (2000). *The role of classroom assessment in teaching and learning*. California University of California Publishing.
- 29-Slater, T.E., Ryan, J.M., & Samson, S.L. (1997). Impact and dynamics of portfolio assessment and traditional

- in a college physics course. *Journal of Research in Science Teaching*, 34, 255-271.
- 30-Smith, K., & Tillema, H. (2001). Long term influences of portfolios on professional development. *Scandinavian Journal of Educational Research*, 45, 183-202.
- 31-Snow, R.E. (1989). Toward assessment of cognitive and conative structure in learning. *Educational Researcher*, 18, 8-14.
- 32-Sommers, J. (1991). Bringing practice in line with theory using portfolio grading in the composition classroom. In: P. Belanoff & M. Dickson (Eds.), *portfolios: Process and product*, Portsmouth: Boynton/Cook Publishers, Inc.
- 33-Supovitz, J.A. & Brennan, R.T. (1997). Mirror, mirror on the wall, which is the fairest test of all? An examination of the equitability of portfolio assessment relative to standarized tests. *Harvard Educational Review*, 67, 472-506.
- 34-Valencia, S. (1992). Implementation Literacy Portfolio Assessment. Paper presented at the Meeting of the Southwest Regional International Reading Association, Tucson, AZ.
- 35-White, E.M.(1994). Portfolios as an assessment concept. In: L. Black, D.A. Diaker, J. Jommers & G. Stygall (Eds.), *New directions in portfolio assessment* (pp. 25-39). Portsmouth: Boynton/Cook Publishers, Inc.
- 36-Wolf, D.P. (1989). Portfolio assessment: Sampling student work. *Educational leadership*, 46, 35-39.

Portfolio assessment as an approach To the traditional

Prepared by

Dr. Abd El Nasser El Sayed Amer
Lecturer of Measurement and Evaluation
Faculty of Education – Ismailia - Suez
Canal University

Abstract

The study focus on portfolio assessment as an integrated approach to standardized tests. It included two samples, the sample of study (126 graduate students) and the cross-validity sample (64 graduate students). The measures included, self assessment questionnaire ($N= 11$, $\alpha=0.56$), attitudes scale toward the characteristics of the learning, instruction, and learner ($N = 17$, $\alpha = 0.86$, 90), achievement test (30 item, $\alpha = 0.88$), and checklist for assessing the activities in portfolio ($N = 20$, $\alpha=0.89$).

The attitude toward the characteristics of the learning instruction and learner is strongly positive interms of portfolio assessment while the attitude is negative toward the previous characteristics interms of standardized tests. There are significance differences between the performance on portfolio assessment and the performance on standardized test, and positive correlation between the performance on the activities included in portfolio and scores on test, for the sample study ($r=0.64$,) and for cross validity sample ($r = 0.52$), and negative correlation between self assessment and performance on portfolio assessment (-0.23 , $P = 0.01$).

The results demonstrate that the predictive Model interms of the scores on achievement test and the scores on portfolio is the best predictor for achievement in the final term. This result assert that the importance of the integration between portfolio assessment and standardized tests for students decision making.