نمذجة البيانات الطولية لتقييم فاعلية اداء المعلمين باستخدام مدخل القيمة المضافة

د. هشام حبيب الحسينى محمد

استاذ علم النفس التربوى المساعد بالمركز القومى للامتحانات والتقويم التربوى

ملخص الدراسة

هدفت هذه الدراسة الى دراسة تقييم فاعلية اداء المعلمين من خلال ربط فاعلية ادائهم بمدى تحقيق طلابهم لنواتج التعلم فى مادتى الرياضيات واللغة العربية ودرجة النمو الحادثة من خلال دراسة طولية لمدة ثلاث سنوات لطلاب الصفوف الثلاثة الأولى من التعليم الابتدائى، ومعرفة مدى الاتساق بين نتائج التقييم باستخدام نموذجين من نماذج القيمة المضافة وهما نموذج الكسب (كأحد نماذج القيمة المضافة أحادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات).

وقد تم اختيار العينة من معلمى الصفوف الثلاثة الاولى من التعليم الابتدائى (معلمى الفصل) حيث يقوم كل معلم بتدريس كل من اللغة العربية والرياضيات، كما يمكن تتبع اداءهم مع طلابهم على مدار ثلاث سنوات حيث بلغت العينة ١٣ معلماً بنهاية الدراسة بعد انتقال احد المعلمين إلى مدرسة اخرى بعد العام الاول من التطبيق. وتكونت عينة الدراسة من طلاب الصف الأول والثانى والثالث الابتدائى فى كل من الدراسة الأساسية وعينة تقنين أدوات الدراسة.

وتم تطبيق مقياس المستوى الاجتماعى الاقتصادى ومقياس الاستعداد اللغوى والرياضى فى صورتهم النهائية فى النصف الاول من العام الدراسى نفسه. بينما طبقت اختبارات نواتج التعلم فى الرياضيات واللغة العربية عبر الدراسة الطولية مع نهاية النصف الدراسى الاول ونهاية نصف الدراسى الثانى على مدار ثلاثة اعوام دراسية هى: ٢٠١٢/٢٠١٣ ، و٢٠١٣/٢٠١٤، و٢٠١٤/٢٠١٥.

وكانت من اهم نتائج الدراسة ما يلى:

١- ان معاملات الارتباط بين نتائج الرياضيات واللغة العربية منخفضة نسبياً فى الصف الثانى والثالث الابتدائى باستخدام نماذج الكسب، بينما كان معامل الارتباط منخفض فى الصف الثالث الابتدائى فقط عندما استخدم نموذج التأثيرات المتعددة.

٢- بالنسبة الى العلاقة بين نتائج استخدام كل من نموذج الكسب ونموذج التأثيرات المتعددة كمؤشر لفاعلية أداء المعلم وترتيب المعلمين وفقا لأداءتهم (من خلال نتائج قياس نواتج التعلم فى مادتى الرياضيات واللغة العربية)، كانت العلاقة الارتباطية بين نتائج النموذجين تتراوح ما بين مرتفعة إلى متوسطة فى جميع الصفوف.

نمذجة البيانات الطولية لتقييم فاعلية اداء المعلمين باستخدام مدخل القيمة المضافة

د. هشام حبيب الحسينى محمد

استاذ علم النفس التربوى المساعد بالمركز القومى للامتحانات والتقويم التربوى

مقدمة :

يعد استخدام البيانات محوراً اساسياً لكل من عمليات التحسين، والفاعلية الكلية للمدرسة overall school practice، وتقديم معلومات لاتخاذ القرارات التربوية المهمة، وتقدم تأثيراً موجباً على المعلمين من خلال تنميتهم مهنياً، كي تساعد على تغيير وجهات النظر والاتجاهات نحو الممارسات التربوية(Wohlstetter, Datnow, & Park,2008) .

وتعتبر نماذج تقييم القيمة المضافة value - added assessment أحد أنظمة تحليل البيانات التى تقدم تقديرات غير متحيزة للنظام المدرسى والمعلمين ودورهم فى نمو أداء المتعلمين فى المواد الدراسية المختلفة. وتقديم معلومات كافية فيما يخص نمو الطلاب ومدى تحقيقهم لنواتج التعلم المستهدفة ودور المعلمين فى نمو اداء المتعلمين من خلال بيانات تراكمية دقيقة، ومن ثم فمن المهم إدراك وفهم تفسير نتائج القيمة المضافة المعقدة بشكل بسيط يمكن تطبيقه فى السياق المؤسسى للمدرسة لتحقيق النمو والتطور المستهدف (Wainer,2004).

وفى ضوء أهمية التقويم التربوي لعناصر العملية التعليمية، تبرز أهمية الحكم على مدى ما يحققه المعلم لتحقيق نواتج التعلم المستهدفة للمتعلمين من خلال النقلة النوعية التي أحدثها تقييم القيمة المضافة حيث انتقل من دراسة التغير في المجموعات cohort-to-cohort analysis إلي دراسة التغير الفردي والاهتمام بجمع البيانات الطولية longitudinal data collection، ومن ثم تستخدم نماذج القيمة المضافة value - added modeling (VAM) لمعرفة مدى فاعلية المدرسة والمعلم (Braun,2005).

وقد يكون الاهتمام المتزايد بالقيمة المضافة تأكيداً لوجهة النظر التى تؤكد أن الأساليب الإحصائية المتطورة يمكن أن تقدم تقييماً لفاعلية المعلمين من خلال مؤشرات تضع فى اعتبارها القوى والعوامل غير التربوية مثل الخلفية الاجتماعية والاقتصادية للمتعلمين (Braun,2006).

ولتحقيق هذا الهدف يتم استخدام نماذج القيمة المضافة حيث تسمح بفصل إسهامات المعلم فى أداء المتعلم من خلال عوامل السياق التى تخرج خارج نطاق التحكم فى الفصول والمؤسسات التعليمية. فكثير من الدراسات أظهرت أن نماذج القيمة المضافة تقدم مزيداً من الدقة لتقدير أداء المعلمين بما يجعلها تتكامل مع الطرق الأخرى التى تستخدم فى تقويم ادائهم. فهذه النماذج تقدم أسساً وقيماً كمية أكثر دقة لأداء المعلم.

مشكلة الدراسة :

يعد الغرض الأساسى من تقويم فاعلية أداء المعلمين أن تظهر عملية التقويم طبيعة العلاقة بين ما يتم تدريسه من جهة وما قد تعلمه الطلاب من جهة أخرى. وقد نتج عن هذا التوجه حركات إصلاحية كثيرة للتعليم من أهمها: حركة قياس الحد الأدنى للكفاءات minimum competency testing movement في السبعينيات، وحركة المعايير standards movement في الثمانينيات، والتقويم التفاعلي وتقويم المردود interactive and impact evaluation في أواخر التسعينيات من القرن العشرين، والتي أدت لظهور "تقييم القيمة المضافة" والتي اعتبرت من أهم حركات إصلاح التعليم وتطوير جودته في أوائل القرن الحادى والعشرين (Zupanc, Urank, Bren,2009).

وهناك افتراضات أساسية يجب أن تتوافر لاستخدام أي نموذج من نماذج القيمة المضافة هى: صدق التكوين الفرضي لدرجات الاختبارات، حيث تعتمد قيمة النتائج التي نحصل عليها من استخدام أي نموذج من نماذج القيمة المضافة علي الخصائص السيكومترية الجيدة لاختبارات التحصيل مثل الصدق والثبات والموضوعية؛ ومعادلة درجات الاختبارات، حيث يعتمد مدخل القيمة المضافة على عدالة المقارنات التي تتم من خلال نماذج القيمة المضافة؛ وعدم تحيز الاختبار، وبقاء أثر المعلم علي الطالب، حيث تفترض نماذج القيمة المضافة بقاء أثر المعلم علي الطالب عبر السنوات الدراسية المتعاقبة (مجدى قاسم وصلاح علام وهشام الحسينى وهشام جاد الرب، ٢٠١١).

وقد أظهرت القيمة المضافة مجموعة من القضايا الأساسية، وإذا لم يتم تحديد دراسة هذه القضايا بشكل واضح ومحدد، سوف تفقد القيمة المضافة الحكم الموضوعى لفاعلية المعلمين وتظهر تعميمات غير صحيحة، وبالتالى قد تعيق الجهود المنظمة لتحسين عمليتى التعليم والتعلم. حيث إن فحص ومناقشة قضايا وإشكاليات استخدام القيمة المضافة فى عمليات التقويم التربوى مازالت محدودة، خاصة فى مصروالوطن العربى حيث لم تلق تحليلات القيمة المضافة الاهتمام الكافى بها – فى حدود علم الباحث – حتى الآن، حيث إن معقولية نتائج دراسات القيمة المضافة هى التى تؤكد الاعتماد عليها فى عمليات التقويم . (Kupermintz, 2002).

فهناك حالة من الجدل طويل الأجل حول تأثيرات المعلمين على الطلاب، وقد نجحت بعض

بحوث نماذج القيمة المضافة في بلورة هذا الجدل إلى حد كبير خلال السنوات القليلة الماضية. كدليل على تأثيرات المعلمين على الطلاب. وتعد هذه الأبحاث ذات أهمية خاصة في هذا المجال. فقد صاغ كل من رايت وهورن وساندرز Wright, Horn, and Sanders في بحثهم المنشور عام ١٩٩٧، تحت عنوان "أثر المعلم على انجاز الطالب لدى عينتين من المدارس من ولاية تينيسى حيث تناول الباحثون في دراستهم عدداً من المشاركين من مراحل دراسية مختلفة، وتوصلوا إلى فاعلية استخدام نماذج القيمة المضافة فى قياسهم لتأثير المعلمين على إنجاز الطلاب. كما توصل كل من روان وكورتيني وميلر Rowan, Correnti, and Miller (2000)، الى تقدير تأثيرات المعلمين باستخدام نماذج القيمة المضافة على ثلاث مجموعات من الطلاب على عينة قومية تمثل بعض المدارس، حيث وجدوا تأثيرات ما بين متوسطة وكبيرة للمعلمين على تعليم طلابهم كمؤشر لفاعلية أداء المعلمين. كما استخدم ريفكين، هانوشيك وكين Rivkin, Hanushek, and Kain (2000)، بيانات من مدارس تكساس لتحديد الحد الأدنى من الارتباط حول تأثيرات المعلمين ووجدوا أن هناك ثمة تأثيرات للمعلمين (McCaffrey,Lockwood, Kortez, Louis, & Hamilton,2003a). مما يشير إلى أنه قد يكون من المهم اتخاذ خطوة لدفع عجلة تقويم المعلمين، وان يكونوا مسئولين بصورة مباشرة عن أدائهم، وهنا نتساءل **كيف يمكننا أن نقوم بهذا؟ كيف يمكننا قياس أداء المعلمين بشكل فردي وربطه بنتائج طلابهم ومدى تحقيقهم لنواتج التعلم المستهدفه؟ وهل يمكن لإجراءات القيمة المضافة المساعدة فى تقويم وتطوير أداء المعلمين بدرجة كبيرة من الدقة؟**

وبالتالى تظهر أهمية نتائج تقييم القيمة المضافة تأثير المعلمين على تعلم المتعلمين ونمو تحصيلهم. حيث تفترض دراسة هذه النماذج أن تأثير المعلمين كبير جداً من خلال حساب متغيرات متعددة فى نمو أداء المتعلمين من خلال فترة زمنية تتراوح ما بين عامين الى ٤ اعوام. وقد أظهرت الدراسات وجود أدلة أن هناك تأثير واضح لمختلف المعلمين على تحصيل المتعلمين، ولكن هذه الدراسات أظهرت صعوبة فى تحديد حجم تأثير المعلمين بدرجة دقيقة وبالتالى تحتاج هذه التأثيرات إلى مزيد من الدراسات. ومن ثم أصبح هناك متطلب اساسى حول الاجابة على مجموعة من التساؤلات مثل : **ما مدى تحقيق المتعلمين لنواتج التعلم؟، ما النواتج التى تم تحقيقها من خلال مؤشرات نتائج الاختبارات؟، ما مدى التباين فى نتائج المجموعات المختلفة من المتعلمين؟، ما مدى التقدم فى ضوء النتائج السابقة لهم؟ والأهم هنا هو ما مدى إسهام المعلمين فى تحقيق هذه النواتج؟** أى أن تقييم القيمة المضافة يركز على نمو تحصيل المتعلم عبر الزمن. ودرجة الاكتساب gain score التى يتم فيها طرح درجات اختبار سابق من درجات اختبار حالى، أى مقارنة مكتسباته بين نقطتين زمنيتين، وهذا الاكتساب learning gain يعد بمثابة تقدير لإسهام المعلم فى تعلمه.

وقد يقفز إلي الذهن هنا سؤال منطقي وبديهي: **كيف يمكن لمدخل تقييم القيمة المضافة تصنيف الإسهامات التي قام بها المعلم في نمو تحصيل المتعلم، من تلك الإسهامات التي ترجع**

إلي المتعلم نفسه؟، الإجابة هنا ترجع للطريقة الجديدة التي تستخدم في تقييم القيمة المضافة والتي لم تكن موجودة من قبل وهي وحدة التحليل: المتعلم الواحد individual student، أو الفصل الواحد، أو المعلم الواحد، أو غيرها من الوحدات. لأن متابعة نمو التحصيل تتم علي مستوي كل طالب علي حدة بدلاً من مجموعات المتعلمين student cohorts، وهو الأسلوب الذي ساد قبل ظهور وانتشار تقييم القيمة المضافة، خيث تقدير جميع أداءات كل الطلاب هي نقطة البداية baseline التي تستخدم لمقارنة أدائهم بها فيما بعد، وبهذا يتم عزل (أو إلغاء أثر) عدد كبير من المتغيرات التي ترجع لخصائص المتعلم. ولكن بالرغم من ذلك لو تابعنا أداء جميع المتعلمين الذين يقوم نفس المعلم بتعليمهم ووجدنا أن معظمهم يحصل علي درجات أقل من المتوقع، هل نستطيع أن نعزي ذلك للصدفة؟ بالطبع لا، فمن العدل هنا أن نقول أن هناك أدلة قوية تشير إلي أن المعلم قد يكون السبب وراء مستوي التحصيل المتوقع. وقد استفاد مدخل تقييم القيمة المضافة كثيراً من توظيف الأساليب الإحصائية الحديثة في عزل المتغيرات التي ترجح لخصائص المتعلم عن تقييم دور المعلم فى نمو تحصيل المتعلم.

وقد أشار كل من شيرنكفيلد وستفلبيم Shrinkfield & Stuffelbeam (1995)، الى أنه لا يوجد موضوع شغل الكثيرين مثل مدى صدق فاعلية أداء المعلمين اعتماداً على نتائج تعليم الطلاب (مخرجات التعلم). وهو ما أشار اليه كل من ميلمان وسكولوك Millman (1997) &scholoch، من ظهور نقد متزايد حول الربط بين مخرجات التعلم وفاعلية عمل المعلم. ومن ثم ظهرت أهمية اجراء المزيد من الدراسات لصدق مقاييس تقويم المعلم باستخدام نماذج القيمة المضافة لفاعلية التدريس teaching effectsiveness التى تحلل نتائج درجات الطلاب وتقديرات التأثيرات الفردية للمعلمين على درجات الكسب score gains وعلاقة ذلك ببعض المتغيرات المرتبطة بالطلاب مثل المستوى الاجتماعى الاقتصادى وقدرات وسمات الطلاب McCaffrey,Lockwood, Kortez, .Louis, & Hamilton,2003b)

وقد ظهرت مجموعة من طرق قياس فاعلية أداء المعلمين باستخدام نماذج القيمة المضافة وأقدم هذه الطرق هى التى تعتمد على درجات الطلاب فى نفس العام (درجات الكسب) لتقدير فاعلية اداء المعلمين من خلال أداء ونتائج الطلاب، وهذه الطرق لا تضع فى اعتبارها اى عوامل أخرى قد تكون لها تاثير فى أداء الطلاب وكذلك مستوى المعرفة السابقة للطلاب خاصة عندما نكون بصدد المقارنة وترتيب مستوى اداء المعلمين بناء على مستوى الصف الدراسى (Coleman,Campbell& .Kilgore, 1982)

ومن بين التحديات في تقدير تأثيرات المعلمين فصل تأثيرات المعلم عن مصادر التنوع الأخرى في انجاز الطلاب مثل خلفية الطالب أو تأثيرات النظام المدرسي، ومن ثم ظهرت نماذج متعددة المتغيرات من خلال البيانات الطولية، لدراسة تأثيرات متغيرات خلفية الطلاب على نمط معقد

نسبيا في تفاعل عدد من العوامل مثل المستوى الاجتماعى الاقتصادى للطلاب وسمات واستعددات الطلاب المختلفة، والعلاقة بين السمات وتأثيرات المعلم الحقيقية، ونوع النموذج المستخدم. حيث يمكن تحديد مستوى تأثير المعلم بأنه متوسط نمو تحصيل الطلاب لمعلم الفصل مقارنة بجميع المعلمين فى نفس الصف ونفس المادة واضعين فى الاعتبار هذه المتغيرات المصاحبة ودرجاتهم السابقة على اختبارات معيارية مقننة standardsized tests (Crawford, Schlader, Penuel, & Toyama, 2008) (Ballou, Sanders & Wright, 2003) .

وهناك جدل أساسي حول استخدام كل من **نماذج أحادية المتغيرات** (واهمها نموذج درجات الكسب)؛ **ونماذج متعددة المتغيرات** (وأهمها نموذج التأثيرات المتعددة)، مثل دراسات تام، (2003) Thum ، وروان، واخرون (٢٠٠٢)، وروجوسا، (1995) Rogosa؛ وآليسون، Allison (1990)، وبريك وويزبيرج، (Bryk & Weisberg 1976). وقد أظهر لورد (1969) Lord أنه حتى خلال المواقف البسيطة، يمكن أن تعطي هذه النماذج نتائج مختلفة. وبالتالي فإن كلاهما ليس صحيحاً بالضرورة مع كل تطبيق أو أن أي منهما "يتفوق" على الآخر. ويشير بريك وويزبيرج (١٩٧٧) إلى إننا يجب الا ندافع عن منهج على حساب الآخر كوسيلة لتقدير تأثيرات المعلمين وان هذه النماذج تحتاج إلى مزيد من الفحص ودراسة العلاقة بينها من حيث النتائج التى تتوصل إليها (Thum, 2003) , (Rowan, Correnti, & Miller 2002), (Olson, 2002).

وهناك عدد محدود من الدراسات اهتمت بفحص فاعلية أداء المعلم باستخدام القيمة المضافة، فقد وجدت دراسة لاكوود واخرون Lackwood et. el. 2004 ، ان حساسية تقدير تاثيرات المعلمين باستخدام نماذج مختلفة من نماذج القيمة المضافة كانت ضعيفة، اى يوجد اتساق بين نتائج النماذج، فى المقارنة بين تقديرات تاثير المعلمين كما تقاس بنواتج التعلم للطلاب. كما وجد شميتز Schmitz, 2007، اخترفات ضئيلة فى تقديرات النماذج لتأثيرات المعلمين، وقد توصل إلى وجود علاقات متوسطة بين نتائج النماذج المختلفة فى مادتى الرياضيات ومهارات القراءة. بينما أظهرت دراسة تكوى واخرون Teckwe et. el. 2004 ، عدم اتساق النتائج بين نتائج النماذج المختلفة باستخدام نماذج أحادية المتغيرات ومتعددة المتغيرات فى مادة الرياضيات Lockwood, Louis, & McCaffery, (2004), (Tekwe, Carter, Ma, Algina, Lucas, Roth, 2004).

ومن ثم تتحدد الصعوبة فى تحديد النموذج (نماذج أحادية المتغيرات فى مقابل نماذج التأثيرات المتعددة) الذى يقدم صورة واضحة لتقييم فاعلية اداء المعلم، فكل نموذج له نقاط القوة ونقاط الضعف ويقدم وصفاً مختلفاً لتقييم القيمة المضافة، ومن ثم نحتاج إلى دراسة لصدق هذه النماذج والمقارنة بين نتائجها باستخدام البيانات الطولية فى ضوء افتراضات نماذج القيمة المضافة. ومن ثم تحديد النموذج المناسب الذى يقدم بيانات عن انجازات الطالب كمدخلات لحساب تأثير impact المعلمين وترتيبهم ranking بين أقرانهم وفقا لمتوسط كسب الطلاب. ومن ثم معرفة الصدق التقاربى convergent

والتمييزى discriminant لنماذج القيمة المضافة المختلفة عن طريق استخدام محكات مستقلة لقياس فاعلية أداء المعلمين من خلال المتغيرات المرتبطة بالطلاب، وذلك من خلال عمليات تدمج تقديرات متوسط الاداء والكسب مع متوسط اداء كل طالب لمعلم معين من خلال حساب تقدير متوسط الطلاب فى كل مادة وكل صف، ومتوسط اداء الطلاب لنفس المعلم (Tekwe, Carter, Ma, Algina, Lucas, Roth, 2004).

وتتحدد مشكلة الدراسة فى صعوبة تحديد النموذج الذى يقدم صورة واضحة لتقييم فاعلية اداء المعلمين فكل نموذج له نقاط قوة ونقاط ضعف ومن ثم نحتاج إلى دراسات تساعدنا فى التوصل إلى نموذج صادق أواستخدام نماذج مختلفة يكون استخدامها مقنعاً للمعلمين ومن ثم تفسير نتائج تقييم أداء المعلمين فى نسق متكامل من خلال دراسات طولية فى ضوء افتراضات نماذج القيمة المضافة.

إلا أن هذا يتضمن استخدام نتائج الطلاب في تقييم المعلمين، ومن ثم يجب ان نضع في الاعتبار ان تكون هذه الاختبارات دقيقة وعادلة فى التقييم كأهم إجراء واجب اتباعه عند استخدام نماذج القيمة المضافة.، مع الوضع فى الاعتبار أيضا بعض الافتراضات الأساسية مثل اختيار العينة، والخلفية الاجتماعية والاقتصادية للطلاب، واستعداداتهم للدراسة من أجل اتخاذ الأحكام والقرارات الدقيقة.

ومن ثم يمكن تحديد مشكلة الدراسة فى الأسئلة التالية:

١- هل توجد علاقة ارتباطية بين تقديرات فاعلية اداء المعلمين فى كل من مادتى الرياضيات واللغة العربية (كما تقاس بنتائج تقويم نواتج التعلم لطلابهم) من خلال استخدام نموذج الكسب كأحد نماذج القيمة المضافة أحادية المتغيرات؟

٢- هل توجد علاقة ارتباطية بين تقديرات فاعلية أداء المعلمين فى كل من مادتى الرياضيات واللغة العربية (كما تقاس بنتائج تقويم نواتج التعلم لطلابهم) من خلال استخدام نموذج التأثيرات المتعددة كأحد نماذج القيمة المضافة متعددة المتغيرات؟

٣- هل توجد علاقة ارتباطية بين تقديرات نموذج الكسب (كأحد نماذج القيمة المضافة أحادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات) فى تقييم فاعلية أداء المعلمين كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادة الرياضيات؟

٤- هل توجد علاقة ارتباطية بين تقديرات نموذج الكسب (كأحد نماذج القيمة المضافة احادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات) فى تقييم فاعلية اداء المعلمين كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادة اللغة

العربية؟

٥- هل توجد علاقة ارتباطية بين ترتيب المعلمين وفقا لتقديرات فاعلية ادائهم (كما تقاس بنتائج تقويم نواتج التعلم لطلابهم لكل من مادتى الرياضيات واللغة العربية) من خلال استخدام نموذج الكسب كأحد نماذج القيمة المضافة أحادية المتغيرات؟

٦- هل توجد علاقة ارتباطية بين ترتيب المعلمين وفقا لتقديرات فاعلية ادائهم (كما تقاس بنتائج تقويم نواتج التعلم لطلابهم لكل من مادتى الرياضيات واللغة العربية) من خلال استخدام نموذج التأثيرات المتعددة كأحد نماذج القيمة المضافة متعددة المتغيرات؟

٧- هل توجد علاقة ارتباطية بين ترتيب المعلمين وفقا لتقديرات نموذج الكسب (كأحد نماذج القيمة المضافة أحادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات) فى ضوء تقييم فاعلية ادائهم كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادة الرياضيات؟

٨- هل توجد علاقة ارتباطية بين ترتيب المعلمين وفقا لتقديرات نموذج الكسب (كأحد نماذج القيمة المضافة أحادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات) فى ضوء تقييم فاعلية ادائهم كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادة اللغة العربية؟

هدف الدراسة :

تهدف هذه الدراسة الى دراسة تقييم فاعلية اداء المعلمين من خلال ربط فاعلية ادائهم بمدى تحقيق طلابهم لنواتج التعلم فى مادتى الرياضيات واللغة العربية ودرجة النمو الحادثة من خلال دراسة طولية لمدة ثلاث سنوات لطلاب الصفوف الثلاثة الأولى من التعليم الابتدائى، ومعرفة مدى الاتساق بين نتائج التقييم باستخدام نموذجين من نماذج القيمة المضافة وهما نموذج الكسب (كأحد نماذج القيمة المضافة أحادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات).

اهمية الدراسة :

١- قد تفيد هذه الدراسة فى تقديم معلومات نظرية واحصائية حول استخدام نماذج القيمة المضافة فى تقييم فاعلية أداء المعلمين من خلال عرض نماذج القيمة المضافة وافتراضتها وطرقها الإحصائية باختلاف اتجاهاتها، كوسيلة تقييم جديدة بجانب الطرق المتعارف عليها بما قد ينعكس على رفع أداء المعلمين من أجل تحقيق نواتج التعلم المستهدفة لطلابهم.

٢- ان دراسة نماذج تحليل القيمة المضافة قد تساعد الباحثين والقائمين على عمليات تقويم المعلمين فى معرفة إسهامات المعلمين فى تقدم المتعلمين من خلال قيمة النمو السنوي

لهم مقارنة ببدايته الأولى كخط أساس baseline، اى أنها قد تقدم مؤشرات عادلة لا يوجد بها تحيز Bias لنتائج تقييم فاعليةأداء المعلمين من خلال نتائج طلابهم، بما يساعد فى اتخاذ قرارات تربوية تستند إلى بيانات واقعية طولية وذلك بالنسبة لكل طالب وكل معلم.

مصطلحات الدراسة :

١- **القيمة المضافة:** طريقة لتحليل بيانات الاختبارات التحصيلية ومراجعة نتائجها فى الفِرَق السابقة، للتنبؤ بالنمِو المتوقع للمتعلمين، اى أنها طريقة لتحليل بيانات اختبارات الطلاب للتأكد من نمو الطلاب فى التعلم بمقارنة الوضع الحالي للطلاب بأداءاتهم السابقة. ويستفاد من هذه النتائج فى التقييم الموضوعي للمعلمين.

٢- **نماذج القيمة المضافة أحادية المتغيرات:** ويطلق عليها اسم "نماذج تهيئة المتغيرات المصاحبة" و"نماذج درجات الكسب"، حيث تحدد هذه النماذج التقديرات الحالية كنتيجة للتقديرات السابقة والربط الصريح بين درجات الطلاب وتأثيرات معلميهم الحاليين فقط من خلال تحديد الدرجات الحالية كوظائف خطية للمتغيرات، **ومن أمثلتها نموذج الكسب.**

٣- **نماذج القيمة المضافة متعددة المتغيرات:** يتسم المفهوم التحليلي الرئيسي الثاني من خلال الربط بين تقديرات الطلاب وتأثيرات المعلمين خلال السنوات المتعددة واضعين فى الاعتبار المتغيرات المصاحبة مثل استعداد الطلاب للدراسة والمستوى الاجتماعى الاقتصادى لهم .(Haegeland, & Kirkeboen, 2008)، **ومن أمثلتها نموذج التأثيرات المتعددة.**

٤- **نمو أداء الطلاب:** هى التغيرات التى حدثت فى أداء كل متعلم على حدة، خلال فترة زمنية محددة، وذلك بفحص الزيادة التى حدثت فى تحصيل كل طالب من اختبار الى آخر فى مادة دراسية معينة، وتتبع هذه الزيادة بانتقال المتعلم من فرقة الى أخرى.

٥- **فاعلية اداء المعلم:** يمكن تحديد فاعلية أداء المعلم فى ضوء هذه الدراسة بأنها متوسط نمو التحصيل المقدر لطلاب معلم الفصل مقارنة بجميع المعلمين فى نفس المادة ونفس الصف الدراسى، واضعين فى الاعتبار الخلفية الاجتماعية والاقتصادية للطلاب وقدراتهم وسماتهم الشخصية وتقديراتهم السابقة على الاختبارات المعيارية المقننة standardized tests بما يسمح بازالة التباين ويعطى مزيداً من الدقة فى المقارنات بين المعلمين.

الإطار النظرى والدراسات ذات الصلة:

لقد اهتم خبراء التربية بأساليب تقويم المعلم لأهمية دوره فى العملية التعليمية لضمان جودة مخرجات التعلم. وقد تعددت هذه الأساليب ففى بعض الدراسات يطلب من مديري المدارس أوالموجهين التربويين أن يحددوا أفضل المعلمين وأسوؤهم، وفى البعض الآخر يقومون بترتيب المعلمين حسب الجدارة، وفى البعض الثالث تستخدم طرقاً أكثر تفصيلاً تتطلب إعطاء وزن وقيمة لكل عنصر من عناصر التقييم (McCaffrey, Lockwood, Kortez, Louis, & Hamilton,2003b).

وقد استخدمت فى السنوات الأخيرة **أساليب وطرق متعددة لتقويم المعلم منها**

أ) تحليل الموقف التعليمى داخل الفصل والتفاعل بين المعلم والتلاميذ، وذلك عن طريق:

١- **طريقة التدريس المصغر** Microteaching، وفيها يعطى المعلم موضوعاً محدوداً وقصيراً يستغرق ١٠ دقائق فقط ويسجل هذا الموقف بشريط فيديو على أن يعرض للنقد والتحليل لتقويم مدى نجاح المعلم.

٢- **ملاحظة التفاعل الحادث فى بيئة الفصل،** وذلك بين المعلم وتلاميذه من حيث الأسئلة والأجوبة وطريقة الحوار الذى يدور بين المعلم وتلاميذه.

٣- **الملاحظة عن طريق العرض التليفزيونى،** من خلال عرض لدرس نموذجى على معلم الفصل وبتوجيه أسئلة له لمعرفة رأيه فى بعض المواقف التعليمية.

ب) تقويم المعلم عن طريق كفاياته، يتفق معظم المهتمين بسيكولوجية التعلم على أن المعلم الكفء، هو الذى يحدث التغيرات المطلوبة فى سلوك التلاميذ فى إطار الأهداف التربوية، وبالتالى يتوقف مدى تأثير المعلم فى تلاميذه ومدى تحقيقه للأهداف المرغوبة على مدى تمكنه من الكفايات التدريسية.

إلا أن الحكم على كفاية المعلم بمحك تعلم التلاميذ ليس كافياً نظراً لوجود متغيرات عديدة تؤثر فى هذا المحك مثل ذكاء التلاميذ وطبيعة المادة الدراسية ومدى توافر الإمكانات المدرسية وغيرها من العوامل.

ج) تقويم المعلم عن طريق تحديد أدواره، من الأساليب المستخدمة فى تقويم المعلم، تحديد الأدوار المنوطة بالمعلم والتى تمكنه من أداء مهامه الوظيفية ومسئولياته المهنية. ورغم اختلاف هذه الأدوار حسب الظروف الثقافية بالمجتمع وطبيعة المادة الدراسية، ورغم تداخل هذه الأدوار وتكاملها، فقد حدد المتخصصون أدوار المعلم التى تساعده فى أداء مهامه المختلفة بكفاءة بأنه: ناقل للمعرفة، ومسئول عن النمو المتكامل للتلاميذ، ومتابع للتغيرات الحديثة فى محتويات المناهج الدراسية، ومشارك فى عمليات التجديد التربوى، ومشارك فى عملية الإدارة المدرسية، ومسئول عن حفظ النظام،

ومسئول عن تقويم المتعلمين، ومشترك فى تخطيط المناهج، ومنفذ للمنهج.

د) تقويم المعلم عن طريق تقديرات رؤسائه، ويعتمد هذا الأسلوب على استطلاع رأى كل من مدير المدرسة والمدرس الأول والموجه الفنى للوقوف على ما يتمتع به المعلم من سمات شخصية ومدى مشاركته فى العمل المدرسى ومدى تأثيره فى المجتمع التربوى بصفة عامة. وكذلك الوقوف على إتقانه للمادة العلمية وقدرته على التدريس.

هـ) تقويم المعلم عن طريق تقديرات الزملاء، وتتضمن هذه الطريقة التعرف على آراء زملاء المعلم من المعلمين للتعرف على النواحى المختلفة المرتبطة بطابع شخصية المعلم، والمتعلقة بأدائه المهنى. وتفيد هذه الآراء فى إعطاء صورة صادقة عن علاقة المعلم بزملائه ومركزه فى الأسرة المدرسية.

و)تقويم المعلم من خلال تأثيره فى التلاميذ، فالتلاميذ أكثر صلة بمعلمهم فى المواقف التعليمية المختلفة، وغالبا ما تكون أحكام التلاميذ على معلمهم ذات دلالة كبيرة، خصوصا إذا كان هؤلاء التلاميذ على درجة كافية من النضج. ويتم التعرف على آراء التلاميذ حول معلمهم عن طريق استفتاءات بدون ذكر أسمائهم أو بعد تخرجهم.

ز) تقويم المعلم من خلال رأيه فى نفسه (التقويم الذاتى، يستلزم تقويم المعلم دراسة مقتضيات ومتطلبات التدريس، ومن ثم تحديد المؤهلات الواجب توافرها فى المعلم والتى فى ضوئها يمكن تحديد واجباته التى ينبغى أن يقوم بها. وقد أعدت قوائم وبطاقات لتقويم المعلم ذاتياً. وإذا كانت تلك القوائم والاختبارات لها أهميتها فى تقويم المعلم إلا أنها صممت بحيث تقيس الصفات المرغوبة فى المعلم سواء كانت تلك الصفات شخصية أم مهنية أو تمثل استعدادات لدى القائم بالتدريس , (McCaffrey, Lockwood, Kortez, Louis, & Hamilton,2003a) (Kelly, 2012) .

يتضح مما سبق أن أكثر أساليب تقويم المعلم شيوعاً هو أسلوب تقويم المعلم من خلال ملاحظة وقياس كفاءاته المهنية والتدريسية.

استخدام القيمة المضافة فى تقييم أداء المعلمين:

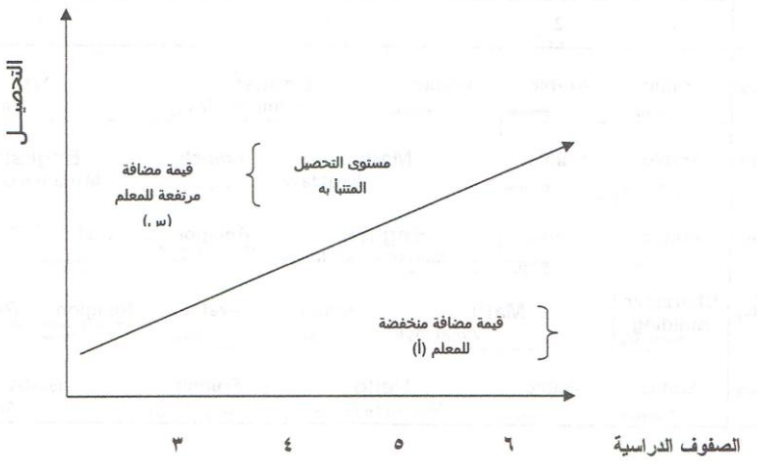
مفهوم القيمة المضافة فى المجال التربوى

يُعرِّف ساندرز (Sanders, (2000، القيمة المضافة بأنها طريقة لتحليل بيانات الاختبارات مع مراجعة نتائجها فى الصفوف السابقة للتنبؤ بنمو المتعلمين، والتعرف على أسباب هذا التوقع لنواتج التعلم المستهدفة، كما يمكن التعرف على إسهام المؤسسة فى تقدم المتعلمين نحو تحقيق أهداف التعليم المحددة (مثل، التحصيل المعرفي cognitive achievement). وبغض النظر إذا

كان المتعلم قد بدأ العام الدراسي بمستوى يفوق فرقته الدراسية أو أدنى منها. فالنمو السنوي المتوقع للمتعلم يعد فحوى منهجية القيمة المضافة، وبذلك تلقى نتائج هذا التقييم الضوء على إسهام المؤسسات التعليمية فى تقدم تعلم كل طالب (Sanders, 2000).

ويعد مفهوم تقييم القيمة المضافة نموذج نمو يُستخدم فى تحليل نتائج تقييم المتعلم بطريقة تحدد "القيمة value " التي تسهم بها المؤسسة التعليمية فى تقدُّم تعلُّمه خلال مدة دراسية معينة، بغض النظر عما إذا كان قد بدأ العام الدراسي بمستوى يفوق فرقته أو أدنى منها، فالنمو السنوي المتوقع للطالب يعد الأساس فى تقييم القيمة المضافة (Braun,2006),(Wainer, 2004) .

وهناك عدد من الطرق تستخدم لتقييم القيمة المضافة لأداء المعلم علي تحصيل المتعلمين، اعتماداً علي درجات الاختبارات التحصيلية المقننة في نهاية العام الدراسي بعد الضبط الإحصائي لأثر مستوي المتعلم المبدئي في بداية العام الدراسي، وتعتمد كثير من هذه الطرق علي نماذج الانحدار، ويستخدم نوع من النماذج الإحصائية تسمي النماذج الخطية الهرمية Hierarchical Linear Models (HLM) وهذه النماذج لا تفترض ثبات المعلم علي جميع المتعلمين، بل تسمح بتقدير أثر المعلم علي كل طالب.



شكل (١) القيمة المضافة للمعلمين

ويوضح شكل (١) مدرسة ابتدائية بها معلمين اثنين فقط في الصف الرابع هما المعلمة/ س، والمعلمة/ أ، فنجد المعلم س فوق المستوى المحدد من تحقيق مستوى الكفاءة بالمعايير، اى ينتجون

القيمة المضافة بمعدل أعلى من المتوسط، وأما المعلمة/ أ، فإنها تحت المستوى المحدد من تحقيق مستوى الكفاءة بالمعايير(Hershberg, 2005). وذلك بما يعكس التوافق العام في الآراء بين الباحثين على أن فعالية المعلم تختلف اختلافًا كبيرًا حتى داخل المدارس. وهذا هو أحد الأسباب التي تجعل هناك اهتماماً كثيراً بالقيمة المضافة للمعلم. وأنه من الممكن تحديد المعلمين الموجودين تحت الخط ومساعدتهم على تحسين أدائهم وتنميتهم مهنياً (Harris,2011).

أسس القيمة المضافة :

أ – يسمح الاختبار السنوى للمتعلمين للباحثين بتتبع "النمو المتوقع" للمتعلمين.

ب– يحقق المعلمون الذين حقق متعلموهم أكثر من المتوقع "قيمة مضافة مرتفعة".

ج– لا تعد القيمة المضافة اختبار إنما طريقة إحصائية لحساب مقدار النمو للمتعلمين عبر فترة زمنية محددة من خلال نتائج الاختبارات.

وبالتالي هناك مخرجات للقيمة المضافة تتمثل فى:

١– نمو المتعلم خلال العام وعبر السنوات السابقة.

٢– نتائج الفصول اعتمادا على نتائج المتعلمين خلال السنوات الثلاثة الأخيرة.

٣– تقديم إسهامات المعلم فى نمو المتعلمين مع الأخذ فى الاعتبار العوامل التى تخرج عن نطاق عمل المعلم والمؤسسة مثل العوامل الاجتماعية والاقتصادية.

٤– تقدم تقديرات لأداء المتعلمين الخاصة بكل معلم، وبالتالي يمكن وصف الفروق فى أداء المعلم عبر فترة زمنية معينة، مع توقع ارتفاع متوسط أداء المتعلمين فى فصولهم مما يشير إلى "فاعلية المعلم" (مجدى قاسم وآخرون، ٢٠١١).

افتراضات نماذج تقييم القيمة المضافة :

هناك افتراضات أساسية على الأقل ينبغى أن تتوافر قبل استخدام أى من هذه النماذج وهى (Sanders, 2000):

١– استخدام الاختبارات المعيارية المقننة فى تقييم القيمة المضافة

إن المدارس التى تعتمد فى سياستها على التقويم البنائي تكون أكثر فاعلية. لذلك، من المهم أن يكون هناك تقويم بنائي وتقييم نهائي فى كل مدرسة بشكل متوازن، مع استخدام اختبارات **محكية المرجع** لتفسير النتائج (التقييم من أجل اتقان التعلم) مما يزيد من فاعلية عمليتي التعليم والتعلم.

تقييم القيمة المضافة تعتبر نظرة جادة ذات قيمة، وأداة جديدة نسبياً لفحص تقديرات الطلاب

على الاختبارات التي تعتبر طريقة للتأكد من أن جميع الطلاب قد تعلموا. كما تقدم طريقة موضوعية لتقويم المعلمين لأنها تركز على نمو الطلاب خلال فترة زمنية طويلة، أكثر من التركيز على الدرجة الحالية للاختبارات التحصيلية، فتقييم القيمة المضافة ليست طريقة لتقييم الوضع الحالي للطلاب، ولكنها طريقة لمعرفة مسار نمو الطالب عبر فترة زمنية محددة.

والقيمة المضافة هنا تركز فقط على تحليل نتائج الاختبارات ولكنها لها فوائد كثيرة منها:

- **التحديد:** تقييم القيمة المضافة تغير تركيز إحصاءات التعليم من التحليل وفقاً للمتغيرات الديمجرافية، إلى الإجابة على سؤال جوهري هو: ما مدى تقدم وتحسن أداء الطلاب؟.
- **المقارنات العادلة:** عن طريق التركيز على نمو الطالب، تقدم القيمة المضافة وسيلة لفهم نمو الطلاب وتحقيق المعلم لأهداف التعلم. وهذه مهمة للمدارس التى يوجد بها مجموعة من الطلاب منخفضي التحصيل (انور محمد الشرقاوى ، ١٩٩٦).

٢- إجراء تكافؤ لدرجات الاختبار Test Score Equating :

يعتمد مدخل القيمة المضافة على المقارنة بين تقديرات اختبارية طولية عبر الزمن خلال عدة سنوات متلاحقة. ولكن يفترض ضمنياً فى أى نموذج من نماذج القيمة المضافة أن تتدرج جميع هذه التقديرات الاختبارية على ميزان تدرج واحد one common scale لتحقيق نفس الغرض، وهو عدالة المقارنات التى تتم من خلال هذه النماذج. وبالتالي فإنه قبل استخدام تلك التقديرات الاختبارية لابد من إجراء تكافؤ رأسي vertical scaling لها، وإجراء تكافؤ أفقي horizontal scaling لدرجات الاختبارات المختلفة فى الفرقة الدراسية الواحدة وهو ما تحققه نظرية الاستجابة للمفردة فى تدريج المقاييس.

٣- البيانات المفقودة Missing Data لا تؤثر فى النتائج :

نظراً لأن تقييم القيمة المضافة يعتمد على جمع كم كبير من البيانات، فإنه ليس من المستغرب أن تكون بعض هذه البيانات مفقودة Missing Data لسبب أو لآخر، كأن يغادر المتعلم المؤسسة وينتقل لمؤسسة أخرى ثم يعود مرة أخرى إلى نفس المؤسسة، مما يعنى فقدان بيانات تحصيل المتعلم خلال فترة غيابه، أو نتيجة لغياب بعض المتعلمين عن الاختبار، أو عندما يحتمل أن يؤثر حجم البيانات المفقودة (كأن يكون حجم البيانات المفقودة ١٠% أو أكثر) فى دقة النتائج مثل تأثيره فى تباين الدرجات وفى تحيز البيانات، ولذلك تفترض نماذج القيمة المضافة أن البيانات المفقودة ترجع إلى أسباب عشوائية missing at random بمعنى أنها لا تتعلق بالتأثيرات التي تم تقديرها، أو إلى تباين البيانات، أو تحيز النتائج بعامة. (Tekwe, Carter, Ma, Algina, Lucas, Roth, 2004).

٤- بقاء أثر المعلم فى المتعلم Persistence of Teacher Effect :

تفترض نماذج القيمة المضافة بقاء أثر المعلم فى المتعلم عبر الأعوام الدراسية المتعاقبة، فإذا كان قد حصل أحد المتعلمين فى إحدى المؤسسات على كسب تحصيلى فى الفرقة الرابعة مقداره خمس وحدات فوق متوسط اداء المتعلمين فى نفس المنطقة التعليمية، فإنه من المتوقع أن يكون الكسب التحصيلى لنفس المتعلم خمس وحدات، حتى لو كان معلم الصف فى الفرقة الخامسة معلماً متوسطاً من حيث القيمة المضافة. ولا يعد هذا الافتراض من الافتراضات صعب التحقيق فى معظم البيانات (Doran, and Cohen.2005).

٥- تحديد العوامل التى تقع خارج نطاق سيطرة من يخضع للتقييم:

٥-١ **العوامل الاجتماعية الاقتصادية فى علاقتها بنمو التحصيل:** بالرغم من أن درسات عديدة تؤكد علي أن المستوي الاقتصادي الاجتماعي من أهم العوامل المؤثرة في تحصيل المتعلمين، فعلي سبيل المثال، توصل هيرشبرج وزملاءه (Hershberg, Simon, & Lea-Kruger, 2004) إلي أن المعلم يعتبر أحد أهم المنبئات ذات الكفاءة المرتفعة بنمو تحصيل المتعلمين حتى بالمقارنة بغيرها من المتغيرات التي كانت سابقاً تعتبر علي درجة عالية من الأهمية مثل مستوي دخل الأسرة. وهنا تبرز أهمية استخدام المدخل الإحصائي المناسب للقيام بذلك، إلا وهو تقييم القيمة المضافة. (Braun, 2005), (Rivikin, 2007).

٥-٢ **استعداد الطلاب للدراسة** : . فالاستعداد هو الاستطاعة أو القيمة التنبؤية للقدرة إذ أن الاستعداد الخاص هو إمكانية تحقيق القدرة بمعنى أن الاستعداد سابقة على القدرة وهو لازم لها، فالقدرة ما هى إلا قدح للاستعداد بعوامل البيئة (أحمد زكى صالح، ١٩٩٢: ٦٥٦). وقياس الاستعداد يساعد على التنبوء بالأداء اللاحق، أي ما يمكن للفرد أداؤه مستقبلا إذا ما أتيحت له الظروف المناسبة. (Childs & McKay, 2001)

القيمة المضافة وقياس نمو المتعلمين

ومن المهم ان نتفهم هنا أن "التقييم من أجل التعلم ASSESSMENT FOR LEARNING " لا يأتى نتائجه الفعالة من خلال نتائج تقييم أسبوعية أو شهرية أو نصف سنوية أو حتى سنوية، إنما من خلال **مدى زمنى أطول** وعبر تقييمات متعددة (القيمة المضافة) (Stiggins, 2005).

(١) الأداء والنمو:

الأداء هو ما تسهم به المدرسة والمعلم فى مخرجات التعلم. وهذا يتطلب القياس الدقيق للأداء الذي يضع مقارنات جيدة وعادلة بين المدارس وبين المعلمين. إن الصعوبة الرئيسية فى استخدام تقدير اداء الطلاب فى الاختبارات لتقويم المعلمين، وذلك من خلال مقارنة أداء الطلاب هذا العام بالعام السابق. فإذا حصل المعلمون على تقديرات أداء منخفضة، يكون ذلك نتيجة أن طلابهم حصلوا على درجات منخفضة فى العام السابق كنتيجة لجهود التعلم الضعيفة فى هذا العام السابق (Harris,

.2011)

(٢) قياس الأداء:

تتحدد مستويات التحصيل التى تعد مقبولة أو متميزة استناداً إلى تحليل ما يجب أن يعرفه الطلبة، وما يمكنهم أداؤه، وليس استناداً إلى الأداء النسبي للطلبة مقارنة بأقرانهم. وهذا يتطلب تقديم أمثلة واقعية عن مستوى ونوعية أداء الطلبة الذي يجب أن تتضح للتحقق من تمكنهم من نواتج المحتوى Content Outcomes (Bresciani, M., J., 2006) .

(٣) تقويم الأداء:

يتم تقويم الأداء Performance Assessment، أو التقويم القائم على الأداء Performance based Assessment، عن طريق الأهداف المتمركزة حول سلوك المعلم، ويمكن النظر إلى الهدف، وناتج التعلم من هذه الزاوية كوجهين لعملة واحدة. فالهدف يرتبط مباشرة بالناتج أو السلوك النهائي للمتعلم، ويعبر عنه كما يجب أن يكون، وهذا الأخير يتبع الهدف، ويتجه نحوه، وتكون الأهداف تبعاً لذلك بمثابة المدخلات والنواتج، أو المحصلات التعليمية بمثابة المخرجات لعملية واحدة هى عملية التعلم (Harris, 2011).

(٤) تحديد مستويات الأداء

تعد الدرجات المعيارية الأفضل لقياس نمو الطلاب، حيث تعطى طريقة لمقارنة القياس بين المدارس والطلاب المختلفة، وطريقة للمقارنة قى ضوء قياس "الكفاءة" أو "الإتقان".

ربط الطلاب بالمعلمين:

تقدر تأثيرات المعلم من خلال ملاحظة ناتج طلابهم، ويتطلب التقدير ربط الطلاب بمعلميهم. فإذا حضر جميع الطلاب مدرسة واحدة فقط وتعلموا بواسطة معلم واحد فقط في كل عام، فإن ربط اداءهم بأداء المعلم سوف يوفر ببساطة سجلات إدارية (فئة القوائم) يتم الحفاظ عليها بدقة. مع ذلك، فإن العالم الحقيقي ليس بهذه البساطة، فيغير الطلاب المدرسة في منتصف العام كما يغيرها المدرسون أيضا. يدرس بعض المعلمين في فريق أو يتشاركون المسئولية لتعليم مادة دراسة ما.

يوجد العديد من افتراضات وضع النماذج المحتملة للتعامل مع المشكلة المعقدة لربط اداء الطلاب بأداء المعلمين. يمكن الحفاظ على الدرجات القبلية لتحديد مشاركة كل معلم لمدخلات الطلاب، وكلا المدرسين يمكن إدراجه في النموذج من أجل إنجاز الطلاب، أو النمو الموزون من خلال مساهمتهم في المدخلات. على نحو مشابه، وهناك افتراض آخر أن الطلاب الذين يحولون أو لديهم فريق تدريس يتمتعون بناتج مماثل في التوزيع للطلاب الآخرين، إذن فيمكن استثناء هؤلاء الطلاب من البيانات، ولذا يمكن وضع نموذجهم دون رابط لأي معلم .(Goldhaber, & Brewer, 2000) .(Tucker, & Strong 2005)

توفر العديد من تطبيقات نماذج القيمة المضافة إما ذكرا غير صريح للنماذج من أجل وضع نماذج للطلاب الذين يحولون أو بطريقة أخرى ليس لديهم روابط بالمعلمين (Rowan, Correnti, & Miller, 2002, Wohlstetter, Datnow,. & Park, 2008) ، أو تستثني الطلاب ذوي البيانات غير مكتملة (Rivikin, Hanushek, & kain, 2000) .

نمذجة البيانات الطولية لتقدير تأثيرات المعلمين

أن استخدام تقدير تأثيرات المعلمين من خلال تشكيل البيانات الطولية longitudinal حول إنجاز الطلاب أظهر عدداً من القضايا الإحصائية والسيكومترية المهمة حول مدى فاعلية تقديرات تأثيرات المعلمين. من ثم **تتمحور الأسئلة أولاً** حول إمكانية أن تكون التقديرات الإحصائية لتأثيرات المعلمين قريبة من التأثيرات الواقعية حتى يتم فهمها وإدراكها؛ **ثانياً**، حول ارتباط هذه التقديرات نسبياً بإجراءات التقديرات الأخرى، وفي حالة اختلاف التقديرات عبر النماذج، يجب في هذه الحالة مراعاة الشك المترتب على هذه النتائج؛ **أخيراً**، يجب ألا ترتبط هذه الأخطاء في التقدير بخصائص الطلاب أو المعلمين أو غيرها من العوامل الأخرى إلى الحد الذي يؤدي إلى انحراف التأثيرات التقديرية عن الواقعية، والذى قد يؤدي إلى تشويه الهدف من استخدام نماذج القيمة المضافة، والذي يتمثل في عزل تأثيرات المعلمين عن العوامل الأخرى المؤثرة على إنجاز الطالب (Tucker, & Strong 2005).

النماذج الأساسية لتحليل البيانات المطولة لتقدير تأثيرات المعلمين:

تتراوح نماذج تقييم القيمة المضافة بين النماذج البسيطة التي تفترض أن جميع التأثيرات ثابتة، إلي النماذج المعقدة التي تسمي بالنماذج ذات التأثيرات المختلطة mixed-effects models والتي لا تفترض أن جميع التأثيرات ثابتة. وجميع هذه النماذج تستخدم درجات الطلاب على الاختبارات التحصيلية المقننة، أو درجات الكسب فى الاختبارات التحصيلية gain score، كمتغيرات تابعة، أو نواتج outcomes لأثر المدرسة، والمعلم... وغيرها. وجميع هذه النماذج تسعي لتقدير التأثيرات السببية causal effect للمعلم لتحديد ما "أضافه" المعلم من "قيمة" لتعلم الطلاب (Deb Roy, & Bates, 2003).

ويوجد اتجاهين لتحليل البيانات المطولة لتقدير تأثيرات المعلمين:

١- **النمذجة التسلسلية للنتائج الأحادية** (أحادي المتغيرات) أو "نماذج تهيئة المتغيرات المصاحبة": التي تعكس الدرجات الحالية اعتماداً على اداءهم السابق، حيث يتم تقدير اداء الطلاب في الفصول سنوياً.

٢- **النمذجة المشتركة لدرجات السنوات المتعددة** (متعدد المتغيرات): من خلال تحديد تأثيرات المُدخلات الإضافية- في حالة النظر إلى التأثيرات الإضافية للمعلمين- حول الإنجاز

التعليمي المسئول عن الانجاز السابق.

ونظراً لاختلاف الفروض، تنضوي هذه التوجهات على عدد من جوانب القوة والضعف المختلفة، ولا يوجد أي منها يتفوق على الآخر تحت كافة الظروف. لكن بصفة عامة، يعتبر النموذج متعدد المتغيرات هو النموذج المفضل عادةً كنموذج إحصائي لتقدير تأثيرات المعلمين نظراً لمرونته وكفاءة استخدام البيانات المتاحة.

اولاً: **نماذج أحادية المتغيرات**، تعد هذه النماذج أقل نماذج القيمة المضافة تعقيداً، وتتميز عملياته الحسابية بالبساطة، ويسهل فهم نتائجها. ويطلق عليها اسم "نماذج تهيئة المتغيرات المصاحبة" و"نماذج درجات الكسب". تحدد نماذج تهيئة المتغيرات تقدير اداء الطلاب الحالى كنتيجة لادائهم السابق والربط الصريح بين درجات الطلاب وتأثيرات معلميهم الحاليين فقط من خلال تحديد الدرجات الحالية كوظائف خطية للمتغيرات. ويعتمد هذا النموذج على طرح درجة سابقة لكل متعلم فى اختبار مادة دراسية معينة من تقدير حالى على اختبار آخر فى نفس المادة للحصول على درجة الكسب gain score، ثم يُحسب متوسط الكسب للصف المدرسي، ومن أمثلة الدراسات التى استخدمت هذه النماذج كل من روان، كورينتي وميلر (2002) **(RCM)**,Rowan, Correnti & Miller؛ وساندرز وريفرز (1996)Sanders & Rivers ؛ ودالاس DVAAS؛ ويبستر وميندرو Webster & Mendro (1997) Sanders, Saxion, Schneider, Dearden, & Horn 2002.

وتفترض هذه النماذج ما يلي:

$$yt= mt+ byt-1+ Tt+et \qquad (4.1)$$

حيث ترمز yt إلى درجة الطالب خلال الزمن t= 1,2,…..,p، بينما تشير mt إلى متوسط درجة الطالب التي ربما تعتمد على خصائص الطالب وغيرها من المتغيرات الأخرى، وتكون Tt هي تأثيرات المعلمين وet هي الأخطاء التراكمية. وتستخدم هذه النماذج بيانات كل عام دراسي على حده بحيث يكون الاستخدام الوحيد للبيانات من السنوات المتعددة هو من خلال متغير العام السابق.

بدلاً من ذلك، تحدد نماذج درجات الكسب التحسن في درجات عام واحد فقط (الدرجة الحالية الأقل من الدرجة السابقة) بصورة منفصلة لكل عام دراسي وربط تحسن الطلاب بتأثيرات المعلمين خلال العام الحالي. وخلال نماذج القيمة المضافة الحديثة (مثل RCM)، يتم قياس هذا التحسن من بداية صف دراسي إلى بداية الصف الدراسي التالي، ويفترض نموذج التحسن في الدرجات ما يلي:

$$Yt-yt-1= mt+Tt+et \qquad (4.2)$$

حيث تشبه المختصرات المستخدمة تلك الخاصة بنموذج المتغيرات. ونظراً لصياغة كل

سنة بطريقة منفصلة عن الأخرى، وتستخدم المفاهيم المتشابهة (مثل: المتوسط أو تأثير المعلمين).

ثانياً: النمذجة متعددة المتغيرات،

يتسم المفهوم التحليلي الرئيسي الثاني لتقدير تأثير المعلمين والذي نطلق عليه اسم "النمذجة متعددة المتغيرات" من خلال الربط بين درجات الطلاب وتأثيرات المعلمين خلال السنوات المتعددة. وتشمل الأمثلة الحالية لهذا المفهوم نموذج نظام **تينيسى** TVAAS لتقييم القيمة المضافة **متعدد الطبقات** والنماذج مشتركة التصنيفات لروان وكورينتي وميلر (٢٠٠٢)، وروديبوش وبريك Raudenbush & Bryk (2002)، ماكفري ولوك وود McCaffrey & Lockwood (2003).

وتشير النماذج المطولة متعددة المتغيرات للزمن بالرمز t، تدل قيمة yit على درجة الطالب، بينما تدل mt على متوسط درجة العينة وTt إلى معلم الطالب وeit إلى مصطلح الخطأ التراكمي. ونعرض النماذج على مدى ثلاث سنوات من البيانات لكل مادة دراسية (مثل الرياضيات أو القراءة).

ويفترض نموذج التصنيف التقاطعي أن إنجاز الطالب ينمو خطياً بمرور الوقت وأن معايير هذا الاتجاه التنموي ترتبط بالطلاب. (Raudenbush & Bryk, 2002)، ويصبح النموذج على النحو التالي:

$$yi1= mi+ bi+ T1+ei1$$
$$yi2=mi+2bi+t1+T2+ei2$$
$$yi3= mi+3bi+T1+T2+T3+ei3 \qquad (4.3)$$

ويفترض أن تكون تأثيرات المعلمين مستقلة عن المتغيرات الأخرى في النموذج. ويستخدم نموذج نظام تينيسى TVAAS لتقييم القيمة المضافة ما يُعرف باسم النموذج **متعدد الطبقات** لحساب السنوات التالية لتأثيرات المعلمين والتي تعتمد على طبقات السنوات السابقة على النحو التالي:

$$yi1= m1+ T1+ei1$$
$$yi2= m2+ T2+T1+ei2$$
$$yi3= m3+T3+T2+T1+ei3 \qquad (4.4)$$

تشمل السمة الأساسية للنموذج متعدد الطبقات مثل نموذج نظام تينيسى لتقييم القيمة المضافة الأساسي بأن المعلمين يؤثرون على انجاز الطلاب عندما يكون هؤلاء الطلاب في فصولهم، ويستمر هذا التأثير بنفس القوى على مدى جميع سنوات الاختبار في المستقبل (Sanders, Saxion, & Horn, 2002, Ballou, Sanders, & Wright, 2003).

وقد قدم كل من ماكفري وآخرين McCaffrey, et all (2003) بديلاً للنموذج متعدد الطبقات يسمح بتقدير قوة استمرار تأثيرات المعلمين خلال السنوات التالية. ويمكن صياغة النموذج

على النحو التالي:

$$yi1= m2+T1+ei1$$
$$yi2= m2+T2+a21T1+ei2$$
$$yi3= m3+T3+a32T2+a31T1+ei3 \quad (4.5)$$

وتحدد معايير a21 و a32 وa31 استمرارية تأثيرات المعلمين كما كانت في السنة السابقة على درجات العام الحالي. نطلق على ذلك النموذج اسم "نموذج الاستمرار" لأنه يصيغ استمرارية تأثيرات المعلمين.

فصل تأثيرات المدرسين والمدارس السابقين عن تأثيرات المعلم المقدرة.

تشمل النماذج متعددة المتغيرات تأثيرات المعلمين السابقين بالإضافة إلى تأثيرات المعلمين الحاليين في النموذج من أجل الدرجات كل عام. (McCaffrey, D.F., Lockwood, J., Kortez, D., Louis, T.A., & Hamilton, l. (2003)، Ballou, D., Sanders, W., & Wright, P. (2003)).

يسمح نموذج الثبات(4.5)، كما وصفنا سابقا، لتأثيرات المعلم في تقديرات المستوى أن تتغير مع مرور الوقت، مع معالم تصف التغييرات على الآثار التي قدرت من البيانات. علاوة على ذلك، يؤدي نموذج الثبات إلى مكاسب تعتمد على تأثيرات المعلم للعام السابق، على عكس نموذج الطبقات أو النموذج المصنف طبقا لأكثر من سمة، ولفهم هذا الأمر، انظر في نموذج درجات طلاب السنة الثانية والثالثة. من المعادلة ٤,٥ التي تقدم لهم من خلال:

$$Y2 = M2 + T2 + a21T1 + e2 \quad (4.6)$$
$$Y3 = M3 + T3 + a32T2 + a31T1+ e3 \quad (4.7)$$

نحصل على المعادلة التالية لدرجات التحسن عند طرح المعادلة ٤,٦ من المعادلة ٤,٧:

$$Y3 - Y2 = (M3 – M2) + T3 + (a32 – 1) T2 + (a31 - a21) T1 + e3 - e2$$

ومن ثم، تعتمد المكاسب من الصف الثاني إلى الصف الثالث على مدرس الصف الثاني خلال الفصل المدرسي T2 (a32 – 1) ومعلم الصف الأول خلال الفصل المدرسي (a31 – a21) T1. تؤثر تأثيرات المعلم للصفين الأول والثاني على كل من درجات المستوى للصف الثالث والدرجات من الصف الثاني إلى الثالث، هذا مع نموذج الثبات لماكفري وآخرون (٢٠٠٣)، وبالتفصيل، عندما تكون a32 < 1، إذا كانت T2 موجبة إذن فإن مدرس الصف الثاني له مساهمة سلبية على الدرجات، والعكس إذا كانت T2 سالبة. يعد هذا شكل من الانحدار للمتوسط لأن تأثيرات المعلم تبتعد مع مرور الوقت وتنحدر درجات اختبار الطلاب تجاه الأداء الذي لم يعد

يتأثر بمعلم الصف الثاني، وعلى العكس، إذا كانت a32 > 1 ، إذا كانت T2 موجبة فإن لمعلم الصف الثاني تأثيرا إيجابيا على النمو خلال الصف الثالث – والعكس إذا كانت T2 سالبة. يزيد المعلمون الفعالون السابقون من سرعة عملية النمو ويقلل المدرسون غير الفاعلون منها. (McCaffrey, Lockwood, Kortez, Louis, & Hamilton, 2003, Deb Roy, & Bates, 2003).

تعليق عام

نماذج تهيئة المتغيرات المصاحبة، نظراً لقدرة هذه النماذج على تقسيم نمذجة الدرجات إلى أجزاء عديدة تعتمد على أداء الطلاب في الفصول، يتسم نموذج تهيئة المتغيرات بسهولة التحديد وملاءمة الاستخدام عن طريق أي حزمة برمجية إحصائية للنماذج الخطية المختلطة. ويتسم النموذج بقبول منطقي كبير بسبب إمكانية تفسيره كبداية جميع الطلاب في نفس المستوى الأساسي من الانجاز. إلا أن التقديرات سوف تكون حساسة للسنة المختارة كقيمة لبداية التقدير (Rivers, 1999).

ويكمن العيب الرئيس في نموذج تهيئة المتغيرات في أن ملائمة النماذج بصورة منفصلة لكل عام من البيانات تتجاهل معلومات مهمة عديدة. ويتجاهل النموذج المعلومات حول اداءات الطلاب في السنوات الأخرى والتي يمكن أن تفسر العوامل الطلابية الفردية وتقلل أخطاء بناء العينات وربما الانحراف في تقدير تأثير المعلمين خلال العام الحالي. كما يتجاهل النموذج حقيقة أن الدرجات الناتجة في المستقبل ترتبط بالمعلومات حول المعلمين السابقين. ويؤدي تجاهل هذه المعلومات إلى ما يطلق عليه الإحصائيون اسم "التقديرات غير الملائمة" والتي تشير إلى إمكانية أن تؤدي تقنية بديلة (مثل النمذجة المشتركة للعديد من سنوات جمع البيانات) إلى تقديرات ذات أخطاء أقل في بناء العينة. وسوف تؤدي عدم كفاءة التقديرات إلى أخطاء أكبر في التأثيرات التقديرية للمعلمين وتزيد من صعوبة الاستنتاجات حول المعلمين بالمقارنة مع التقديرات الفعالة الجيدة.

ولو كانت قيم الخطأ التراكمي في نموذج (4.1) مستقلة عبر السنوات المختلفة، لكانت التقديرات فعالة أيضاً. لكن ذلك يحدث فقط تحت ظروف تقييدية (وبالتالي غير واقعية) لاستمرار تأثيرات المعلمين بمرور السنوات والنمط الوظيفي للنمو في الإنجاز والتشابه بين درجات الطلاب.

كما يمكن أن يمثل خطأ القياس في درجات الاختبار مشكلة لتقدير تأثيرات المعلمين في النموذج متعدد المتغيرات. فعلى الرغم أن درجات الاختبار تقيس الإنجاز بصورة ضعيفة، إلا أن المحللين الذين يطبقون نموذج تهيئة المتغيرات في تقدير تأثيرات المعلمين سواء ضمنياً أو بصورة مباشرة **يفترضون صحة قياس الإنجاز خلال** السنوات السابقة. وعند تجاهل خطأ القياس في درجات السنوات السابقة، يحدث انحراف في تقديرات معاملات النموذج. تشير النتائج التحليلية إلى أنه لو تم توزيع الطلاب عشوائياً على الفصول وكانت توزيعات القيم الخالية من الأخطاء للطلاب في إنجاز

السنوات السابقة متشابهة عبر الفصول، فلن يؤدي أي خطأ في القياس إلى أخطاء نظامية في تأثيرات المعلمين. لكن يفترض المعلمون عادةً أن الطلاب لا يتم توزيعهم عشوائياً وأن توزيع الإنجاز السابق يتنوع عبر الفصول. وفي هذه الحالة، عندما تكون توزيعات القيم الخالية من الأخطاء للإنجاز السابق مختلفة عبر الفصول، يمكن أن يؤدي الخطأ في القياس إلى أخطاء نظامية في تأثيرات المعلمين.

لكن جميع النماذج البديلة تفترض كذلك وجود تأثيرات ثابتة للمعلمين عبر الطلاب، وبالتالي لا يتضح لنا إذا كانت أي من البدائل يمكنها تقديم تقديرات أفضل لمتوسط تأثيرات المعلمين إذا كانت التأثيرات غير ثابتة أو أن الطلاب يختلفون عبر الفصول.

ويتمثل العيب الثالث للنماذج تهيئة المتغيرات المصاحبة في استبعاد درجات الاختبارات التي يفقدها الطلاب خلال السنة السابقة أو الحالية. وتتسم هذه البيانات الجزئية هنا أيضاُ بعدم الكفاءة والفاعلية. فضلاً عن ذلك، إذا كان الطلاب الذين يفقدون الدرجات يختلفون بصورة نظامية عن الطلاب الآخرين، يمكن أن تنحرف التقديرات.

أما النماذج متعددة المتغيرات، فلها مميزات عديدة تسمح للمحللين باستكشاف مجموعة متنوعة من الافتراضات حول ثبات آثار المُعلم فضلا عن المتغير المتبقي لنتائج الطالب، ومن خلال تصميم النموذج الأخير بوضوح، تستخدم النماذج البيانات بفاعلية لاستغلال المتغير المتبقي أيضاً. تكون النماذج قوية بالنسبة للمتغيرات المحذوفة في بعض الظروف (McCaffrey, Lockwood, Kortez, Louis, & Hamilton, 2003, Deb Roy, & Bates, 2003).

ومن ثم يمكن تحديد فاعلية اداء المعلمين باستخدام النماذج السابقة من خلال مجموعة من النقاط المهمة:

- تفسير تأثير المعلم من خلال مستوى الطالب المبدئى داخل الفصل حيث يأتى كل طالب بفكرة ومعارف ومهارات مختلفة وعادات تعلم وممارسات مختلفة والتدريس الفعال سوف يساعد كل طالب للوصول إلى أقصى تقدم ممكن وفقا لإمكاناته وقدراته.
- ويعتبرمفهوم الاستعداد aptitude مهم جدا فالافراد يختلفون بشكل واسع فى الاستعداد والاستفادة من التدريس ومن ثم يجب أن نضع فى الاعتبار عند تقويم المعلمين استعداد الطلاب.
- وعندما نضع كل هذه المتغيرات فى الاعتبار عند تقويم المعلمين بين مجموعات الطلاب التى يدرس لها معلمون مختلفون فنحن أمام نظام عادل ومناسب للتقويم
- نماذج التأاثيرات المتعددة يضع فى اعتباره كل هذه المتغيرات من خلال معالجة المتغيرات القبلية للتحصيل من خلال دراسى خط الأساس base line ومن ثم يتم حساب تقدم مستوى الطلاب بالنسبة لأداءات الطلاب السابقة (معدل الكسب).

دراسات تناولت تقييم القيمة المضافة للمعلمين

يوضح مضمون الأبحاث الحديثة حول نماذج القيمة المضافة أن المعلمين يؤثرون بطرق مختلفة على تعلم الطلاب وتطور قدراتهم على الإنجاز. وتتسم التأثيرات التي أشارت إليها هذه الأبحاث بأنها كبيرة ومسئولة عن نسبة مرتفعة من التنوع في النمو. ويمكن تناول الأبحاث من خلال ثلاثة محاور هى:

١- أهمية تأثير المعلمين.

٢- تأثيرات المعلمين التراكمية والمستمرة.

٣- تنوع تأثير المعلمين بتنوع إنجاز الطلاب.

أ - أهمية تأثير المعلمين

دراسة رايت، هورن وساندرز (Wright, Horn, and Sanders (1997, 2006 :

سعى كل من رايت، هورن وساندرز إلى إظهار أن المعلمين يمثلون العامل الأهم تأثيراً على تعليم الطلاب. ويصيغ هؤلاء الباحثون في دراستهم ذات التصميم المنهجي والتي تشمل الإنجاز والتأثيرات العشوائية للمعلمين. ويقيس الباحثون إسهامات جميع المتغيرات في هذه النماذج باستخدام ما يسمونه "درجة المعيارية زد z-score". لا أن دراسة رايت، هورن وساندرز لم تقدم أي دليل على أن التأثيرات التقديرية للمعلمين وعناصر التنوع المرتبطة بها لا تتعرض للانحراف كنتيجة للعوامل التعليمية الأخرى للتعليم والتي لا يتناولها النموذج.

ويشير كل من رايت، هورن وساندرز مفهوم الدرجة المعيارية "زد" للإشارة إلى قياس حجم أثر المتغير المصاحب على الدرجات. وبالنسبة لحساب التأثيرات العشوائية للمعلمين والتفاعلات العشوائية بين المعلمين والانجاز، تعادل درجة المعيارية "زد" عنصر التنوع التقديري للأثر مقسوماً على الخطأ القياسي التقديري. وبالنسبة لمتغيرات التأثيرات الثابتة، يمكن الحصول على درجات المعيارية "زد" من خطوتي الحساب التاليتين:

أولاً: قيم p الناتجة من حاصل قسمة الإحصائيات F لاختبار التأثيرات الثابتة.

ثانياً: تعادل درجات المعيارية "زد" حساب التوزيع المعياري العادي الذي يتوافق مع أقل من نصف قيم P (١ مطروحاً منه قيم p مقسوماً على ٢). بمعنى أن قيم p الناتجة عن الاختبار f، يتم معاملتها كما لو كانت قيم p من اختبارين للتوزيع العادي.

كرر الباحثون تحليلهم ٣٠ مرة من خلال تلخيص أنماط النتائج عبر النماذج. وقد سجلت تأثيرات المعلمين أعلى درجة المعيارية "زد" في ٢٠ من بين إجمالي ٣٠ نموذج، وحقق عنصر الإنجاز أعلى درجة المعيارية "زد" في النماذج العشرة المتبقية. واستنتج من ذلك أن أهم العوامل المؤثرة على تعليم الطلاب هو المعلم" (Wright., horn, & Sanders,1997).

وعلى الرغم أن الباحثين قادوا عملية تطوير "نموذج التأثيرات المتعددة layered effects model" معقد والذي استخدمه نظام تينيسى TVAAS للتقييم القيمة المضافة في الحصول على

درجات الاختبارات فورياً من العديد من المشاركين وسنوات الاختبار المتعددة حول كل طالب (ساندرز، ساكستون وهورن، ١٩٩٧Sanders, Saxton, and Horn)، إلا أنهم لا يستخدمون ذلك النموذج في دراستهم. بدلاً من ذلك فقد استخدموا نموذجاً أبسط وأكثر تقليدية للتحسن الناتج عن التأثيرات الثابتة لسمات الطالب والفصل والتأثيرات العشوائية للمعلمين Wright, horn, & Sanders, (2006), Wright, horn, & Sanders, 1997).

دراسة روان، كورنتي وميلر (2000) Rowan, Correnti, and Miller (RCM):

استخدم كل من روان، كورنتي وميلر **أربعة** نماذج مختلفة لكل مادة ومجموعة وصف دراسي مع البيانات لاكتشاف مصادر التنوع في درجات الاختبار. يتكون النموذج **الأول** من تحليل تنوع ثلاثي المستويات مع الطلاب داخل الفصول أو المعلمين والفصول داخل المدارس والمدارس كمصادر التنوع الثلاثة؛

أما النموذج **الثاني** فهو عبارة عن نموذج للتشابه حيث تعادل درجة الطالب في الاختبار وظيفة خطية إضافية لدرجة السنة السابقة وتأثير المعلمين وتأثير المدرسة والمتغيرات المرتبطة بالطلاب والخطأ الناتج عنها.

$$Yiti = \alpha + \beta ytijt -1 + \gamma xiji + \eta it + \Theta I (ij) + eijt$$

- بحيث تكون (yiti) هي درجة الرياضيات أو القراءة الخاصة بالطالب j في المدرسة I لتطبيق الاختبار t،
- بينما يمثل الرمز (ŋit) تأثير المدرسة عند تطبيق ذلك الاختبار،
- وتكون ((Θ i(ij) هي تأثير معلم الطالب،
- في حين ترمز (xiji) إلى مجموعة من خصائص الطالب والتي ربما يتنوع البعض منها بمرور الوقت
- كما يعبر الرمز (eijt) عن الخطأ الناتج.
- ويتم تضمين تأثيرات المدرسة والمعلمين كتأثيرات عشوائية؛

أما النموذج **الثالث** فهو نموذج لاكتساب الدرجات حيث يشبه تحسن درجات الطلاب خلال عام واحد الوظيفة الخطية لدرجة السنة السابقة وتأثير المعلم وتأثير المدرسة والمتغيرات المرتبطة بالطلاب والأخطاء الناتجة. وكما ورد في دراسة مكارثي وآخرين (٢٠٠٣) .McCaffrey et al، يعتبر هذا النموذج مكافئ جبرياً للنموذج الثاني الذي وضعه كل من روان، كورنتي وميلر (RCM) نظراً لأن نموذج التحسن في الدرجات يشمل درجات العام السابق كمتغير إضافي في النموذج، وبالتالي لا نناقش هذا النموذج في استعراضنا؛

أما النموذج **الأخير** الذي استخدمه روان، كورنتي وميلر (RCM) فهو عبارة عن نموذج متعدد التصنيفات (رودنبوش وبريك، ٢٠٠٢ Raudenbush and Bryk). وفي هذا النموذج، يتم عرض

درجة الطالب j في المدرسة I في الوقت t، yijt على النحو التالي:

$$yijt = \alpha + \beta t + \xi t2 + \alpha i + \beta it + \alpha ij + \beta ijt + \gamma xijt + \Theta\ 1(ij) + ... + \Theta\ t(ij) + eijt$$

- حيث ترمز كل من (αi و βi) إلى الخط المائل والمنحدر الدال على المدرسة،
- أما (αij و βij) فترمز إلى الخط العشوائي المائل والمنحدر للطالب،
- كما يدل (xijt) على مجموعة من خصائص الطلاب والتي يمكن أن تتنوع البعض منها بمرور الوقت،
- وتشير (Θ 1(ij) حتى Θ t(ij)) إلى تأثيرات معلمي الطلاب خلال وقت الاختبار ١ إلى t،
- و (eijt) فهي الاختصار الدال على الخطأ الناتج.

وبالتالي، تعد الدرجات بمثابة نماذج ذات وظيفية زمنية عامة (ξt2 + βt+α) بالإضافة إلى الآفاق الزمنية الخطية العشوائية المرتبطة وغير المرتبطة بالمدرسة. ويفترض النموذج عدم وجود تنوع في العنصر اللاخطي للنموذج. ويتم ضمنياً تجسيد أي تنوع في β في قيمة الخطأ الناتج.

يتم إضافة تأثير المعلم Θ t(ij) لكل عام وتظل هذه التأثيرات ظاهرة في النموذج خلال إجراء الاختبارات في المستقبل، وهو السبب الذي يجعل نموذج الدرجة خلال الزمن (ت) يشمل الفصول الدراسية السابقة التي كان المعلمون يعملون فيها مع الطلاب. وعلى الرغم أننا يمكن أن نشير إلى مصطلح Θ t(ij)، يقر روان، كورنتي وميلر بأنها عبارة عن تأثيرات الفصل الناتجة. وكما يتضح في دراسة ماكفري وآخرين (٢٠٠٣)، يشير النموذج السابق متعدد التصنيفات ما يلي:

$$yijt - yijt\ -1 = \beta - \xi + 2\xi t + \beta i + \beta ij +$$

$$\gamma' (xijt - xijt-1) + \Theta\ t(ij) + eijt - eijt - 1$$

وبالتالي فإن التأثير الوحيد للمعلم على اكتساب الدرجات فهو من المعلم الحالي، أما المعلمين الآخرين فلا يسهمون في تلك العملية.

ومن خلال نموذج تحليل التباين ANOVA، وجد نموذج روان، كورنتي وميلر (RCM) أن الفصول (بما فيها من معلمين) مسئولة عن حوالي ما بين ١٢-٢٣% من إجمالي التنوع في انجاز القراءة وما بين ١٨-٢٨% من إجمالي التنوع في إنجاز الرياضيات. أما في نموذج التوافق التباييني، يكون المعلمون (الفصول) مسئولون عن ما بين ٤-١٦% من التباين التوافقي في درجات القراءة وما بين ٨-١٨% من التباين التوافقي في درجات الرياضيات تبعاً للمجموعة والسنة الدراسية. يعتبر التباين التوافقي بمثابة التباين الناتج بعد إجراء التوافق بين خصائص الطلاب والدرجات السابقة.

وبالنسبة لنموذج لتصنيفات المتعددة cross – classified model، وجد روان، كورنتي وميلر

أن التنوع في تأثيرات المعلمين مسئولة عن ما بين ٦٠-٦١% من "التنوع الموثوق" في نمو القراءة ونسبة تتراوح بين ٥٢-٧٢% للتنوع الموثوق في النمو في الرياضيات. يُعرف روان، كورنتي وميلر "التنوع الموثوق" في النمو على أنه نسبة التنوع في المنحنيات العشوائية المدرسية مضافاً إليها التنوع في المنحنيات الفردية العشوائية والتنوع في تأثيرات المعلمين.ويتم استثناء التنوع في الانحرافات حول الاتجاهات الخطية من تحليل التنوع الموثوق.

في المجمل العام، نستنتج أن روان، كورنتي وميلر يقدمون دليل مقنع على تأثيرات المعلمين على الرغم من عدم التحديد الدقيق لحجم هذا التأثير. ويجب تفسير نتائجهم بحرص نظراً لارتباط أحجام التأثير بجزء صغير فقط من التنوع في الدرجات (أو النمو) بدلاً من إجمالي التنوع (Rowan, Correnti, & Miller,2000).

دراسة روفكين، هانوشيك وكين (2000) Rivkin, Hanushek, and Kain (RHK) :

يستفيد روفكين، هانوشيك وكين من المجموعات الطلابية المتعددة ذوي السنوات الثلاث من درجات الاختبار في التخلص التام من تأثيرات العوامل الأخرى بخلاف المعلمين على إنجاز الطلاب. ويجد المؤلفون أن هناك ثمة تأثيرات للمعلمين ويقدرون أن المعلمين مسئولين عن نسبة ٣,٢% على الأقل من التنوع في إنجاز الطلاب. ويعني ذلك أن زيادة وحدة انحراف معياري واحدة إلى تأثير المعلمين ترتبط بحوالي زيادة قدرها ٠,١٨ من زيادة وحدة الانحراف المعياري في الدرجات. واقتصر الباحثين في تحليلاتهم على الطلاب الذين يستمرون في نفس المدرسة ويكملون الاختبارات على مدى ثلاث سنوات متتالية. وبالتالي، يقترح المؤلفون في نتائجهم أن المعلمين يمكنهم الاهتمام ببعض الطلاب في قياسات معينة، لكن يستحيل إجراء التفسير العام للنتائج.

ومن بين القيود الرئيسية التي واجهت دراسة رايت، هورن وساندرز- وإلى حد أقل دراسة روان، كورنتي وميلر- احتمالية أن الآثار التراكمية للطلاب أو المدرسة لا تحظى بالاهتمام من جانب متغيرات النموذج، وبالتالي تربك تقديرات تأثيرات المعلمين. تتناول دراسة روفكين، هانوشيك وكين هذه المشكلة بصورة مباشرة وتحاول تقدير التأثير السببي الحقيقي للمعلمين والذي يختلف عن أي مصادر أخرى للتنوع. استخدمت دراسة روفكين، هانوشيك وكين البيانات حول ٥٠٠٠٠٠ طالب من تكساس من خلال ثلاث مجموعات متعاقبة من ٢١٥٦ للمدارس التمهيدية. وقاموا في الدراسة بالربط بين درجات اختبار الرياضيات لكل طالب خلال ثلاث سنوات تتوافق مع الصفوف الدراسية الرابع والخامس والسادس بالنسبة لمجموعتين والصفوف الدراسية الثالث والرابع والخامس للمجموعة الثالثة. يرتبط الطلاب بالمدرسة من خلال السنة الدراسية وليس الفصل. ويتم الحصول على درجات الاختبار عن طريق تطبيق اختبار تكساس لتقييم المهارات الأكاديمية (TAAS) Texas Assessment of Academic Skills، وبالتالي يتم قياس الدرجات خلال السنوات الدراسية على مقياس تنموي واحد. وبدلاً من ذلك، تم تقنين الدرجات بحيث يكون متوسط الدرجات القياسية لكل صف دراسي تعادل

صفر بينما يعادل الانحراف المعياري للدرجات القياسية قيمة واحد صحيح (.Rivikin.,2007)

وبالتالي، يمكننا استنتاج أن هذه الدراسة قد أظهرت تأثيرات المعلمين في تحسن درجات الطلاب القياسية الذين يظلون في نفس المدرسة لمدة ثلاث سنوات متتالية، لكن لا يمكننا افتراض ربط هذا التحسن بالتحسن على أحد الاختبارات التنموية (Rivikin, Hanushek, & kain 2000).

الخلاصة:

وجدت الدراسات الثلاث أن تأثيرات المعلمين على الطلاب تحمل أهمية كبيرة. ويصعب تقييم حجم التأثيرات والأهمية النسبية للمعلمين بالمقارنة مع العوامل الأخرى المؤثرة على التعليم، إلا أن هذه الدراسات تشير إلى إمكانية تفسيرها في ضوء توضيح وجود تأثير للمعلمين.

ب- تأثيرات المعلمين التراكمية طويلة الأجل

- دراسة ساندرز وريفرز (1996) Sanders and Rivers:

يستخدم ساندرز وريفرز في الدراسة مفهوم ثنائي المراحل. أولاً، يقدرون تأثير المعلمين باستخدام صورة من نموذج تعديل المتغيرات المصاحبة covariate في الفصل التالي. ويفترض النموذج أن درجات الطلاب في العام الحالي تعادل وظيفة خطية إضافية لدرجات الطلاب في العام السابق وتأثير المعلمين والخطأ التراكمي:

$$yit = \alpha t + \beta tyit\text{-}1 + \Theta\, t(i) + eit$$

حيث تمثل (yit) درجة الطالب في الرياضيات خلال الصفوف الدراسية t = 3,4,5، بينما تمثل (Θ t(i)) أثر المعلمين على الطلاب في الصف الدراسي t، بينما يرمز إلى نسبة الخطأ التراكمي. ويفترض أن تكون تأثيرات المعلمين عشوائية. وتلائم النماذج المنفصلة درجات الرياضيات في الصفوف الدراسية الثالثة والرابعة والخامسة بحيث يتم السماح لمعاملات النموذج بالتغير مع تغير الصف الدراسي. كما يتم تجاهل أي علاقة ارتباط ممكنة بين الأخطاء التراكمية من نفس الطالب. ويتم تصنيف تقديرات تأثير المعلمين لكل صف دراسي إلى واحد من خمسة تصنيفات لتقدير درجات تأثير كل معلم على مقياس مقسم من ١ إلى ٥، مع تخصيص رقم ١ إلى الخمس الأدنى (الأقل تأثير) ورقم ٥ للمعلمين في الخمس الأعلى (الأكثر تأثير). وخلال المرحلة الثانية، يتم تعديل درجات الطلاب من الصف الدراسي الخامس كوظيفة خطية إضافية لتأثير المعلمين (حيث يتم معاملة التصنيفات الخمسة كمتغيرات تصنيفية) للصفوف الدراسية من الثالثة إلى الخامسة، ودرجات طلاب الصف الثاني والخطأ التراكمي:

$$Yi5 = \mu 5 + \beta 5{,}2yi2 + \Sigma 5\ \Sigma 5\ \gamma gkTigk + \varepsilon it$$

$$g = 3k = 1$$

بحيث تكون k=1-5 تعادل واحد لو كانت نسبة i للمعلمين في الصف الدراسي g ذات درجة

تأثير لـ k وصفر في نسب أخرى.

أما (Ygk) فتشير إلى تأثيرات المعلمين عند درجة تأثير معينة. وعندما تكون بعض نسب (y2k وy3k) غير صفرية، تستمر تأثيرات المعلمين خلال اختبارات السنوات التالية. ويتم استخدام بيانات نفس الطلاب في المرحلتين الأولى والثانية. (كوبرمينتز، ٢٠٠٢ Kupermintz). (Sanders & Rivers, 1996).

وفي الدراسة مشابهة باختيار منهج لتحسين درجات الطلاب وتطبيق منهج ساندرز وريفرز على البيانات المتشابهة. وأجريت مقارنة بين تقديرات ساندرز وريفرز لتأثير ثلاثة من أكثر المعلمين تأثير (Tg5 = ١ عندما تكون قيمة الصف الدراسي g= ٣، ٤، و٥)، على التأثير الحقيقي كما يحدده نموذج جمع البيانات. احتوى نموذج جمع البيانات على معايير لتحديد كل من: (أ) تأثيرات المعلمين على الطلاب، (ب) استمرار تأثيرات المعلمين على الدرجات المستقبلية، و(ج) تفاوت النسب المتوسطة لتأثيرات ومتغيرات الطلاب في الفصول أو المتغيرات الأخرى بخلاف تأثيرات المعلمين والتي نتج عنها تحسن الدرجات.

نستنتج من ذلك أن ساندرز وريفرز يقدمان الدليل على وجود واستمرار تأثيرات المعلمين أو الفصول بالنسبة للطلاب الذين يجتازون الاختبارات، إلا أنه غالباً ما يحدث مبالغة إلى حد ما في تقدير حجم التأثيرات. .(McCaffrey, Lockwood, Kortez, Louis, & Hamilton, 2003)

- دراسة نموذج ريفرز Rivers (1999) :

نظراً لحجم تأثيرات ساندرز وريفرز وفروض هذه النتيجة والتناقض مع المنهج المستخدم، يحاول المؤلفون الآخرون تكرار نفس النتيجة لكن بتعديل بسيط. تكرر ريفرز (١٩٩٩) نفس التصميم مع إجراء عدد من التغييرات المهمة للتغلب على بعض الانتقادات الموجهة لدراسة ساندرز وريفرز وقد وجدوا استمراراً في تأثيرات المعلمين. ونظراً لوجود ثمة علاقة وهمية ممكنة بين الفصول ونتائج الطلاب ونظراً للخطأ في قياس متغير درجات الاختبارات الأربعة في نهاية كل عام دراسي، ربما يكون هناك انحرافاً ايجابياً في دراسة ريفرز نتوقع لها أن تكون ما بين صغيرة إلى متوسطة على أقصى تقدير.

وبدلاً من استخدام نموذج التوافق البسيط الذي اتبعته دراسة ساندرز وريفرز لتقدير تأثير المعلمين، تستخدم ريفرز تقديرات تأثيرات المعلمين من نموذج نظام تينيس لتقييم القيمة المضافة. ويعد نظام تينيسى TVAAS لتقييم القيمة المضافة نموذجاً معقداً للتوزيع المشترك لبيانات درجات الاختبارات المطولة للطلاب. وتوزع هذه النماذج الدرجات لعدد حتى خمس مواد دراسية (الرياضيات

والقراءة واللغة والعلوم والدراسات الاجتماعية) حتى إجمالي اختبارات ست سنوات. وعلى الرغم أن النموذج لا يشمل أي من خصائص الطلاب أو الفصول أو المدرسة، فأنه يسمح بالارتباط بين درجات نفس الطالب وهو ما يعمل على التعديل الجزئي لخصائص الطلاب المحذوفة. يشمل نموذج نظام تينيس لتقييم القيمة المضافة معيار منفصل لمتوسط كل نظام مدرسي أو إدارة مدرسية، وبالتالي ترتبط التأثيرات التقديرية للمعلمين بالمعلمين الآخرين في نفس المنطقة التعليمية.

يتمثل **الفرق الثاني** بين دراسة ريفرز وتلك التي أجراها ساندرز وريفرز في أن ريفرز استخدمت مجموعتين من الطلاب بدلاً من مجموعة واحدة لتقدير استمرار تأثيرات المعلمين. قدمت المجموعة الأولى تقديرات لتأثير المعلمين من نموذج نظام تينيس لتقييم القيمة المضافة. أما المجموعة الثانية المنفصلة من الطلاب فقد قدمت تقديرات حول تأثير المعلمين. وبالتالي، تم تقدير قيم تأثير المعلمين وتأثير تنوع التأثير من مجموعتين منفصلتين من الطلاب.

يكمن **الفارق الرئيس** الأخير بين ريفرز ودراسة ساندرز وريفرز في استخدام ريفرز للنتائج حول اختبار مختلف عن الاختبار المستخدم في تقدير التأثير ، ويتم قياس النتيجة خلال نهاية العام الدراسي التاسع بينما يتم قياس تأثيرات المعلمين خلال السنوات الدراسية السادسة والسابعة والثامنة.

وجدت دراسة ريفرز أن تأثيرات المعلمين من جميع السنوات الدراسية الأربع ترتبط بدلالة إحصائية بالدرجات التي يحصل عليها الطلاب في نهاية الصف التاسع. ويتناقص تأثير معلمي الصفوف الدراسية الخامسة والسادسة بالمقارنة مع درجات الطلاب في الصف الرابع. بمعنى أن معلمي الصفوف الدراسية الخامس والسادس كانوا أكثر تأثيراً بالنسبة للطلاب ذوي الدرجات الأقل خلال الصف الرابع. ويبلغ حجم تأثير معلمي الصف السادس حوالي ٢,٥ مرة أكبر بالنسبة للطلاب في الربع الأدنى بالمقارنة مع الربع الأعلى في اختبار الصف الرابع (Rivers, 1999).

وبالتالي، تقدم دراسة ريفرز بصورة متوازنة الدليل على أن المعلمين يؤثرون على إنجاز الطلاب في الاختبارات القياسية المطبقة بعد عدة سنوات في المستقبل بالنسبة للطلاب الذين يظلون في نفس النظام التعليمي لمدة ٦ سنوات (Rivers, 1999).

- **دراسة ميندرو، وجوردن، وجوميز، وآندرسون وبيمبري Mendro, Jordan, Gomez, Anderson, and Bembry (MJGAB) (1998):**

وجدت دراسة ميندرو، جوردن، جوميز، أندرسون وبيمبري على عكس نماذج ساندرز وريفرز وريفرز تأثيرات كبيرة ومستمرة للمعلمين عبر المجموعات المتعددة وفي كل من درجات القراءة والرياضيات. فقد قاموا بإجراء ٤ تحليلات منفصلة على هذه المجموعات. يتناول التحليل الأول درجات الرياضيات لجميع الطلاب الذين اجتازوا اختبار الرياضيات خلال ٤ سنوات متتالية. أما

التحليل الثاني فيتناول درجات القراءة لجميع الطلاب على مدى ٤ سنوات مكتملة تم فيها اجتياز اختبار القراءة. وتكرر التحليلات الثالثة والربعة هذه التحليلات وتحدد العينة بجميع الطلاب الذين لديهم درجات اختبار على مدى خمس سنوات متعاقبة في القراءة أو الرياضيات. ويتم استبعاد طلاب الصف الرابع من التحليلين الأخيرين. وتستخدم جميع التحليلات اختبار (إيوا ITBS) للمهارات الأساسية كمقياس للانجاز.

واستخدم الباحثون نظام دالاس للمحاسبة (DVAAS) لتقدير تأثير المعلمين من خلال ثلاث مراحل، حيث تقوم **المرحلة الأولى** بحذف تأثيرات ما يُعرف باسم "متغيرات الإنصاف" من درجات العام الحالي والسابق ونسب الحضور والغياب. تشمل متغيرات الإنصاف كل من الأصل وإتقان اللغة (الإتقان المحدود للغة الانجليزية والأصول الزنجية والأسبانية وغيرها)، والنوع والتفاعلات ذات الترتيب الأول والثاني لهذه المتغيرات. كما تشمل متغيرات الإنصاف كذلك مقاييس مستوى التوافق مع الدخل ومعدلات الفقر وحضور الجامعة (يتم تحديدها عبر مستوى التوافق للطلاب). وبالنسبة لدرجات القراءة والرياضيات خلال العامين الحالي والسابق.

وشملت **المرحلة الثانية** متغيرات المدرسة المتضمنة في النموذج كل من التنقل بين المدارس وتكدس الفصول ونسب الطلاب الأقليات ونسب الطلاب السود ونسب الطلاب الأسبان ومتوسطات مستوى التوافق عبر متغيرات الإنصاف. كما يشتمل النموذج على تأثيرات عشوائية للمدرسة وتأثيرات عشوائية على المؤشر الذي تغير بتغير المدارس (نتائج المرحلة الأولى لنتائج السنة السابقة ونسبة الحضور للمدرسة). تسمى الآثار الناتجة للدرجات الفردية للطلاب من نماذج المرحلة السابقة "النسب الباقية للمرحلة الثانية" ويتم إدخارها للمرحلة النهائية للتقدير.

وتقدر **المرحلة الثالثة** تأثيرات المعلمين كمتوسطات للفصول خلال النسب المتبقية للمرحلة الثانية. ويترتب عن الإجراء تقديرات منفصلة لتأثيرات المعلمين على درجات اختبارات ITBS للقراءة والرياضيات. ويتم عرض التفاصيل حول تأثيرات معلمي دالاس في دراسة ويبستر وميندرو (١٩٩٧).

يستخدم نموذج ميندرو، وجوردن، وجوميز، وآندرسون وبيمبري (MJGAB) التقديرات النهائية للمرحلة الثالثة كمقاييس لتأثير المعلمين. وبالنسبة لتحليلات بيانات السنوات الأربع، يحصل المعلمون على تقييمات تتراوح من ١–٥ تتوافق مع التأثيرات التقديرية للمرحلة الثالثة، كما تشتمل النماذج على درجات الاختبار القبلي كمتغير، وتشمل تقييمات المعلمين لكل من السنوات الثلاثة أو الأربعة وفقاً للتحليل. ويتم تضمين تقييمات المعلمين في النموذج كسلسة مكونة من ٤ متغيرات أو متغيرين للمؤشرات وفقاً لنسب التقييم الثلاث أو الخمس. ولا تذكر دراسة ميندرو، جوردن، جوميز، آندرسون وبيمبري النماذج بل تشير إلى الدلالة العالية لتأثيرات المعلمين. (Mendro, Jordan, Gomez, Anderson, & Bembry 1998).

خلاصة:

يشير ثبات النتائج عبر دراسات ساندرز وريفرز، ريفرز وميندرو، جوردن، جوميز، آندرسون وبيمبري وكين إلى أن تأثيرات المعلمين من السنوات السابقة ترتبط بالدرجات المستقبلية للاختبار – حتى ولو بعد مرور عدد من السنوات. وربما يكون حجم التأثيرات مبالغاً فيه في دراسة ساندرز وريفرز، إلا أن نتائج ساندرز وريفرز غير واردة عندما يكون حجم التأثيرات الحقيقية هو صفر. بينما يؤكد ريفرز على وجود تأثيرات دائمة للمعلمين. كما تساند نتائج مينـدرو، جوردن، جوميز، آندرسون وبيمبري وكين التي تتحكم في كثير من متغيرات الطلاب لتفسير نتائج تأثير المعلمين.

جـ- اختلاف تأثيرات المعلمين تبعاً لمستوى تحصيل الطلاب :

يستكشف ساندرز وريفرز ليس فقط التأثيرات التراكمية للمعلمين لكن أيضاً مدى التنوع في تأثير المعلمين بتنوع مستوى إنجاز الطلاب. وأظهرت نتائجهم أن الطلاب ذوي المستوى الأدنى للإنجاز هم أول المستفيدين من المعلمين الأكثر تأثيراً.

وجد الباحثون أنه في داخل كل مستوى لإنجاز الطلاب تظهر زيادة في متوسط أداء الطلاب كنتيجة لتأثير فاعلية المعلمين. كما ترتبط الدرجات سلبياً بالإنجاز – الطلاب ذوي المستويات الأدنى من الإنجاز يحققون أعلى مستويات الكسب. ويفسر الباحثون هذه النتيجة على أنها توضح أن الطلاب ذوي المستويات الأدنى من الانجاز هم أول المستفيدين من المعلمين الأعلى تأثيراً وكفاءة (McCaffrey, Lockwood,Kortez, Louis, & Hamilton, 2003).

جدول (١) ملخص أبحاث تأثيرات المعلمين المتراكمة على نتائج الطلاب

أ – أهمية تأثير المعلمين	
الدراسات المستعرضة: رايت، هورن وساندرز (١٩٩٧) روان، كورينتي وميلر (٢٠٠٢) ريفكين، هانوشيك، وكين (٢٠٠٠)	تسعى جميع هذه الدراسات باستخدام مناهج متنوعة إلى توضيح وجود تأثيرات للمعلمين. ويصعب تقييم حجم هذه التأثيرات وأهمية المعلمين بالمقارنة مع العوامل الأخرى المؤثرة في التعليم، إلا أن استعراضنا يشير إلى إمكانية تفسير هذه البحوث في ضوء إظهار وجود تأثير للمعلم.
ب – تأثيرات المعلمين كتأثيرات تراكمية وطويلة الأجل	
الدراسات التي تم استعراضها: ساندرز وريفيرز (١٩٩٦) ريفرز (١٩٩٩) مينـدرو وآخرون (١٩٩٨) كين (١٩٩٨)	تهدف هذه الدراسات إلى توضيح هيمنة تأثيرات المعلمين السابقين على الإنجاز المستقبلي للطلاب. يشير ثبات النتائج عبر هذه الدراسات أنها توضح استمرارية تأثيرات المعلمين. ولكن ربما يكون هناك إفراط في توضيح حجم هذه التأثيرات يحتاج إلى مزيد من الدراسات
جـ- اختلاف تأثيرات المعلمين وفقاً لمستوى انجاز الطالب	
الدراسات التي تم استعراضها: ساندرز وريفرز (١٩٩٦)	تهدف هذه الدراسة إلى توضيح أنه كلما انخفض مستوى إنجاز الطلاب، تزداد احتمالات أن يكونوا أول المستفيدين من المعلمين الأكثر تأثير.

إجراءات الدراسة :

عينة الدراسة :

تتكون عينة الدراسة من مجموعتين: الاولى، مجموعة تقنين الادوات، والثانية عينة الدراسة الاساسية.

جدول (٢)

عينة تقنين الأدوات

المدرسة	الصف الدراسى	عدد الطلاب	متوسط العمر
مدرسة محمد فريد للتعليم الاساسى– إدارة منشأة ناصر	الاول الابتدائى	٩٠	٦,٦
	الثانى الابتدائى	٩٦	٧,٧
	الثالث الابتدائى	٩٦	٨,٦
مدرسة السيدة خديجة للتعليم الاساسى – إدارة المعادى	الاول الابتدائى	٩٢	٦,٥
	الثانى الابتدائى	٨٤	٧,٤
	الثالث الابتدائى	٨٢	٨,٣

جدول (٣)

عينة الدراسة الاساسية

سنوات الدراسة الطولية	مدرسة محمد فريد للتعليم الاساسى– إدارة منشأة ناصر	مدرسة السيدة خديجة للتعليم الاساسى – إدارة المعادى	مدرسة محمد فريد للتعليم الاساسى– إدارة منشأة ناصر	مدرسة السيدة خديجة للتعليم الاساسى – إدارة المعادى	مدرسة محمد فريد للتعليم الاساسى– إدارة منشأة ناصر	مدرسة السيدة خديجة للتعليم الاساسى – إدارة المعادى	اجمالى عينة الطلاب فى الدراسة الأساسية	متوسط العمر	عينة المعلمين
العام الدراسى ٢٠١٢/٢٠١٣ الصف الاول الابتدائى	٣٢٩	٣٢٢					٦٥١	٦,٥	١٤
العام الدراسى ٢٠١٣/٢٠١٤ الصف الثانى الابتدائى			٣٢٧	٣٢٠			٦٤٧	٧,٦	١٣
العام الدراسى ٢٠١٤/٢٠١٥ الصف الثالث الابتدائى					٣٢٧	٣٢٠	٦٤٧	٨,٦	١٣

تم اختيار العينة من معلمى الصفوف الثلاثة الاولى من التعليم الابتدائى (معلمى الفصل) حيث يقوم كل معلم بتدريس كل من اللغة العربية والرياضيات، كما يمكن تتبع اداءهم مع طلابهم على مدار ثلاث سنوات حيث بلغت العينة ١٣ معلماً بنهاية الدراسة بعد انتقال احد المعلمين إلى مدرسة اخرى بعد العام الاول من التطبيق.

وقد تكونت عينة الدراسة من طلاب الصف الأول والثانى والثالث الابتدائى فى كل من الدراسة الأساسية وعينة تقنين أدوات الدراسة.

وقد تم تطبيق مقياس المستوى الاجتماعى الاقتصادى ومقياس الاستعداد اللغوى والرياضى فى صورتهم النهائية فى النصف الاول من العام الدراسى نفسه. بينما طبقت اختبارات نواتج التعلم فى الرياضيات واللغة العربية عبر الدراسة الطولية مع نهاية النصف الدراسى الاول ونهاية نصف الدراسى الثانى على مدار ثلاثة اعوام دراسية هى: ٢٠١٢/٢٠١٣ ، و٢٠١٣/٢٠١٤، و٢٠١٤/٢٠١٥.

أدوات الدراسة :

تتكون أدوات الدراسة من مقياس المستوى الاجتماعى الاقتصادى، واختبارات الاستعداد للدراسة فى الجانب اللغوى والجانب الرياضى، واختبارات تحصيلية محكية المرجع فى مادتى الرياضيات واللغة العربية.

١- مقياس المستوى الاجتماعى الاقتصادى

تم استخدام مقياس المستوي الاقتصادي الاجتماعي كأحد أهم العوامل المصاحبة عند استخدام نموذج التأثيرات المتعددة (كأحد نماذج المتغيرات المتعددة للقيمة المضافة).

ويتكون المقياس من من خمسة ابعاد هى: الوسط الاجتماعى، والمستوى التعليمى للوالدين، ومستوى المعيشة، والجو الاسرى. وبلغ ثبات الاستمارة بطريقة اعادة الاختبار معامل ارتباط قدره ٠,٧٥٩ وبطرقة ثبات الفا ٠,٨ وهو معامل ثبات مرتفع نسبياً.

٢- مقياس الاستعداد اللغوى والرياضى اعداد الباحث

أ – إعداد عبارات الاختبار :

قام الباحث بإعداد مقياس الاستعداد لقياس الاستعداد الدراسى للمرحلة الابتدائية كأحد المتغيرات المصاحبة المهمة عند استخدام نماذج القيمة المضافة لتقييم فاعلية اداء المعلمين وذلك كالتالى ويتكون اختبار الاستعداد الرياضى من ٣٠ مفردة فى صورته الاوليه ويتضمن ثلاثة مكونات هى:

- يميز بين الأشكال المتشابهة والمختلفة.
- التعرف على الاعدادو العمليات الحسابية البسيطة .
- يقارن خواص الأجسام باستخدام لغة الرياضيات، مثل (الطول- الحجم- السعة- الوزن- المساحة- الزمن-.....).

الاستعداد اللغوى و ويتكون من ٣٠ مفردة فى صورته الاوليه ويتضمن ثلاثة مكونات هى:

- ينطق الحروف الهجائية من مخارجها الصحيحة
- يميز أوجه التشابه والاختلاف في رسم الحروف الهجائية.
- يربط بين الصورة والكلمات الدالة عليها.

ب- تحكيم المقياس

- تم عرض مكونات كل من الاستعداد اللغوى والرياضى على مجموعة من المحكمين في القياس والتقويم التربوى وبعض معلمى رياض الاطفال ومعلمى الفصول بالمرحلة الابتدائية، وذلك للحكم على صياغة أسئلة من حيث البساطة فى الصياغة ومناسبة الصور المستخدمة فى المقياس بما يناسب المرحلة العمرية الصغيرة. وكذلك الحكم على ارتباط المفردة بالبعد الفرعى. وتم تعديل الأسئلة وفق آراء السادة المحكمين.

ج- الدراسة الاستطلاعية :

تم تطبيق المقياس على عينة صغيرة من الطلاب مكونة من (١٢) طالباً بالصف الاول الابتدائى بهدف معرفة العبارات التى يجد الطلاب صعوبة فى فهمها أو أية صعوبات أخرى، وقد كان من نتيجة الدراسة الاستطلاعية تعديل صياغة (٤) عبارات وتوضسح بعض الصور فى بعد الاستعداد اللغوى وتعديل (٣) عبارات فى بعد الاستعداد الرياضى.

د-صدق الملاءمة (كما توفره نظرية الاستجابة للمفردة)

إن تدرج مفردات مجموعة من الاختبارات – يفترض أنها تقيس نفس السمة – على ميزان تدرج واحد مشترك باستخدام نظرية الاستجابة للمفردة يعني أن هذه المفردات تحقق شرط أحادية البعد أي أنها تعرف متغيراً واحداً. وتتحقق أحادية البعد بتحقق صدق المحتوى وبملاءمة كل من الأفراد والمفردات للنموذج (لأسس القياس الموضوعي) (أمينة كاظم،١٩٨٨، ١٩٩٥). وقد تم حذف الأفراد غير الصادقين والمفردات غير الملائمة آلياً. وذلك كالتالى:

١- تم استخدام مجموعة من البرامج الإحصائية في تحليل وتدريج استجابات مجموعات الدراسة على الاختبار وهى الحزمة الإحصائية SPSS وبرنامج ونستبس WINSTEPS.

٢- التدريج المبدئي على ميزان تدريج واحد لكل من مقياس الاستعداد ببعديه اللغوى والرياضى كل منهما على حدة.

٣- حذف الافراد غير الملائمين الذين تجاوزا محكات الملاءمة الاحصائية، حيث ا تجاوز الحد الفاصل (+٢) يعنى ان نمط استجابات الفرد يختلف عما يتوقع منه. وتجاوز الحد الفاصل (-٢) فيعنى ان نمط استجابات الفرد متسق بدرجة غير واقعية.

٤- مراجعة إحصاءات الملاءمة وحذف المفردات غير الملائمة والتي كانت قيمة كا٢ الخاصة بها دالة إحصائياً.

٥- التدريج النهائي لكل مقياس بعد حذف الأفراد غير الصادقين والمفردات غير الملائمة.

٦- تقدير بارمترات صعوبة وتمييز المفردات والخطأ المعياري لتقدير كل منها.

وفيما يلي عرض للمؤشرات الإحصائية لمفردات لمقياس الاستعداد اللغوى والرياضى في صورتها النهائية، وذلك بعد حذف الأفراد والمفردات غير الملائمة.

وقد تراوحت الصعوبات لبعد الاستعداد اللغوى مابين (-١،٩٥) وحتى (٤،٦٧) بمتوسط قدره (٠،٧٦) لوجيت والذي يشير إلى المدى الواسع لصعوبات المفردات، وتراوح الخطأ المعياري لصعوبات المفردات ما بين (٠،٠٤) و(٠،١٩) الذي يشير إلى ارتفاع قيم دالة المعلومات للمفردات.

بينما تراوحت الصعوبات لبعد الاستعداد الرياضى مابين (-١،٩٥) وحتى (٣،١٦) بمتوسط قدره (٠،٧٧) لوجيت والذي يشير إلى المدى الواسع لصعوبات المفردات، وتراوح الخطأ المعيارى لصعوبات المفردات ما بين (٠،٠٥) و(٠،٣٥) الذي يشير إلى ارتفاع قيم دالة المعلومات للمفردات.

هـ - البنية العاملية لمقياس الاستعداد للدراسة :

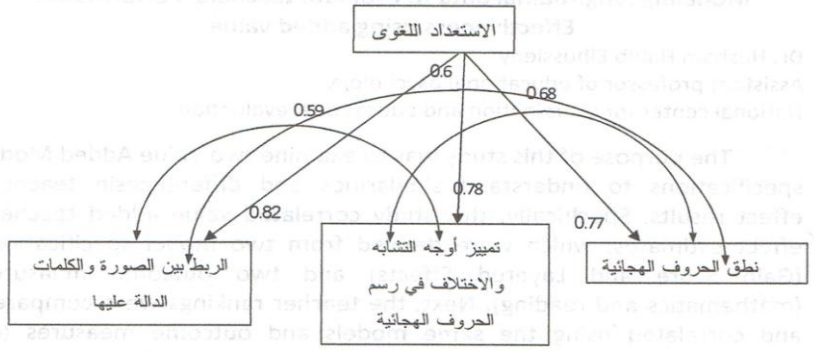
استخدم التحليل العاملى التوكيدى للتحقق من البنية العاملية لمقياس الاستعداد اللغوى والرياضى، وأظهرت نتائج التحليل الإحصائى للبيانات ما يلى:

جدول (٤)

مؤشرات حسن المطابقة للتحليل العاملى التوكيدى لمقياس الاستعداد اللغوى والرياضى

مؤشرات المطابقة	مدى المؤشر	قيمة المؤشر	
		الاستعداد اللغوى	الاستعداد الرياضى
درجات كا٢ X2	حيث القيمة المنخفضة مع عدم الدلالة الإحصائية تشير إلى اتفاق النموذج مع البيانات	٤٠,٥٦٦	٤٩,٤
درجات الحرية		٣٣	٣٣
جذر ومتوسط درجات خطأ الاقتراب RMSEA	حيث القيمة ٠,٠٥ فأقل تعنى اتفاق جيد مع النموذج	٠,٠٤	٠,٠١
مؤشر حسن المطابقة GFI	إذا زادت القيمة عن ٠,٩ تعنى أن النموذج يتفق مع البيانات	٠,٩٧	٠,٩٧
مؤشر حسن المطابقة المصحح بدرجات الحرية AGFI	إذا زادت القيمة عن ٠,٩ تعنى أن النموذج يتفق مع البيانات	٠,٩٧	٠,٩٦
مؤشر المطابقة غير المعيارى NNFI	إذا زادت القيمة عن ٠,٩ تعنى أن النموذج يتفق مع البيانات	٠,٩٦	٠,٩٨
مؤشر المطابقة المقارنة CFI	إذا زادت القيمة عن ٠,٩ تعنى أن النموذج يتفق مع البيانات	٠,٩٦	٠,٨٩

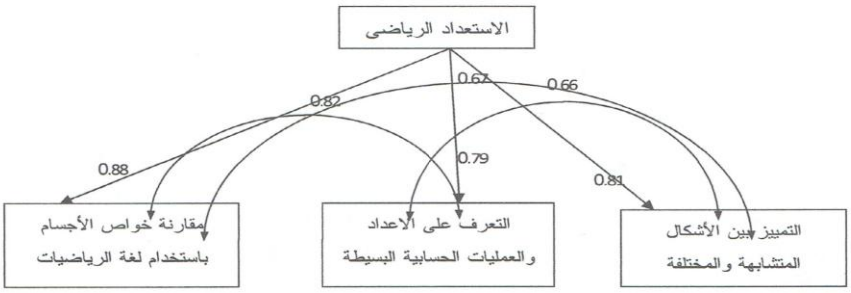
يتضح من الجدول السابق أن مؤشرات المطابقة لمكونات مقياس الاستعداد اللغوى والرياضى، تظهر حسن المطابقة للبيانات.



شكل (٢) : نتائج التحليل العاملى التوكيدى لمقياس الاستعداد اللغوى

شكل (١) : نتائج التحليل العاملى التوكيدى لمقياس الاستعداد اللغوى

يتضح من الشكل (١) ان مسارات تشبعات العوامل تحقق البنية العاملية لمقياس الاستعداد اللغوى حيث اظهرت تشبع قدره (٠,٧٧) لعامل نطق الحروف الهجائية، وتشبع قدره (٠,٧٨) لعامل تمييز اوجه التشابه والاختلاف فى رسم الحروف الهجائية، وتشبع قدره (٠,٨٢) لعامل الربط بين الصورة والكلمات الدالة عليها.



شكل (٣): نتائج التحليل العاملى التوكيدى لمقياس الاستعداد الرياضى

شكل (٢): نتائج التحليل العاملى التوكيدى لمقياس الاستعداد الرياضى

يتضح من الشكل (٢) ان مسارات تشبعات العوامل تحقق البنية العاملية لمقياس الاستعداد الرياضى

حيث اظهرت تشبع قدره (٠,٨١) لعامل التمييز بين الاشكال المتشابهة والمختلفة، وتشبع قدره (٠,٧٩) التعرف على الاعداد والعمليات الحسابية البسيطة، وتشبع قدره (٠,٨٨) لعامل مقارنة خواص الاجسام باستخدام لغة الرياضيات.

و – ثبات مقياس الاستعداد

اظهر معامل ثبات كيودر ريتشاردسون الى ان معامل ثبات القدرة الذى بلغ (٠,٩٨) ومعامل ثبات تقدير القدرة (٠,٧٩) مما يشير الى ثبات مقياس الاستعداد اللغوى. بينما بلغ معامل ثبات القدرة للاستعداد الرياضى (٠,٩٨) ومعامل ثبات تقدير القدرة (٠,٧٩) مما يشير ان مقياس الاستعداد يتميز بتقديرات ثابته للصعوبة.

٣– الاختبارات التحصيلية فى اللغة العربية والرياضيات (محكية المرجع) اعداد الباحث

تم بناء بنكين لكل من مادة الرياضيات واللغة العربية لصفوف من الاول إلى الثالث الابتدائى، حتى يتم سحب منها اختبارات مقننة محكية المرجع كاحد اسس استخدام نماذج القيمة المضافة لانها تقيس مدى اكتساب الطالب للمهارات او المفاهيم بغض النظر عن مكانته بين اقرانه.

أ– اعداد الاطار المفاهيمى ومفردات البنكين

- تم تحديد نواتج التعلم لمادتى اللغة العربية والرياضيات فى الصفوف الثلاثة بالمرحلة الابتدائية من خلال اهداف المنهج ونسبة تمثيل الاهداف وفقا لوحدات كل منهج دراسى.
- تم مراجعة الاطار المفاهيمى وجدول المواصفات للاختبارات وتحكيمها من خلال بعض موجهى المادة ومتخصصى المواد بالمركز القومى للامتحانات ومركز تطويرالمناهج.
- تم اعداد مفردات الاختبارات بحيث تغطى نواتج التعلم فى الصفوف الثلاثة فى اللغة العربية والرياضيات بالاستعانة بمجموعة من معلمى وموجهى المادتين.
- تم تحكيم المفردات من حيث الصياغة وملاءمتها لقياس الاهداف الخاص بكل صف دراسى من خلال خبراء القياس بالمركز القومى للامتحانات والتقويم التربوى وكليات التربية.

ب–تدريج البنك

تم اعداد ٦ صور اختبارية لكل مادة (اللغة العربية – الرياضيات) بواقع اختبارين (نصف دراسى اول – نصف دراسى ثانى) لكل صف دراسى من الصفوف الثلاثة الاولى بالمرحلة الايتدائية مع وجود مفردات مشتركة بين الصور الاختبارية المختلفة فى الصفوف الثلاثة لكل مادة حتى يتم تدريج اسئلة كل مادة فى بنك اسئلة واحد وتدريج واحد تبعا لصعوبة كل سؤال على متصل واحد بوحدة قياس معرفة (اللوجيت).

تم التطبيق الاستطلاعى لتدريج وتقنين الاختبارات (اللغة العربية والرياضيات) خلال العام الدراسى ٢٠١٣/٢٠١٢.

- تم استخدام مجموعة من البرامج الإحصائية في تحليل وتدريج استجابات مجموعات الدراسة على الاختبار وهى الحزمة الإحصائية SPSS وبرنامج ونستبس WINSTEPS.

من خلال التحليل الأحصائى تم حذف البيانات التامة والصفرية، والافراد غير الملائمين، والمفردات غير الصادقة (غير الملائمة لقياس الموضوعى) فى تدرجها على متصل المتغير موضع القياس (اللغة العربية- الرياضيات).

ج- صدق الملاءمة للاختبارات اللغة العربية والرياضيات

تم اتباع جميع الخطوات التى ذكرها فى مقياس الاستعداد لحساب صدق الملائمة. وقد تراوحت الصعوبات بنك مادة اللغة العربية مابين (-٥،٦٨) وحتى (٦،٠٤) بمتوسط قدره (٠،٧٦) لوجيت والذي يشير إلى المدى الواسع لصعوبات المفردات، وتراوح الخطأ المعياري لصعوبات المفردات ما بين (٠,١٤-) و(١،٨٣) الذي يشير إلى ارتفاع قيم دالة المعلومات للمفردات.

بينما تراوحت الصعوبات بنك مادة الرياضيات مابين (-١،٩١) وحتى (٤،١٩) بمتوسط قدره (٠،٧٧) لوجيت والذي يشير إلى المدى الواسع لصعوبات المفردات، وتراوح الخطأ المعيارى لصعوبات المفردات ما بين (٠،٠٥) و(٠،٣٥) الذي يشير إلى ارتفاع قيم دالة المعلومات للمفردات.

د- ثبات بنك اختبارات اللغة العربية والرياضيات

اظهر معامل ثبات كيودر ريتشاردسون الى ان معامل ثبات القدرة بلغ (٠,٩٧) مما يشير الى ثبات مفردات بنك اللغة العربية. بينما بلغ معامل ثبات القدرة لبنك الرياضيات (٠,٧٩) مما يشير ان مقياس الاستعداد يتميز بتقديرات ثابته.

هـ- سحب الصور الاختبارية لاختبارات اللغة العربية والرياضيات

بعد الانتهاء من تدريج بنك اللغة العربية والرياضيات للصفوف الثلاثة بالمرحلة الابتدائية تم حصر المفردات وتصنيفها (سهلة- متوسطة - صعبة) على أساس المتوسط والانحراف المعيارى وهو مايوضحه الجدول التالى:

جدول (٥)

مستويات تصنيف بنك اللغة العربية والرياضيات

مستوى المفردة	قيمة الصعوبة	
	من	إلى
سهلة	أصغر قيمة للصعوبة	أصغر من م−ع
متوسطة	م−ع	م+ع
صعبة	أكبر من م+ع	أكبر قيمة للصعوبة

- في ضوء استقلال القياس التي يوفرها نموذج راش تم سحب صور اختبارية متعادلة من البنك الكلي للغة العربية والرياضيات بعد أن أصبحت مفردات كل بنك على مقياس متدرج واحد حيث تم سحب الصور الاختبارية بناء على: دالة معلومات كل اختبار، والخطأ المعياري ، والمنحنى المميز، وتقدير متوسط القدرة، والمحتوى المناسب للصف الدراسى.

- تم اعداد جداول مواصفات يشمل اهداف كل مادة ومستويات الصعوبة للاسئلة وذلك لعدد صورتين اختباريتين لكل صف دراسى من الصفوف الثلاثة الاولى بالمرحلة الابتدائية فى كل مادة دراسية (اللغة العربية− الرياضيات) بواقع امتحان لنصف العام وامتحان لنهاية العام على مدار ثلاثة أعوام دراسية ٢٠١٢/٢٠١٣، و٢٠١٣/٢٠١٤، و٢٠١٤/٢٠١٥.

- تم سحب الصور الاختبارية بناء على الخطوات السابقة مع مراعاة وجود مفردات مشتركة بين جميع الصور فى جميع الصفوف فى المادة الواحدة (اللغة العربية− الرياضيات) بما يضمن تعادل القياس على مدار الدراسة الطولية وموضوعية القياس كأحد شروط استخدام نماذج القيمة المضافة.

تصميم الدراسة الطولية:

١− اختيار نماذج القيمة المضافة لتقييم فاعلية أداء المعلمين:

تم اختيار نموذج درجات الكسب gain score model (كنموذج احادى المتغيرات)، ونموذج التأثيرات المتعددة layerd effects model (كنموذج متعدد المتغيرات) لعدة أسباب:

- يسمح نموذج الكسب كنموذج مبسط بالمقارنة بين التأثيرات العشوائية لتأثيرات المعلمين ونموذج التأثيرات المتعددة كنموذج أكثر تعقيدا يضع فى اعتباره المتغيرات المصاحبة مثل استعددات الطلاب والمستوى الاجتماعى الاقتصادى عند حساب تأثيرات المعلمين باختلاف سياق التعلم من خلال قياس مدى تحقيق طلابهم نواتج التعلم على مدار ٣ سنوات دراسية.

تم استخدام هذه النماذج فى بعض الدراسات السابقة بما يسمح بالمقارنة بين النتائج لمعرفة مدى صدق هذه النماذج فى تقييم فاعلية اداء المعلمين.

٢- اعداد بنك أسئلة الاختبارات المعيارية المقننة (الرياضيات واللغة العربية) :

- إذا كانت درجات اختبارات الطلاب ستستخدم لمقارنة أداء المعلم، عندئذ يجب أن تكون الاختبارات معيارية لجميع الطلاب كي تكون المقارنة أمرًا ممكنًا. وهذا يعني، على سبيل المثال، أن التقييمات التي تتم داخل الفصل الدراسي التي يصممها ويستخدمها المعلمون بشكل مفرد لا يمكن استخدامها للتقييم العادل والموضوعى. حيث تختلف طريقة إعداد هذه الاختبارت من معلم لآخر.

- لذا استخدم فى تدريج الأسئلة واحد من أهم أساليب القياس التربوي الحديثة وهو نموذج (راش Rasch)، نظراً لما يحققه من موضوعية فى القياس تتمثل فى خطية التدرج وأحادية القياس واستقلاليته. بما أتاح الفرصة لتدريج أسئلة كل بعد من الأبعاد المكونة لمهارات الرياضيات واللغة العربية للصفوف الثلاثة من المرحلة الابتدائية.

- أن التدريج واسع المدى الذي يوفره نموذج (راش) يتيح الفرصة لقياس التغير فى تقدير قدرة الفرد (تحصيله) أو الأفراد على مدى زمني مثل قياس معدل النمو عبر الصفوف الدراسية، ومعدلاته (صلاح علام، ٢٠٠٤)، (Harris, 2011). والتى استخدمت فى القياسات المتعددة على مدار الدراسة الطولية على ميزان تدرج مشترك له وحدة قياس واحدة. والتى تعتبر أحد افتراضات استخدام نماذج القيمة المضافة.

٣- دراسة خط الاساس:

- تعتبر أبسط طريقة لحساب "درجة" القيمة المضافة (النمو)، وهى طرح قياس قاعدي baseline أو قبْلى من قياس مُلاحظ (بعْدى)، ومقارنته بمستوى تحصيله الراهن (الملاحظ observed)، ولتحديد نموه عبر الزمن. ولمعرفة إسهام كل معلم فى هذا النمو. (Jackubowski, 2008).

- ومن ثم تم تطبيق أول اختبارت لطلاب الصف الاول الابتدائى فى مادتى اللغة العربية والرياضيات بنهاية الصف الدراسى الاول فى العام الدراسى ٢٠١٣/٢٠١٢ وتعتبر تقديرات الافراد على هذا الاختبار هى خط الاساس لمستوى الطلاب فى القياسات التتبعية على مدار الدراسة الطولية.

- كما تم تطبيق مقاييس المتغيرات المصاحبة covariate (المتغيرات المستقلة التى قد يكون لها تأثير على نواتج التعلم)، والتى تشمل مقياس المستوى الاجتماعى الاقتصادى ومقياس الاستعداد ببعديه اللغوى والرياضى خلال النصف الأول من العام الدراسى ٢٠١٣/٢٠١٢ عند استخدام نموذج التأثيرات المتعددة ويعتبر هذه أيضا تحقيقا لأحد افتراضات استخدام نماذج القيمة المضافة.

٤- جمع البيانات الطولية:

تم جمع البيانات الطولية عن طريق تطبيق الاختبارت المتعادلة القياس فى مادتى الرياضيات واللغة العربية فى نهاية النصف الدر اسى الأول والثانى لنفس الطلاب بداية من العام الدراسى

٢٠١٣/٢٠١٢ (الصف الاول الابتدائى)، والعام الدراسى ٢٠١٤/٢٠١٣ (الصف الثانى الابتدائى)، والعام الدراسى ٢٠١٥/٢٠١٤ (الصف الثالث الابتدائى).

٥- تحليل البيانات:

- استخدام برنامج SPSS,20 وبرنامج SAS,9.2.

- تم الربط فى ملفات إدخال البيانات بين نتائج الطلاب وكود معلم الفصل لنفس الطلاب.

- أول خطوة للتحليل تشمل تحليل أحادى المتغيرات univariate analysis، والتى تشمل توزيع التكررات وتقديم صورة وصفية تشمل متوسط درجات مادتى الرياضيات واللغة العربية فى كل صف دراسى على مدار الثلاث سنوات.

- ثم خطوة تحليل متعدد المتغيرات multivariate analysis، وتم استخدام المعادلات الإحصائية الخاصة بنموذج درجات الكسب ونموذج التأثيرات المتعددة والتى عرضت فى الإطار النظرى والتى شملت أرقام من (4.1) إلى (4.7).

- للاجابة على الأسئلة من ١ و ٢ تم حساب معاملات الارتباط بين تقديرات فاعلية المعلمين فى مادتى الرياضيات واللغة العربية فى كل نموذج على حدة فى كل من النموذجين المختارين.

- للإجابة على الأسئلة من ٣ و ٤ تم حساب معاملات الارتباط بين تقديرات فاعلية المعلمين فى النموذجين المختارين لكل مادة على حدة.

- للإجابة على الأسئلة من ٥ و ٦ تم حساب معاملات الارتباط الرتبى لسبيرمان sperman rank corerelation، لترييب المعلمين وفقا لنتائج المادتين فى كل نموذج على حدة.

- للإجابة على الأسئلة من ٧ و ٨ تم حساب معاملات الارتباط الرتبى لسبيرمان، لترييب المعلمين فى النموذجين المختارين لكل مادة على حدة.

نتائج الدراسة ومناقشتها

اولاً: فاعلية أداء المعلمين (كما تقاس بنتائج نواتج التعلم فى الرياضيات واللغة العربية) باستخدام نماذج القيمة المضافة المختلفة

يتناول هذا الجزء الإجابة على أسئلة الدراسة من الأول إلى الثالث عن طريق اختبار العلاقة بين تقديرات فاعلية المعلمين فى اللغة العربية والرياضيات المقاسة بنواتج التعلم فى المادتين باستخدام معاملات ارتباط بيرسون بين كل من نموذج الكسب ونموذج التأثيرات المتعددة. وكذلك من خلال

نتائج النموذج الواحد فى كل مادة دراسية عبر الصفوف الدراسية المختلفة.

السؤال الأول: هل توجد علاقة ارتباطية بين تقديرات فاعلية أداء المعلمين فى كل من مادتى الرياضيات واللغة العربية (كما تقاس بنتائج تقويم نواتج التعلم لطلابهم) من خلال استخدام نموذج الكسب كأحد نماذج القيمة المضافة أحادية المتغيرات؟

جدول (٦)

مصفوفة معاملات ارتباط بيرسون بين نتائج الرياضيات واللغة العربية باستخدام نموذج الكسب (من الصف الاول الى الصف الثالث الابتدائى)

		رياضيات		اللغة العربية	
	االمادة	من الصف الاول الى الثانى الابتدائى	من الصف الثانى الى الثالث الابتدائى	من الصف الاول الى الثانى الابتدائى	من الصف الثانى الى الثالث الابتدائى
رياضيات	من الصف الأول الى الثانى الابتدائى	٠,١			
	من الصف الثانى الى الثالث الابتدائى	٠,١٣	٠,١		
اللغة العربية	من الصف الأول الى الثانى الابتدائى	٠,٤٣	-٠,٠٩	٠,١	
	من الصف الثانى الى الثالث الابتدائى	-٠,٠٧	٠,١٥	٠,٤٠	٠,١

يظهر الجدول رقم (٦) الارتباطات بين تقديرات فاعلية اداء المعلمين باستخدام نموذج الكسب من خلال تقديرات مادتى اللغة العربية والرياضيات، حيث كانت معامل الارتباط بين نتائج المادتين ٠,٤٣ بين الصف الاول والثانى الابتدائى، بينما كان معامل الارتباط ٠,١٥ فقط بين الصف الثانى والثالث الابتدائى.

السؤال الثانى: هل توجد علاقة ارتباطية بين تقديرات فاعلية اداء المعلمين فى كل من مادتى الرياضيات واللغة العربية (كما تقاس بنتائج تقويم نواتج التعلم لطلابهم) من خلال استخدام نموذج التأثيرات المتعددة كأحد نماذج القيمة المضافة متعددة المتغيرات؟

جدول (٧)

مصفوفة معاملات ارتباط بيرسون بين نتائج الرياضيات واللغة العربية باستخدام نموذج التأثيرات المتعددة (من الصف الاول الى الصف الثالث الابتدائى)

		رياضيات			اللغة العربية		
المادة	الصف الدراسى	الصف الاول	الصف الثانى	الصف الثالث	الصف الاول	الصف الثانى	الصف الثالث
رياضيات	الصف الأول	٠,١					
	الصف الثانى	-٠,٦٩	٠,١				
	الصف الثالث	-٠,٣٧	٠,٢٨	٠,١			
اللغة العربية	الصف الاول	٠,٧٧	-٠,١٩	-٠,٢٥	٠,١		
	الصف الثانى	-٠,٢٣	٠,٥٠	٠,٠٧	٠,١٤	٠,١	
	الصف الثالث	-٠,١٩	٠,١٤	٠,٢١	-٠,٣١	٠,١٨	٠,١

يظهر الجدول رقم (٧) الارتباطات بين تقديرات فاعلية أداء المعلمين باستخدام نموذج التأثيرات المتعددة من خلال تقديرات مادتى اللغة العربية والرياضيات، حيث كانت معامل الارتباط بين نتائج المادتين ٠,٧٧ فى الصف الأول وكانت ٠,٥ فى الصف الثانى، بينما كانت ٠,٢١ فقط فى الصف الثالث.

السؤال الثالث: هل توجد علاقة ارتباطية بين تقديرات نموذج الكسب (كأحد نماذج القيمة المضافة أحادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات) فى تقييم فاعلية أداء المعلمين كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادة الرياضيات؟

جدول (٨)

مصفوفة معاملات ارتباط بيرسون بين نتائج نموذج الكسب ونموذج التاثيرات المتعددة فى مادة الرياضيات (من الصف الأول الى الصف الثالث الابتدائى)

		نموذج الكسب		نموذج التأثيرات المتعددة	
نموذج القيمة المضافة المستخدام	الصف الدراسى	من الصف الاول الى الثانى الابتدائى	من الصف الثانى الى الثالث الابتدائى	الصف الثانى الابتدائى	الصف الثالث الابتدائى
نموذج الكسب	من الصف الأول الى الثانى الابتدائى	٠,١			
	من الصف الثانى الى الثالث الابتدائى	٠,١٣	٠,١		
نموذج التأثيرات المتعددة	الصف الثانى الابتدائى	٠٨٤	٠,٢١	٠,١	
	الصف الثالث الابتدائى	٠,١٣	٠,٩٦	٠,٢٨	٠,١

يظهر الجدول رقم (٨) الارتباطات المقارنة بين نموذج الكسب و نموذج التأثيرات المتعددة لتقديرات فاعلية أداء المعلمين من خلال تقديرات مادة الرياضيات، حيث كانت معامل الارتباط بين نتائج النموذجين ٠,٨٤ فى الصف الثانى وكانت ٠,٩٦ فى الصف الثالث.

السؤال الرابع: هل توجد علاقة ارتباطية بين تقديرات نموذج الكسب (كأحد نماذج القيمة المضافة أحادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات) فى تقييم فاعلية أداء المعلمين كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادة اللغة العربية؟

جدول (٩)

مصفوفة معاملات ارتباط بيرسون بين نتائج نموذج الكسب ونموذج التاثيرات المتعددة فى مادة اللغة العربية (من الصف الأول الى الصف الثالث الابتدائى)

		نموذج الكسب		نموذج التأثيرات المتعددة	
		من الصف الاول الى الثانى الابتدائى	من الصف الثانى الى الثالث الابتدائى	الصف الثانى الابتدائى	الصف الثالث الابتدائى
نموذج الكسب	من الصف الأول الى الثانى الابتدائى	٠,١			
	من الصف الثانى الى الثالث الابتدائى	-٠,٤٠	٠,١		
نموذج التأثيرات المتعددة	الصف الثانى الابتدائى	٠,٧٤	-٠,١٢	٠,١	
	الصف الثالث الابتدائى	-٠,١٨	٠,٩٢	٠,١٨	٠,١

يظهر الجدول رقم (٩) الارتباطات المقارنة بين نموذج الكسب و نموذج التاثيرات المتعددة لتقديرات فاعلية اداء المعلمين من خلال تقديرات مادة اللغة العربية، حيث كانت معامل الارتباط بين نتائج النموذجين ٠,٧٤ فى الصف الثانى وكانت ٠,٩٢ فى الصف الثالث.

ثانياً: ترتيب المعلمين

هذا الجزء يتناول الاجابة على أسئلة الدراسة من الخامس الى الثامن من خلال ترتيب المعلمين من خلال قياس نواتج التعلم فى الرياضيات واللغة العربية وتقديرات نموذج الكسب والتأثيرات المتعددة باستخدام الارتباط الرتبى لسبيرمان.

السؤال الخامس: هل توجد علاقة ارتباطية بين ترتيب المعلمين وفقا لتقديرات فاعلية ادائهم (كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادتى الرياضيات واللغة العربية) من خلال استخدام نموذج الكسب كأحد نماذج القيمة المضافة أحادية المتغيرات؟

جدول (١٠)

مصفوفة معاملات الارتباط الرتبى لبيرسون بين نتائج الرياضيات واللغة العربية باستخدام نموذج الكسب (من الصف الأول الى الصف الثالث الابتدائى)

		رياضيات		اللغة العربية	
		من الصف الأول الى الثانى الابتدائى	من الصف الثانى الى الثالث الابتدائى	من الصف الأول الى الثانى الابتدائى	من الصف الثانى الى الثالث الابتدائى
رياضيات	من الصف الأول الى الثانى الابتدائى	٠,١			
	من الصف الثانى الى الثالث الابتدائى	-٠,٠١	٠,١		
اللغة العربية	من الصف الأول الى الثانى الابتدائى	٠,٥١	-٠,٠٥	٠,١	
	من الصف الثانى الى الثالث الابتدائى	-٠,٠٤	٠,٠٧	-٠,٢٤	٠,١

يظهر الجدول رقم (١٠) معاملات الارتباط الرتبى بين ترتب المعلمين وفقا لتقديرات فاعلية ادائهم من خلال قياس نواتج التعلم فى مادتى الرياضيات واللغة العربية باستخدام نموذج الكسب، حيث كانت معامل الارتباط ٠,٥١ بين الصف الأول والثانى الابتدائى، بينما كان معامل الارتباط ٠,٧ بين الصف الثالث والرابع الابتدائى.

السؤال السادس: هل توجد علاقة ارتباطية بين ترتيب المعلمين وفقا لتقديرات فاعلية ادائهم (كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادتى الرياضيات واللغة العربية) من خلال استخدام نموذج التأثيرات المتعددة كأحد نماذج القيمة المضافة متعددة المتغيرات؟

جدول (١١)

مصفوفة معاملات الارتباط الرتبى لبيرسون بين نتائج الرياضيات واللغة العربية باستخدام التأثيرات المتعددة (من الصف الأول الى الصف الثالث الابتدائى)

		رياضيات			اللغة العربية		
		الصف الأول	الصف الثانى	الصف الثالث	الصف الأول	الصف الثانى	الصف الثالث
رياضيات	الصف الأول	٠,١					
	الصف الثانى	-٠,٧٢	٠,١				
	الصف الثالث	-٠,٣٥	٠,١٨	٠,١			
اللغة العربية	الصف الأول	٠,٧٤	-٠,٢٨	-٠,٢٦	٠,١		
	الصف الثانى	-٠,٢٧	٠,٤٦	٠,٠٦	-٠,١٧	٠,١	
	الصف الثالث	-٠,٢٢	٠,١٧	٠,٢٢	-٠,٣٧	٠,٢٩	٠,١

يظهر الجدول رقم (١١) معاملات الارتباط الرتبى بين ترتب المعلمين وفقا لتقديرات فاعلية

ادائهم من خلال قياس نواتج التعلم فى مادتى الرياضيات واللغة العربية باستخدام نموذج التأثيرات المتعددة، حيث كانت معامل الارتباط ٠,٧٤ فى الصف الاول وكانت ٠,٤٦ فى الصف الثانى، بينما كانت ٠,٢٢ فقط فى الصف الثالث.

السؤال السابع: هل توجد علاقة ارتباطية بين ترتيب المعلمين وفقا لتقديرات نموذج الكسب (كأحد نماذج القيمة المضافة احادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات) فى ضوء تقييم فاعلية ادائهم كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادة الرياضيات؟

جدول (١٢)

مصفوفة معاملات الارتباط الرتبى لبيرسون بين نتائج نموذج الكسب ونموذج التأثيرات المتعددة فى مادة الرياضيات (من الصف الأول الى الصف الثالث الابتدائى)

		نموذج الكسب		نموذج التأثيرات المتعددة	
		من الصف الاول الى الثانى الابتدائى	من الصف الثانى الى الثالث الابتدائى	الصف الثانى الابتدائى	الصف الثالث الابتدائى
نموذج الكسب	من الصف الأول الى الثانى الابتدائى	٠,١			
	من الصف الثانى الى الثالث الابتدائى	-٠,٠١	٠,١		
نموذج التأثيرات المتعددة	الصف الثانى الابتدائى	٠٧٩	٠,١٠	٠,١	
	الصف الثالث الابتدائى	٠,٠١	٠,٩١	٠,١٨	٠,١

يظهر الجدول رقم (١٢) معاملات الارتباط الرتبى بين ترتب المعلمين وفقا لتقديرات فاعلية ادائهم من خلال استخدام كل من نموذج الكسب و نموذج التأثيرات المتعددة من خلال قياس نواتج التعلم فى مادة الرياضيات، حيث كانت معامل الارتباط ٠,٧٩ فى الصف الثانى، بينما كانت ٠,٩١ فقط فى الصف الثالث.

السؤال الثامن: هل توجد علاقة ارتباطية بين ترتيب المعلمين وفقا لتقديرات نموذج الكسب (كأحد نماذج القيمة المضافة احادية المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج القيمة المضافة متعددة المتغيرات) فى ضوء تقييم فاعلية ادائهم كما تقاس بنتائج تقويم نواتج التعلم لطلابهم فى مادة اللغة العربية؟

جدول (١٣)

مصفوفة معاملات الارتباط الرتبى لبيرسون بين نتائج نموذج الكسب ونموذج التاثيرات المتعددة فى مادة اللغة العربية (من الصف الأول الى الصف الثالث الابتدائى)

		نموذج الكسب		نموذج التأثيرات المتعددة	
		من الصف الأول الى الثانى الابتدائى	من الصف الثانى الى الثالث الابتدائى	الصف الثانى الابتدائى	الصف الثالث الابتدائى
نموذج الكسب	من الصف الأول الى الثانى الابتدائى	٠,١			
	من الصف الثانى الى الثالث الابتدائى	-٠,٠١	٠,١		
نموذج التأثيرات المتعددة	الصف الثانى الابتدائى	٠٧٩	٠,١٠	٠,١	
	الصف الثالث الابتدائى	٠,٠١	٠,٩١	٠,١٨	٠,١

يظهر الجدول رقم (١٣) معاملات الارتباط الرتبى بين ترتب المعلمين وفقا لتقديرات فاعلية ادائهم من خلال استخدام كل من نموذج الكسب و نموذج التأثيرات المتعددة من خلال قياس نواتج التعلم فى مادة اللغة العربية، حيث كانت معامل الارتباط ٠,٧٦ فى الصف الثانى وكانت ٠,٨٣ فى الصف الثالث.

ملخص النتائج

يقدم الجدول رقم (١٤) نتائج فاعلية أداء المعلمين من خلال قياس نواتج التعلم فى العلوم والرياضيات واضعين فى الاعتبار أهم افتراضات نماذج القيمة المضافة (مراعاة الدرجات المفقودة فى القياسات المتعددة طوال فترة الدراسة الطولية – استخدام الاختبارات المعيارية المقننة – قياس استعداد الطلاب للدراسة – قياس المستوى الاجتماعى الاقتصادى للمتعلمين) من خلال البيانات الطولية على مدار ثلاث سنوات، مع الربط بين قياس نواتج التعلم الطلاب وفاعلية اداء المعلم (معلم الفصل من الصف الأول الى الصف الثالث الابتدائى). وذلك باستخدام نموذج الكسب (كأحد نماذج أحادى المتغيرات) ونموذج التأثيرات المتعددة (كأحد نماذج متعددة المتغيرات).

جدول (١٤)

ملخص نتائج معاملات الارتباط ومعاملات الاتباط الرتبى لفاعلية اداء المعلمين باستخدام نموذج الكسب ونموذج التاثيرات المتعددة للقيمة المضافة

العلاقة بين نتائج نواتج التعلم فى الرياضيات واللغة العربية			العلاقة بين نتائج نموذج الكسب ونموذج التأثيرات المتعددة كما تقاس بنواتج التعلم		
	معاملات الارتباط	معاملات الارتباط الرتبى		معاملات الارتباط	معاملات الارتباط الرتبى
نموذج الكسب			الرياضيات		
من الصف الأول الى الثانى الابتدائى	٠,٤٣	٠,٥١	الصف الثانى	٠,٨٤	٠,٧٩
من الصف الثانى الى الثالث الابتدائى	٠,١٥	٠,٠٧	الصف الثالث	٠,٩٦	٠,٩٦
نموذج التاثيرات المتعددة			اللغة العربية		
الصف الأول	٠,٧٧	٠,٧٤			
الصف الثانى	٠,٥٠	٠,٤٦	الصف الثانى	٠,٧٤	٠,٧٦
الصف الثالث	٠,٢١	٠,٢٢	الصف الثالث	٠,٩٢	٠,٨٣

ويمكن أن تلخص نتائج الدراسة الحالية فيما يلى:

٣- ان معاملات الارتباط بين نتائج الرياضيات واللغة العربية منخفضة نسبياً فى الصف الثانى والثالث الابتدائى باستخدام نماذج الكسب، بينما كان معامل الارتباط منخفض فى الصف الثالث الابتدائى فقط عندما استخدم نموذج التأثيرات المتعددة.

٤- اظهرت العلاقة بين نتائج استخدام كل من نموذج الكسب ونموذج التأثيرات المتعددة كمؤشر لفاعلية أداء المعلم وترتيب المعلمين وفقا لأداءتهم (من خلال نتائج قياس نواتج التعلم فى مادتى الرياضيات واللغة العربية)، كانت العلاقة الارتباطية بين نتائج النموذجين تتروح ما بين مرتفعة إلى متوسطة فى جميع الصفوف.

ويتضح من نتائج الدراسة ما يلى:

١- عند تقييم فاعلية أداء المعلم من خلال قياس نواتج التعلم فى الرياضيات واللغة العربية، اظهرت النتائج معاملات ارتباط منخفضة او متوسطة فى الصفوف المختلفة مما يظهر أن استخدام نتائج قياس نواتج التعلم فى المواد المختلفة أكثر حساسية فى قياس فاعلية أداء المعلمين والقدرة على الحكم على أداءتهم وان كانت أقل او أكثر فاعلية، وذلك عند استخدام

نموذج الكسب كأحد نماذج أحادية المتغيرات للقيمة المضافة. واتفقت النتائج تقريبا عند استخدام نموذج التأثيرات المتعددة أن معاملات الارتباط تكون منخفضة فى الصفوف الأعلى. وهو ما يتفق مع دراسة شيمتز Schmitz 2007، التى أظهرت نتائج مشابهة من حيث حساسية استخدام نواتج التعلم فى تقييم فاعلية المعلم حيث أظهرت الدراسة معاملات ارتباط منخفضة بين نتائج الرياضيات ومهارات الكتابة فى دراسة طولية لمدة عامين. وإن اختلفت النتائج ايضاً مع توقعات ماكفرى واخرون (٢٠٠٣)، بأن الفروق بين نتائج نماذج القيمة المضافة المختلفة قد تظهر مع الدراسات الطولية التى تستغرق على الأقل ٣ سنوات اى كما تم فى الدراسة الحالية.

٢- عند تقييم فاعلية أداء المعلم من خلال استخدام أكثر من نموذج من نماذج القيمة المضافة، وجدت الدراسة الحالية علاقة ارتباطية مرتفعة بين نتائج تقدير النموذجين فى الرياضيات واللغة العربية، وهو ما يتفق مع دراسة تكوى واخرون Tekwe et.al. 2004، التى توصلت من خلال دراسة العلاقة بين نتائج استخدام نماذج القيمة المضافة المختلفة علاقة مرتفعة كما تقاس بنواتج التعلم فى الرياضيات إلى أن العلاقة بين نتائج النماذج المختلفة تتراوح ما بين ٠,٩٦ إلى ٠,٩٨. وهو ما يتفق مع دراسة بريجز وويكس Briggs & Weeks 2011، التى اختبرت فاعلية المعلم وفاعلية المدرسة باستخدام ثلاثة نماذج للقيمة المضافة ووجد العلاقة بينهم تتراوح ما بين ٠,٤٧ إلى ٠,٩٣. اى أن العلاقة بين نتائج النماذج المتعددة متشابه إلى حد كبير.

وربما تشير هذه العلاقات المرتفعة الى أن معدل الكسب فى تقديرات فاعلية المعلم (نموذج الكسب) يرتبط بقوة مع متوسط تقديرات فاعلية أداء المعلم (نموذج التأثيرات المتعددة)، من ثم لم تظهر فروق بين النموذجين عند استخدام نواتج التعلم لتقييم فاعلية أداء المعلم.

٣- تتفق نتائج هذه الدراسة مع الدراسات السابقة فى أن تقييم فاعلية اداء المعلم من خلال قياس نواتج التعلم قد يظهر الفروق بين اداء المعلمين وترتيبهم وفقا لتحقيق طلابهم نواتج التعلم ودرجة نموأداء الطلاب عبر فترة زمنية معينة، مع حساب معدل الكسب بالنسبة إلى متوسط الكسب على مستوى المادة الدراسية او حتى مستوى الصف الدراسى.

٤- نتيجة لعدم وجود فروق تقريبا بين نتائج نماذج القيمة المضافة المختلفة، قد يكون استخدام النماذج الاكثر تعقيدا مثل نموذج التأثيرات المتعددة افضل لسبب واحد وهو انه يضع فى اعتباره المتغيرات المصاحبة مثل المستوى الاجتماعى واستعداد الطلاب وقدراتهم مما يرجع نتائج قياس نواتج التعلم كمؤشر للقيمة المضافة للمعلم إلى فاعلية ادءا المعلم داخل الفصل الدراسى دون اى عوامل اخرى، بشرط استخدام اختبارات معيارية مقننة فى اطار نظرية

الاستجابة للمفردة التى تظهر درجة الطالب وقدرته الفعلية عند اداءه على الاختبار، من ثم عدم تحيز النتائج عند تقييم فاعلية أداء المعلم.

٥- اظهرت النتائج ايضا اهمية استخدام نماذج القيمة المضافة للمعلم من خلال قياس نواتج التعلم فى المواد الدراسية المختلفة من خلال البيانات الطولية على مدار عامين على الأقل وحساب معدل الكسب السنوى كأحد الوسائل المهمة فى تقييم فاعلية اداء المعلم.

التوصيات

- ضرورة الاهتمام بتقييم اداء المعلمين خلال فترة زمنية طويلة وعد الاقتصار على التقييم اليومى او السنوى فقط، حيث يظهر التقييم على فترة زمنية اطول جهد المعلمين فى الاداء وتحسن اداءهم.
- وضع تحقق نواتج التعلم للطلاب كمحك اساسى لتقييم اداء المعلمين ودرجة التحسن التى حققها معهم المعلم.

بحوث مقترحة

- استخدام مدخل القيمة المضافة فى تقييم اثر التدريب وبرامج التنيمية المهنية للمعلمين.
- دارسة اثر متغيرات البيئة المدرسية فى تقويم نواتج التعلم باستخدام نماذج القيمة المضافة.
- ادراك اداء المعلمين لاستخدام مداخل القيمة المضافة لتقييم فاعلية اداءهم.
- تقويم فاعلية برامج مدارس المتفوقين ومدارس التكنولوجيا باستخدام نماذج القيمة المضافة.

المراجـــــــع :

المراجع العربية:

١. **أحمد زكى صالح (١٩٩٢).** علم النفس التربوى، الطبعة الرابعة عشر، النهضة المصرية. **القاهرة.**

٢. **أمينة محمد كاظم (١٩٨٨).** دراسة نظرية نقدية حول القياس الموضوعي للسلوك (نموذج راش)، سلسلة الكتب المتخصصة. الكويت: مؤسسة الكويت للتقدم العلمي.

٣. **أمينة محمد كاظم (١٩٩٥).** اتجاهات معاصرة فى بناء بنوك الاسئلة. القاهرة: جامعة عين شمس (ط٢).

٤. **أنور محمد الشرقاوى (١٩٩٦).** الاختبارات المرجعية إلى محك، وسائل جديدة فى القياس النفسى والتربوى، فى أنور الشرقاوى وآخرون، اتجاهات معاصرة فى القياس والتقويم النفسى والتربوى. القاهرة: الانجلو المصرية.

٥. **صلاح الدين محمود علام (٢٠٠٤).** التقويم التربوى البديل ، أسسه النظرية، والمنهجية وتطبيقاته الميدانية. القاهرة: دار الفكر العربى.

٦. **مجدى قاسم، وصلاح الدين علام، وهشام الحسينى، وهشام جاد الرب (٢٠١١).** جودة التعليم فى ضوء القيمة المضافة. القاهرة: دار الفكر العربى.

المراجع الاجنبية:

7. **Allal, L., Mottier Lopez, L. (2005).** Formative assessment of learning: A review of publication in french. Formative assessment – improving learning in secondary classrooms, PP. 241-264. Paris: OECD.
8. **Ballou, D., Sanders, W., & Wright, P. (2003).** Controlling for student background in value-added assessment of teachers. Journal of educational and behavioral statistics.
9. **Braun, H.I. (2006).** Background paper: the use of value added models for school improvement. Paris: OECD.
10. **Braun, H.I. (2005).** Value-added modeling; what does due diligence require?. In., R. Lissitz (eds), value added models in education: theory and applications. Maple Grove, Minnesota: JAM press.
11. **Braun, H.I. (2005).** Using student progress to evaluate teachers: a primer on value-added models. Educational tesing service (ETS), policy information center. From: www.ets.org/research/pic.
12. **Bresciani, M., J., (2006).** Outcomes-based academic and co-curricular program review. Vriginia : stylus publishing.
13. **Burger, J.M., Perry, B., & Eagle, M., (2009).** Aggregating and analyzing classroom achievement data supplemental to external testes. In., T.J. Kowalski & T.J. Lasley ll. New York: Routledge. (ch. 20), pp. 317-343.
14. **Carlin, B.P. & Louis. T.A. (2000).** Bayes and empirical bayes methods for data analysis. (2nd ed. New York: chapman & Hall.
15. **Childs, G. & McKay, M. (2001).** Boys starting school disadvantaged: implications from teacher's ratings of behaviour and achievement in the first two years. British Journal of Educational Psychology 71, 303-314
16. **Coleman, J.S, Campbell, T.E,& Kilgore, S.B, (1982**). High achievement: public and other private schools compared, New York.
17. **Crawford, V., M., Schlader, M., Penuel, W., & Toyama, Y, (2008).** Supporting the art of teaching in a data – rich, high- performance learning environment. In., Ellen B. Mandinch, & Margert, Honey (eds), data – driven school improvement: linking data and learning (ch. 7, pp. 109–129). New York, teachers college, Columbia university, inc.
18. **Cromy., A.,& Hanson,M. (2000).** An exploratory analysis of school – based student assessment systems. Oak brook, il:north central regional educational laboratory (NCREL).
19. **Crundwell, R.M. (2005).** Alternative stratigies for large scale student

assessment in canada: is value-added assessment one possible answer. Canadian journal of educational administration and policy, may 24.

20. **Datnow, A., Park, V., (2009).** School system strategies for supporting data use. In., T.J. Kowalski & T.J. Lasley ll. New York: Routledge. (ch. 12), pp. 191-206.
21. **Deb Roy, S. & Bates. D. (2003).** Computational methods for multiple level linear mixed-effects models. University of Wisconsin-madison, department of statistics technical report no. 1076. http://www.stat.wisc.edu/bates/reports/TR1076_letter.pdf.
22. **Doran, H. C. and J. Cohen. (2005).** The confounding effects of linking bias on gains estimated from value added models. In., R. Lissitz (eds), value added models in education: theory and applications. Maple Grove, Minnesota: JAM press.
23. **Doran, H. C. and L. T. Izumi. (2004).** Putting education to the test: a value added model for California. San Francisco: pacific research institute.
24. **Drury, D., & Doran, H. (2003).** The value of value- added analysis. Policy research brief, national school boards association (NSBA). From: www.nsba.org.
25. **Evergreen Freedom F oundation (2003).** Value added assessment. School director handbook.
26. **Goldhaber, D.D. & Brewer, D.J. (2000).** Does teacher certification matter? High school teacher certification ststus and student achievement. Educational evaluation and policy analysis, 22, 129-145.
27. **Haegeland, T. and Kirkeboen, L. (2008).** School performance and value- added indicators – what is important of controlling for socioeconomic background?: a background report for the OECD project on the development of value-added models in education system.
28. **Hamilton,L., Halverson, R. & Wayman, J. (2009).** Using student achievement data to support instructional decision making (NCEE 2009- 4067). Washington, DC: national center for education evaluation and regional assistance, institute of education science, u.s. department of education. Retrieved From,:http://ies.ed.gov/ncee/wwc/publications/practiceguides/.
29. **Harris, D., N. (2011).** Value – added measures in education: what every educator needs to know. Masssachusetts, Cambridge, Harvard education press.
30. **Hershberg, H. (2005).** Value-added assessment: powerful diagnostics to improve instruction and promote student achievement. American association of schoo administrators, women administrators conference, "leadership in a time of change.
31. **Hill, R., Gong. B., Marion, S., and DePascale, C. (2005).** Using value added tables to explicity value sugent growth. http://

www.nciea.org/cgibin/pubspage.cgi?sortby=pub_data,accessed 10 January 2006.

32. **Jackubowski, M. (2008).** Implementing value-added models of school assessment. RSCAS working paper 2008/06, European university institute.
33. **Kelly, S., (2012).** Assessing teacher quality. London: teacher college press.
34. **Kupermintz, H, (2002).** Teacher effects as a measure of teacher effectiveness: construct validity considerations in TVAAS (Tennessee value-added assessment system). Los Angeles: center for the study of evaluation, university of California epidemiologic data. Joint statistical meeting, august 3,2003,san Francisco, California.
35. **Lockwood, J., Louis, T., & McCaffery, D. (2002).** Uncertainty in rank estimation: implication for value-added modeling accountability systems. Journal of educational and behavioral statistics, 27(3), 255-270.
36. **Mandinch, E., B., & Honey, M. (2008).** Data- driven decision making: an introduction. In., Ellen B. Mandinch, & Margert, Honey (eds), data – driven school improvement: linking data and learning (ch. 1, pp. 1–9). New York, teachers college, Columbia university, inc.
37. **McCaffrey, D.F., Lockwood, J., Kortez, D., Louis, T.A., & Hamilton, l. (2003)**. evaluating value-added models for teacher accountability. Santa monica: RAND corporation.
38. **McCaffrey, D.F., Lockwood, J., Kortez, D., Louis, T.A., & Hamilton, l. (2003).** Model for value-added modeling of teacher effects. Journal of educational and behavioral statistics.
39. **Mendro, R., Jordan, H., Gomez, e., Anderson, M., & Bembry, K. (1998)**.an application of multiple linear regression in determining longitudinal teacher effectiveness. Paper presented at the 1998 annual meeting of the AERA, san diego, CA.
40. **Olson, L. (2002).** Education Scholars Finding New Value. in Student Test Data. Education week, 22 (12), 1-14.
41. **Organization for economic co-operation and development . (2005).** Formative assessment of learning: A review of publication in french. Formative assessment – improving learning in secondary classrooms. Paris: OECD.
42. **Raudenbush & Bryk, (2002).** Ahierarchical model for studying scholl effects. Sociology of education, 59, 1-17.
43. **Rivers, j.C. (1999).** The impact of teacher effect on student math competency achievement, dissertation, university of Tennessee, Knoxville. Ann arbor, MI: university microfilms international, 9959317, 2000.
44. **Rivikin, S.G. (2007).** Value –added analysis and education policy. national center for analysis of longitudinal data in education research. From: http//www.uipress.org.
45. **Rivikin, S.G., Hanushek, E.A., & kain, J.F. (2000).** Teachers, schools,

and academic achievement. Cambridge, MA: national bureau of economic research, NBER working paper, w6691.

46. **Rowan, B., Correnti, R., & Miller, R.J. (2002).** Wate large-scale survey research tells us about teacher effects on students achievements: insights from the prospects study of elementary schools. Teachers college record, 104, 1525-1567.
47. **Sanders, W. (2000).** Value – Added Assessment from Student Achievement Data: Opportunities and Hurdles. Journal of Personnel Evaluation in Education, 14 (4), 329 – 339.
48. **Sanders, W., Saxion, A. M., Schneider, J. G., Dearden, B. L., & Horn. S. (2002).** Effects of building change on indicators of student academic growth. From: http/www.mdk12.org/practices/ensure/tva/buliding_change_effects.html.
49. **Sanders, W. & Horn. S. (1998).** Research finding from the Tennessee value- added assessment system (TVAAS) database: implications for educational evaluation and research. Journal of personal evaluation in education. 12(3), 247-256.
50. **Sanders, W. & Horn, S. (1998).** The Tennessee value-added assessment system: a quantitative outcomes-based approach to educational assessment. In., J. Millman (ed.), grading teacher, grading school: is student achievement a valid evaluation measure? Thousand oaks, CA: crowin press, Inc., 137-162.
51. **Sanders, W. & Rivers, J.C. (1996).** Cumulative and residual effects of teachers on future student academic achievement. Knoxville, TN: university of Tennessee value-added research center.
52. **Stiggins, R., (2005).** Assessment for learning. Paper presented in ETS/training conference. USA, Portland, septmber, 2005.
53. **Tekwe, C., Carter, R., Ma, C., Algina, J., Lucas, M., Roth, J. (2004).** An Empirical Comparison of Statistical Models for Value. "Added Assessment of School Performance". Journal of Educational and Behavioral Statistics, 29 (1), 11 – 36.
54. **Thum, Y.M. (2003).** Measuring progress towarda a goal: estimating teacher productivity using a multivariate multilevel model for value-added analysis. From: http://nces.ed.gov/nationsreportcard/neapdata/getdata.asp.
55. **Tucker, D., T., & Strong , J. (2005)** .Linking teacher evaluation and student learning. USA: association for supervision and curriculum development.
56. **Wainer, H.(2004).** Introducation to Value-Added Assessment. Journal of Educational and Behavioral Statistics, 29, 1 – 3.
57. **Wohlstetter,P., Datnow, A. & Park, V. (2008).** Creating a system for data – driven decision – making: applying the principal- agent framework. School effectiveness and school improvement, V. 19, PP. 239-259.

58. **Wright, S.p., horn, S.p., & Sanders, W.L. (2006).** Measurement of academic growth of individual students toward variable and meaningful academic standards. In., R. lissitz (eds), longitudinal and value-added models of student performance (pp. 385-406). Maple Grove, MN: JAM press.
59. **Wright, S.p., horn, S.p., & Sanders, W.L. (1997).** Teacher and classroom context effects on student achievement: implication for teacher evaluation. Journal of personal evaluation in education, 11, 57-67.
60. **Zupanc, D., Urank, M., Bren, M. (2009).** Varibality analysis for effectiveness and improvement in classrooms and schools in upper secondary education in slovania: Assessment of/for learning analytic tool. School effectiveness and school improvement, V. 40, N. 1, PP. 89-122.

Modeling longitudinal data to evaluate teachers' Performance Effectiveness using added value

Dr. Hesham Habib Elhussieny

Assistant professor of educational psychology

National center for examination and educational evaluation

The purpose of this study was to examine two Value Added Model specifications to understand similarities and differences in teachers effect results. Specifically, this study correlated value added teachers effect estimates, which were derived from two model specifications (Gain Score and Layered Effects) and two outcome measures (mathematics and reading). Next, the teacher rankings were compared and correlated using the same models and outcome measures (or outcomes). Conducting these analyses (1) demonstrated how stable the value added teacher effect estimates were when the models and the outcomes were altered, (2) determined whether these models could be used interchangeably to compare teachers, This study used longitudinal data collected from three years, which included test scores of students in elementary school (Grades 1 through 3).

Overall, the findings in this study indicated that the teachers effect estimates and teachers rankings were more sensitive to outcome measures than they were to model specifications. The mathematics and reading correlations from the Gain Score Model for teacher effects and teachers rankings were low. Correlations between the Gain Score Model and the Layered Effects Model from mathematics and Reading were moderate to high for each of the grades.